

Springer Proceedings in Mathematics & Statistics

Ashokkumar Patel
Nishtha Kesswani
Madhusudhan Mishra
Preetisudha Meher *Editors*

Advances in Machine Learning and Big Data Analytics I

ICMLBDA 2023, NIT Arunachal Pradesh,
India, May 29-30

 Springer

Springer Proceedings in Mathematics & Statistics

Volume 441

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including data science, operations research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

Ashokkumar Patel • Nishtha Kesswani •
Madhusudhan Mishra • Preetisudha Meher
Editors

Advances in Machine Learning and Big Data Analytics I

ICMLBDA 2023, NIT Arunachal Pradesh,
India, May 29-30

 Springer

Editors

Ashokkumar Patel
Computer & Information Science
University of Massachusetts Dartmouth
Dartmouth, MA, USA

Nishtha Kesswani
Department of Computer Science
Central University of Rajasthan
Ajmer, Rajasthan, India

Madhusudhan Mishra
Electronics & Communication Engineering
North Eastern Regional Institute of Science
Arunachal Pradesh, Arunachal Pradesh,
India

Preetisudha Meher
Department of Electronics and
Communication Engineering
National Institute of Technology
Arunachal Pradesh, Arunachal Pradesh,
India

International Conference on Machine Learning and Big Data Analytics

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-031-51337-4 ISBN 978-3-031-51338-1 (eBook)
<https://doi.org/10.1007/978-3-031-51338-1>

Mathematics Subject Classification: 68T09, 94A16, 62R07

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

In today's data-driven world, machine learning (ML) and big data analytics have become pivotal forces driving innovation and decision-making across various industries. "Key Trends Shaping the Future of ML and Big Data Analytics" offers a comprehensive exploration of the transformative shifts and emerging trends in these domains. From the evolution of deep learning algorithms to the growing importance of ethical AI, this book provides a roadmap for understanding how ML and big data analytics are shaping our future. It brings together leading experts, offering valuable insights and strategies to navigate the dynamic landscape of data science, ensuring that readers stay at the forefront of this ever-evolving field.

Dartmouth, MA, USA
Ajmer, Rajasthan, India
Arunachal Pradesh, India
Arunachal Pradesh, India

Ashokkumar Patel
Nishtha Kesswani
Madhusudhan Mishra
Preetisudha Meher

Organization

Program Committee Chairs

Kesswani, Nishtha	Department of Computer Science, Central University of Rajasthan, Ajmer, Rajasthan, India
Mishra, Madhusudhan	NERIST, Electronics and Communication Engineering, Nirjuli, India
Sambana, Bosubabu	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India

Program Committee Members

Kesswani, Nishtha	Department of Computer Science, Central University of Rajasthan, Ajmer, Rajasthan, India
Ashiq, Abdul	Sri Eshwar College of Engineering, Information Technology, Coimbatore, India
Bokka, Yugandhar	
Bada, Rajasekhar Reddy	New Horizon College of Engineering, Computer Science and Engineering, Bangalore, India
Bandi, Vamsi	Madanapalle Institute of Technology and Science, Artificial Intelligence and Data Science, Madanapalle, India
Bano, Shahana	Robert Gordon University, School of Computing, Aberdeen, UK
Bendi, Dr. Nageswara Rao	Lendi Institute of Engineering and Technology, Department of Mathematics, Vizianagaram, India

Bhattacharya, Sovan	DR. B.C. Roy Engineering College, Computer Science and Engineering, Durgapur, India
Bodepudi, Nageswararao	
Borugadda, Dr. Somasekhar	
Bouabdallaoui, Doha	Hassania School of Publics Works (EHTP), Laboratory of Systems Engineering, Casablanca, Morocco
Camilleri, Rebecca	University of Malta, Computer Information Systems, Msida, Malta
Chemboli, Vaibhav	GITAM University, Department of Computer Science and Engineering, Visakhapatnam, India
Chowdary, Bhavya	
D, Dakshayani Himabindu	VNRVJiet, Information Technology, Hyderabad, India
Dubey, Vineet Kumar	HBTU, Computer Science, Kanpur, India
Duddupudi, Janardhan Santosh Kumar	
Dabbeeru Priyanka,	Aditya Institute of Technology and Management, Computer Science and Engineering, Tekkali, India
Dhotre, Sudhirkumar	
G, Sathya Priyanka	Periyar University, Statistics, Salem - 11, India
Gudla, Sateesh	
Gajjala, Nirmala Joycee	St Anns College for Women, Computer Science, Hyderabad, India
Goyal, Jayanti	
Gubbala, Dr. Kumari	CMR Engineering College, CSE, Hyderabad, India
Hossain, Dr. Syed Akhter	
Inala, Dr. Krishna Pavan	BVRIT Narsapur, Electronics and Communication Engineering Department, Narsapur, India
Ingale, Hemant	
Iyer, Sailesh	
Javvadi, N V R Swarup Kumar	
Jena, Sanjay Kumar	C. V. Raman Global University, Computer Science Engineering, Bhubaneswar, India
K. Vigneswara Reddy	Research Scholar, Annamalai University, CSE, Chidambaram, India
Kale, Kiran	
Krishnareddy, K	Mother Theresa Institute of Engineering & Technology, EEE, Palamaner, India
Karri, Praveen Kumar	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India
Karyemsetty, Nagarjuna	
Khemchandani, Maahi	Dr. Babasaheb Ambedkar Technological University Vidyavihar, Lonere, Maharashtra 402103, Information Technology, Lonere, India
Kumar, Krishan	NIT Kurukshetra, Computer Engineering, Kurukshetra, Indian
Kumar, Achal	Anand Engineering College Agra, Computer Science & Engineering, Agra, India

Lakshmikanthareddy, M Laxmikanth, Pydipala	Vignan Institute of Technology and Science, Department of Computer Science and Engineering, Yadadri Bhuvanagiri District, India
Lokesh, K M, N P Patnaik	GITAM Deemed to be University, Visakhapatnam, Department of Computer Science Engineering, Visakhapatnam, India
Madala, Chandu Jagan Sekhar	
Mahanty, Dr. Mohan	Vignan's Institute of Information Technology(A), Computer science and Engineering, Visakhapatnam, India
Majumder, Swanirbhar Mishra, Madhusudhan	Tripura University, Information Technology, Agartala, India NERIST, Electronics and Communication Engineering, Nirjuli, India
Mohan, Dr. T Murali	Swarnandhra Institute of Engineering and Technology, Seethrampuram, Computer Science and Science, East Godavari, India
Muttipati, Appala Srinivasu	
N M, Jyothi	Koneru Lakshmaiah Education Foundation, Computer Science Engineering, Vaddeswaram, India
Nandwalkar, Dr. Jayant Netto, Ann Nita PYLA, J V G Prakasa Rao	LBSITW, Poojappura, Trivandrum, ECE, Trivandrum, India Lendi Institute of Engineering and Technology, CSE, Vizianagaram, India
Pal, Saurabh Pandit, Dipti Patil, Mithun	N K Orchid College of Engg & Technology , Solapur MH India, Department of Computer Science and Engg. Solapur, India
Patil, Shashikant	ViMEET, Computer Science & Engineering (AI & ML), Raigad, India
R, Ramyadevi	SRM Institute of Science and Technology, Department of Computer Science and Applications, Chennai, India
Rajesh, P K Ramesh, R	Arignar Anna Government Arts College Musiri, Mathematics, Musiri Trichy, India
Ramakrishna, Chithari	CMR institute of Technology (Autonomous), Hyderabad, CSE(DS), Hyderabad, India
SABBI, Simhadri Dakshayani Kanaka Mahalakshmi Sabharwal, Seema	Government Post Graduate College for Women, Department of Computer Science, Panchkula, India
Sagar, Sanjeela Sambana, Bosubabu	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India
Shaik, Mahaboob Hussain	Vishnu Institute of Technology, Computer Science and Engineering, Bhimavaram, India

Sharma, Amita	
Sharma, Anshuman	Kalinga Institute of industrial Technology, CSE, Bhubaneswar, India
Simhadri, Kumaraswamy	Aditya Institute of Technology And Management, EEE, Tekkali, India
Sinha, Anurag	IGNOU, Computer Science, Ranchi, India
Sridhar, B	Lendi Institute of Engineering and Technology, ECE, Vizianagaram, India
Srinivas, Thiruchinapalli	T Srinivas, Associate Professor in Mathematics, Nellore, Inida
Srinivasa Rao, M	Vignan's Institute of Information Technology, Computer Science and Engineering, Visakhapatnam, India
Sundara Siva Rao, Ivaturi	
Tuduku, Vinod Kumar	Lendi Institute of Engineering and Technology, Management Studies, Vizianagaram, India
Talha, Iftakhar Mohammad	
Thirupatirao, Tirilingi	
Tirumala, Pentapati	
Uppalapati, Padma Jyothi	
Usha Bala, Dr. Varanasi	Anil Neerukonda Institute of Technology and Sciences, Computer Science and Engineering, Visakhapatnam, India
Vardhan, Dr. D. Harsha	
Vamsi, T M N	
Vamsi B, Dr	
Vasamsetti, Samuel	
Venunath, M	Pondicherry University, Computer Science, Puducherry, India
Vijaychandra, Joddumahanthi	LIET, Vizianagaram, India
Yegireddi, Dr. Ramesh	
Kantharaju, Kanaparti	VITAP University, Computer Science and Engineering, Andhra Pradesh, India
Rao, Srinivas	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India
Samatha, Badugu	

Reviewers

Kesswani, Nishtha	Department of Computer Science, Central University of Rajasthan, Ajmer, Rajasthan, India
Ashiq, Abdul	Sri Eshwar College of Engineering, Information Technology, Coimbatore, India
Bokka, Yugandhar	
Bada, Rajasekhar Reddy	New Horizon College of Engineering, Computer Science and Engineering, Bangalore, India

Bandi, Vamsi	Madanapalle Institute of Technology and Science, Artificial Intelligence and Data Science, Madanapalle, India
Bano, Shahana	Robert Gordon University, School of Computing, Aberdeen, UK
Bendi, Dr. Nageswara Rao	Lendi Institute of Engineering and Technology, Department of Mathematics, Vizianagaram, India
Bodepudi, Nageswararao	
Borugadda, Dr. Somasekhar	
Bouabdallaoui, Doha	Hassania School of Publics Works (EHTP), Laboratory of Systems Engineering, Casablanca, Morocco
Camilleri, Rebecca	University of Malta, Computer Information Systems, Msida, Malta
Chemboli, Vaibhav	GITAM University, Department of Computer Science and Engineering, Visakhapatnam, India
Chowdary, Bhavya	
Dubey, Vineet Kumar	HBTU, Computer Science, Kanpur, India
Duddupudi, Janardhan Santosh Kumar	
Dabbeeru Priyanka,	Aditya Institute of Technology and Management, Computer Science and Engineering, Tekkali, India
Dhotre, Sudhirkumar	
G, Sathya Priyanka	Periyar University, Statistics, Salem - 11, India
Gudla, Sateesh	
Gajjala, Nirmala Joycee	St Anns college for women, Computer Science, Hyderabad, India
Goyal, Jayanti	
Gubbala, Dr. Kumari	CMR Engineering College, CSE, Hyderabad, India
Hossain, Dr. Syed Akhter	
Inala, Dr. Krishna Pavan	BVRIT Narsapur, Electronics and Communication Engineering Department, Narsapur, India
Ingale, Hemant	
Iyer, Sailesh	
Javvadi, N V R Swarup Kumar	
Jena, Sanjay Kumar	C. V. Raman Global University, Computer Science Engineering, Bhubaneswar, India
K, Vigneswara Reddy	Research Scholar, Annamalai University, CSE, Chidambaram, India
Kale, Kiran	
Krishnareddy, K	Mother Theresa Institute of Engineering & Technology, EEE, Palamaner, India
Karri, Praveen Kumar	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India
Karyemsetty, Nagarjuna	
Khemchandani, Maahi	Dr. Babasaheb Ambedkar Technological University Vidyavihar, Lonere, Maharashtra 402103, Information Technology, Lonere, India
Kumar, Krishan	NIT Kurukshetra, Computer Engineering, Kurukshetra, Indian

Kumar, Achal	Anand Engineering College Agra, Computer Science & Engineering, Agra, India
Lakshmikanthareddy, M Laxmikanth, Pydipala	Vignan Institute of Technology and Science, Department of Computer Science and Engineering, Yadadri Bhuvanagiri District., India
Lokesh, K M, N P Patnaik	GITAM Deemed to be University, Visakhapatnam, Department of Computer Science Engineering, Visakhapatnam, India
Madala, Chandu Jagan Sekhar	
Mahanty, Dr. Mohan	Vignan's Institute of Information Technology(A), Computer science and Engineering, Visakhapatnam, India
Majumder, Swanirbhar Mishra, Madhusudhan	Tripura University, Information Technology, Agartala, India NERIST, Electronics and Communication Engineering, Nirjuli, India
Mohan, Dr. T Murali	Swarnandhra Institute of Engineering and Technology, Seethrampuram, Computer Science and Science, East Godavari, India
Muttipati, Appala Srinivasu	
N M, Jyothi	Koneru Lakshmaiah Education Foundation, Computer Science Engineering, Vaddeswaram, India
Nandwalkar, Dr. Jayant Pyla, J V G Prakasa Rao	Lendi Institute of Engineering and Technology, CSE, Vizianagaram, India
Pal, Saurabh Pandit, Dipti Patil, Mithun	
Patil, Shashikant	N K Orchid College of Engg & Technology , Solapur MH India, Department of Computer Science and Engg, Solapur, India ViMEET, Computer Science & Engineering (AI & ML), Raigad, India
R, Ramyadevi	SRM Institute of Science and Technology, Department of Computer Science and Applications, Chennai, India
Rajesh, P K Ramesh, R	Arignar Anna Government Arts College Musiri, Mathematics, Musiri Trichy, India
Ramakrishna, Chithari	CMR institute of Technology (Autonomous), Hyderabad, CSE(DS), Hyderabad, India
SABBI, Simhadri Dakshayani Kanaka Mahalakshmi Sabharwal, Seema	
Sagar, Sanjeela	Government Post Graduate College for Women, Department of Computer Science, Panchkula, India
Sambana, Bosubabu	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India

Shaik, Mahaboob Hussain	Vishnu Institute of Technology, Computer Science and Engineering, Bhimavaram, India
Sharma, Amita	
Sharma, Anshuman	Kalinga Institute of industrial Technology, CSE, Bhubaneswar, India
Simhadri, Kumaraswamy	Aditya Institute of Technology and Management, EEE, Tekkali, India
Sinha, Anurag	IGNOU, Computer Science, Ranchi, India
Sridhar, B	Lendi Institute of Engineering and Technology, ECE, Vizianagaram, India
Srinivas, Thiruchinapalli	T Srinivas, Associate Professor in Mathematics, Nellore, India
Srinivasa Rao, M	Vignan's Institute of Information Technology, Computer Science and Engineering, Visakhapatnam, India
Sundara Siva Rao, Ivaturi	
Tuduku, Vinod Kumar	Lendi Institute of Engineering and Technology, Management Studies, Vizianagaram, India
Talha, Iftakhar	
Mohammad	
Thirupatirao, Tirlingi	
Tirumala, Pentapati	
Uppalapati, Padma Jyothi	
Usha Bala, Dr.Varanasi	Anil Neerukonda Institute of Technology and Sciences, Computer Science and Engineering, Visakhapatnam, India
Vardhan, Dr. D. Harsha	
Vamsi, T M N	
Vamsi B, Dr.	
Vasamsetti, Samuel	
Venunath, M	Pondicherry University, Computer Science, Puducherry, India
Vijaychandra,	LIET, Vizianagaram, India
Joddumahanthi	
Yegireddi, Dr. Ramesh	
Kantharaju, Kanaparti	VITAP University, Computer Science and Engineering, Andhra Pradesh, India
Rao, Srinivas	Lendi Institute of Engineering and Technology(A), Vizianagaram, Computer Science and Engineering, Vizianagaram, India
Samatha, Badugu	

Contents

Similarity Analysis of Protein Sequences with a New 3D Graphical Representation Technique	1
Kshatrapal Singh, Ashish Kumar, and Manoj Kumar Gupta	
Enhanced Security and Robustness of Data Using Steganography	13
P. LaxmiKanth, O. Sri Nagesh, V. S. S. P. L. Balaji Lanka, and P. Ramamohan Rao	
A TTIG-Based Deep Convolution Combined GAN and CLS for Text-to-Image Synthesis	27
Raswitha Bandi, M. Sumanth, V. Lowkya, V. Manichandana, and K. Srinidhi	
Smart Agricultural Greenhouse System: A Context-Aware Application ..	47
Sujata Swain and Rajdeep Niyogi	
Improving Performance of Plant Disease Detection Using YOLOv7 and YOLOv8	57
Lakshmi Narayana Chintala, K. Sreerama Murthy, and Venkata Ramana Kondapalli	
Detection of Congenital Heart Disease from Heart Sounds Using 2D CNN-BiLSTM with Attention Mechanism	73
Ann Nita Netto, Lizy Abraham, and Saji Philip	
Automated Reviewer Assignment Process Using Machine Learning Technique	87
Sovan Bhattacharya, Arkaprava Mazumder, Ayan Banerjee, Chandan Bandyopadhyay, and Subrata Nandi	
A Method for Detecting Retinal Microaneurysms in the Fundus Using CR-SF and RG-TF	101
S. Steffi	

Pilot Super-Resolution Network (PSRN)-Based Mango Fruit Classification	119
P. V. Yeswanth, Sammeta Kushal, V. Tharun Kumar, N. R. Ackshay, Ravindra Gangudi, Molapally Tharun Kumar, S. Deivalakshmi, Y. Thanya, and K. M. Lokesh Kumar	
Performance Evaluation of QoS in MAODV Routing Protocol in MANETS	133
Satish Dekka, Bosubabu Sambana, K. Narasimha Raju, D. Manendra Sai, M. Pallavi, and Kuppireddy Krishna Reddy	
Minimize the Energy Consumption to Increase the Network Lifetime for Green IoT Environment	151
Satish Dekka, K. Narasimha Raju, D. Manendra Sai, M. Pallavi, and Bosubabu Sambana	
Unveiling Hate Speech: Identifying Toxic Comments Targeting Women in Online Social Media Posts	161
M. Naveen Kumar, Kumari Gubbala, U. Mahender, Hari Keerthana Guna, and Chilaka Venkateswarlu	
Identification of Retinal Fundus in Diabetic Patients Using Deep Learning Algorithms	173
Vidyullatha Sukhavasi and Divya Yeluri	
Smart Health Prediction Using Random Forest	185
Kooragayala Sukeerthi, Kurma Pooja Reddy, Shaik Thasleema, and Paindla Sowjanya	
Vulnerability Assessment and Penetration Testing Using Parrot Operating System	197
Avvaru R. V. Naga Suneetha, J. K. V. Prasanthi, S. Nimisha, and C. H. Meghna	
Assessing NSAID Threat Degree of Unfavorable Medical Reactions Using Machine Learning	209
Raja Vikram Gandham, P. Tharun Kumar, B. Mahesh Gopal, M. Karthik, and Krishan Dev Nidumolu	
Searchable Encryption for Privacy Preserving with Fine-Grained Access Control	225
Manne Archana, Raparathi Pranathi, Kalakota Shreya, and Boodida Nikhitha	
Classification for Disease Gene Association	241
P. LaxmiKanth, J. Akshitha, P. Akshitha, and D. Abhinandhan	

Machine Learning-Based Air Pollution Monitoring and Forecasting	259
Naga Ravindra Babu M, M. Durga Satish, B. V. Prasanthi, S. V. V. D. Jagadeesh, J. N. S. S. Janardhana Naidu, and Immidi Kali Pradeep	
A Novel Parasitic Mushroom-Like Structure with High Gain Microstrip Patch Antenna for Broadband Applications	273
M. Sahaya Sheela, G. Syam Sudheer Babu, S. N. V. Sai Durga Prasad, and M. D. Vasanth Kumar	
Facial Emotion Recognition Using Artificial Intelligence	285
G. Sateesh, Swaroop Sana, S. V. R. Vara Prasad, and Bosubabu Sambana	
A Hybrid Machine Intelligence Demographic Feature Selection Approach to Improve Recommendation System in Social Domain	303
Bandi Vamsi, Mohan Mahanty, and Bosubabu Sambana	
An Exploratory Review of Machine Learning and Deep Learning Applications in Healthcare Management	315
Narasimha Rao Vajjhala and Philip Eappen	
Bone Fracture Prediction Using Machine Learning and Deep Learning Techniques	325
Satya Vamsi Kumar Appala, S. V. V. D. Jagadeesh, M. Durga Satish, and B. Sri devi	
Plant Disease Detection Using Modern Deep Learning Approach: YOLOv7	337
Ayan Banerjee, Arkaprava Mazumder, Ayush Kumar Shaw, Udit Narayana Kar, Sovan Bhattacharya, and Chandan Bandyopadhyay	
Analysis of the Life Insurance Business Performance Based on COVID by Using Machine Learning Algorithms	347
P. Nithya, C. D. Nandakumar, and S. Srinivasan	
An Ensemble Model of Skin Disease Detection Using CNN and Transfer Learning	357
Bhagyalaxmi K., Vemulapally Vennela, N. Tirumal Reddy, and Shaik Saba Maheen	
Session-Based News Recommendation System	367
V. Vemani, Vaibhav Chemboli, and Pusarla Sindhu	
A Fusion-Based Approach for Generating Image Captions	379
Samatha J. and G. Madhavi	

Comparison of Machine Learning Algorithms for Detection of Stuttering in Speech 391
Sarvagna Gudlavalleti, P. Sunitha Devi, Ramyasri Lakka, Rithika Kuchanpally, and Sai Sonali Dudekula

The Evolutionary Impact of Pattern Recognition in Research Applications: A Wide Spectrum Survey 405
Sumit Pal, Sovan Bhattacharya, Bappaditya Mondal, Anjan Bandyopadhyay, Dola Sinha, and Chandan Bandyopadhyay

Prediction of GATE Examination Clearance for Fresh Graduate Candidates: An Advanced Machine Learning Approach 417
Ayan Banerjee, Rachana Das, Puja Kumari, Ankita, Syed Zahir Hasan, and Sovan Bhattacharya

Foreseeing Worker Attrition Using Machine Learning 429
P. LaxmiKanth, P. Maruthi Vara Prasad, S. Jitendra, and A. Yashwanth

Mouse Controlling Using Eyeball Action 445
S. Kranthi Reddy, D. Shivananda Reddy, B. Suresh, and B. Pavan Kumar

Power Quality Improvement by Using Shunt Hybrid Active Power Filter 457
D. V. Kiran, G. Neetha, G. Gowtham, K. Anusha, K. Ravali, and A. Bharath Kumar

Integration of Renewable Energy Systems Into Utility Grid: A Review on Power Quality Issues, Mitigating Devices, and Control Algorithms 467
Joddumahanthi Vijaychandra, Santi Behera, and Lingraj Dora

Traffic Control System-Based Congestion Control and Emergency Vehicle Clearance 481
K. Krishna Reddy, S. Noor Mohammad, J. Divya, C. Vamsi, S. Ameer Basha, and B. Suresh Reddy

QR-Based Authentication for Login and Payment 489
J. Bibiana Jenifer, S. Sivaramakrishnan, Akhil Raula Satish, V. Preran, S. Chirag, and Shamolima Dutta

Smart Irrigation Watering System Using IoT 497
K. Krishna Reddy, G. Faazil, K. Ajith, C. Pavani, J. Sai Tharun, D. Dhanush Gowdu, and T. Sravani

IoT-Based Transmission Line Multiple Fault Detection and Indication to Electricity Board 503
K. Jeevan Reddy, K. Aruna, K. Manoranjitha, M. Bhavani Sankar, C. Ravindra, and G. Sireesha

Design of Off-Board Electric Vehicle Charger Using PV Array Through MATLAB-Simulink	511
N. V. Kishore Kumar, K. Nithya Sri, S. Akram, O. Praveen Kumar, V. Tharun, and G. Jaswanth	
Human Stress Detection in Sleep Mode Compared with Non-sleep Mode Using Machine Learning Algorithms	527
S. A. Sajidha, M. Sanjay, Pillaram Manoj, A. Sheik Abdullah, R. Priyadarshini, V. M. Nisha, and Aakif Mairaj	
Medical Diagnosis Prediction Using Deep Learning	541
Sri Karthik Avala, Simran Bohra, and S. A. Sajidha	
Detecting Hard Landing of Flights: E-Pilots	565
Revelle Akshara, Nomula Aishwarya, M. Karuna Shree, and N. Lahari	
Detection of Glaucoma Using MobileNet, XAI, and IML	577
A. Rakesh, P. Chetan, R. Shailender Raj, E. Ravi Kondale, and V. Kakulapati	
Attainment Expedients of Markovian Heterogeneous Water Heaters in Queueing Models by Matrix Geometry Method	589
P. Syamala, R. Ramesh, and M. Seenivasan	
Identification of Medicinal Plants Using Inception V3 Model	601
Manoj Kumar Mahto, Vinjamoori Manaswini, D. Akshara, and Indhirala Jayasree	
Smart Gardening Using Internet of Things	613
M. G. Mahesh, S. Afsana, K. Divya, P. Menaka, G. Prasad, and M. A. Suhanulla Khan	
Predictive Analytics of Blood Donor Risk Assessment Using Machine Learning Methods	617
V. Kakulapati, Avula Sravan Reddy, S. Mani Teja, and Acha Vamshi	
Risk Analysis of COVID-19 Patients Mortality Rate in Emergency Ward	631
V. Kakulapati, Gadala Praveen, G. Dheeraj, and E. Nagaraju	
Machine Learning Classification Analysis on Leaf Disease Data	645
Dileep Kumar Kadali, M. Srinivasa Rao, Podagatlapalli Vinay, G. A. K. S. Rajeev Kumar, D. V. Naga Raju, and G. Ratnakanth	
Conversion of Type-2 Intuitionistic Fuzzy Sets into Interval-Valued Intuitionistic Fuzzy Sets and Its Implementation in Decision-Making	657
V. Sireesha, N. Annapurna, and V. Anusha	

A Framework for Secure Database and Similarity Comparison in Android 673
Aishwarya Phadtare and Kishor Mane

A Comprehensive Review and a Conceptual Framework for Predicting the Position of the Mobile Sinks in Wireless Sensor Networks 683
Uma Maheswari Gali and Nagarjuna Karyemsetty

Brain–Computer Interface for Multiple Applications Control 695
K. Krishna Reddy, P. Purushotham, S. Rihan, P. Gowtham, N. Madhav, K. Bhargav, and K. Teja Kumar

Predicting Student Academic Performance Using Machine Learning: A Comparison of Classification Algorithms 703
Naresh Bhimavarapu, B. V. Prasanthi, C. H. Lakshmi Veenadhari, M. Durga Satish, Venkata Durga Rao Matta, and Immidi Kali Pradeep

A Novel Approach in Machine Learning for Solar Energy Prediction System 717
Supriya Vaddi, Fathima Sumreen, Vaishnavi Narayanam, Akshara Patlan-nagari, and Peruru Ananya

Real-Time Tomato Leaf Disease Detection and Diagnosis Using Deep Learning-Based Computer Vision Techniques 725
S. Kanakaprabha, P. Arulprakash, G. Ganesh Kumar, T. Udhayakumar, and R. Janani

Index 735

Similarity Analysis of Protein Sequences with a New 3D Graphical Representation Technique



Kshatrapal Singh , Ashish Kumar, and Manoj Kumar Gupta

Abstract In this study, we develop a new three-dimensional (3D) graphical representation technique for protein sequences. Our approach uses the physico-chemical properties of 20 amino acids and a BLOSUM62 matrix. This evolutionary information yields perceptive visualizations. For an analysis of protein similarities, a specific vector from 3D graphical representation curves is also formed. Our proposed method is verified on two datasets: first on eight species' ND6 protein sequences and second on 22 species' ND5 protein sequences. On the same dataset, the phylogeny achieved by our method agrees with that obtained by other researchers. Our approach, unlike older approaches, does not require the alignment of protein sequences.

Keywords Protein sequence · Graphical representation · Clustering · Phylogenetic tree · BLOSUM62 matrix · Numerical characterization

1 Introduction

Two major issues in bioinformatics are the analysis of and the comparison of DNA and protein sequences to effectively establish evolutionary relationships between them. The main objective of an evolutionary study is to monitor the status of the system over a long time period. Repeating evolutionary events is not practical in a laboratory, so the methods of establishing biological sequence are carried out mostly via computation and statistical approaches. Various mathematical methods have

K. Singh (✉)

KIET Group of Institutions, Ghaziabad, India

A. Kumar

I.T.S Engineering College, Greater Noida, India

M. K. Gupta

Shri Mata Vaishno Devi University, Katra, J&K, India

been given by other researchers for translating protein sequences for two- or three-dimensional (2D or 3D) graphical representations. To compare protein sequences, we use various mathematical objects.

For visualization and DNA analysis, an approach that graphically represents the data is a crucial tool. Recently, to find similarities and dissimilarities between a DNA sequence [1–6] and a protein sequence [7–13], data from various graphical representation models proposed by other researchers were incorporated into our algorithms and computational models. The main problem is that other researchers have considered sequences as strings. They also considered four nucleotide bases as a DNA sequence and 20 amino acids as a protein sequence. In our approach, we consider the chemical structure and physicochemical properties of sequence data. The physicochemical properties of amino acids have strong effects on the substitution rate of amino acids. With the help of these properties, therefore, we can directly estimate various distances between various amino acid sequences. In the case of the alignment-free graphical representation approach, the computation cost is very low [14–16].

Graphical representations of a protein sequence are considered highly complex because a protein contains 20 amino acids, unlike a DNA sequence, which consists of four nucleotides [17]. In this research, on the basis of the physicochemical properties of amino acids and the BLOSUM62 matrix, we suggested a new 3D graphical representation approach using protein sequences (two datasets). Our proposed representations yielded perfect visualizations that were free from any circuits. After that and with the graphical curves of amino acid sequences, we generated a particular vector. Based on the generated vectors, an analysis of similarities and evolutionary relationships was also carried out on two datasets. Our results are consistent with evolutionary facts and research carried out by other researchers. These facts indicate that our method can be applicable to longer sequences of various lengths while maintaining good performance.

2 Proposed Models and Methodology

Protein sequences are made up of 20 types of amino acids. Each amino acid contains specific, distinct physicochemical properties [18, 19]. Accordingly, to extract important information from protein sequences, here we propose taking a new graphical approach by collecting features of amino acids and using the BLOSUM62 matrix.

2.1 *A Novel 3D Graphical Representation of Protein Sequences*

The concept of the BLOSUM62 matrix was devised by Henikoff and Henikoff [20], and the matrix works as substitution matrix. The cost of the BLOSUM62

Table 1 Ten properties, and their numerical values, of 20 amino acids

Amino acid	Pr1	Pr2	Pr3	Pr4	Pr5	Pr6	Pr7	Pr8	Pr9	Pr10
A (alanine)	2.34	9.69	7.0	6.0	0.33	−0.06	31	−0.11	0.23	8.1
R (arginine)	2.17	9.04	9.1	10.7	−0.17	−0.17	124	0.07	0.21	10.5
N (asparagine)	2.02	8.80	10.0	5.41	−0.23	0.16	56	−0.13	0.24	11.6
D (aspartic)	2.09	9.82	13.0	2.77	−0.37	0.07	54	−0.28	0.17	13
C (cysteine)	1.71	10.7	4.8	5.07	0.07	0.38	55	−0.18	0.22	5.5
Q (glutamine)	2.17	9.13	8.6	5.65	−0.41	−0.02	85	−0.24	0.26	10.5
E (glutamic)	2.19	9.67	12.5	3.22	−0.25	−0.18	83	−0.06	0.18	12.3
G (glycine)	2.34	9.60	7.9	5.97	0.37	−0.01	3	−0.07	0.16	9.0
H (Histidine)	1.82	9.17	8.4	7.59	−0.07	0.05	96	0.32	0.20	10.4
I (isoleucine)	2.36	9.68	4.9	6.02	0.14	−0.31	111	0.001	0.27	5.2
L (Leucine)	2.36	9.60	4.9	5.98	0.13	−0.26	111	−0.01	0.28	4.9
K (lysine)	2.18	8.95	10.1	9.74	−0.07	−0.37	119	0.04	0.22	11.3
M (methionine)	2.28	9.21	5.3	5.74	−0.09	0.07	105	−0.04	0.25	5.7
F(phenylalanine)	1.83	9.13	5.0	5.48	0.01	0.07	132	0.43	0.23	5.2
P (Proline)	1.99	10.6	6.6	6.3	0.37	−0.03	32.5	−0.01	0.16	8.0
S (serine)	2.21	9.15	7.5	5.68	0.02	0.47	32	−0.15	0.23	9.2
T (threonine)	2.63	10.4	6.6	6.16	0.14	0.34	61	−0.21	0.21	8.6
W (tryptophan)	2.38	9.39	5.2	5.89	0.01	0.05	170	0.49	0.18	5.4
Y (tyrosine)	2.20	9.11	5.4	5.66	−0.14	0.22	136	0.38	0.19	6.2
V (valine)	2.32	9.62	5.6	5.96	0.24	0.21	84	−0.15	0.25	5.9

matrix shows the chance of replacing one amino acid with another. Higher similarity is represented by positive scores and lower similarity by negative scores. In this chapter, we look at ten of the most important physicochemical features of amino acids: Property1 is pK1 (−COOH); property2 is pK2 is −NH₃; property3 is the polar requirement; property4 is the isoelectric point; property5 is the hydrogenation; property6 is the hydroxythiolation; property7 is the molecular volume; property8 is the aromaticity; property9 is the aliphaticity; and property10 is the polarity gains [21–23]. These properties and their numerical values are shown in Table 1.

By using the K-means clustering algorithm, 20 amino acids are classified into various categories on the basis of their chemical properties. The K-means method is an effective approach for natural language processing, data mining, and bioinformatics [24, 25]. To determine the number of clusters, we used the silhouette index for cluster validity. This contains two factors: cohesion and separation, whose values range between −1 and 1, where good clustering effects are represented by values greater than 0. According to this index, we are easily able to find a good number of clusters from the datasets. Table 2 presents the categorization of ten clusters on the basis of ten properties. For property1, pK1 (−COOH), we classified 20 amino acids into seven groups; if two or more acids are classified into this category, then they share the property pK1 (−COOH). If we proceed in same manner for the other properties, similarities among properties can be found in each pair of amino acids.

Table 2 Groups of 20 amino acids

Property number	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
Pr1	AGILMWV	HF	RQEKS	Y T	NP	C	D			
Pr2	AEI	QHMF	SY P	RK	W	C	T	N	GLV	D
Pr3	CILF	QH	E	A	NK	MWYV	GS	PT	R	D
Pr4	ANCQGHIL MFPSTWYV	RK	DE							
Pr5	ILT	HKM	DQ	GP	NE	FPW	V	RY	A	C
Pr6	IL	HMF	W CT	QGP	YV	RE	S	K	AD	N
Pr7	ILM	APS	W	QEV	NDCT	FY	G	H	RK	
Pr8	ARNDCQEG ILKMPSTV	HF	WY							
Pr9	IL	RCH	T AKFS	DGP	EWY	NQMV				
Pr10	NK	CILMFWYV	AGPST	DE	RQH					

The similarity between two amino acids, S_{AB} , is given by

$$S_{AB} = N_{AB} \times M_{AB} \quad (1)$$

where A and B are the two amino acids; N_{AB} is number of similar properties among amino acids A and B; and M_{AB} is the value of amino acids A and B in the BLOSUM62 matrix. By using this equation, we can find all the values for S_{AB} . Similar values of S_{AB} for these amino acids may be numerically distinct from one amino acid to another. We used the linear regression technique to find the various characteristics of each amino acid to compare their degrees of similarity in graphical form. Consequently, we can now find the slope and the intercept of amino acid A. The various slopes, intercepts, and linear equations of 20 amino acids are shown in Table 3.

Let $P = p_1, p_2, p_3, \dots$ and p_n be a random order of proteins consisting of n amino acids. With the help of x_i, y_i , and z_i , we can represent p_i in 3D coordinates. Now 3D representations of protein sequences can be made by following Eq. 2:

$$p_i = \begin{cases} x_i = i \\ y_i = y_{i-1} + l_i \\ z_i = z_{i-1} + t_i \end{cases} \quad (2)$$

where l_i and t_i represent the slope and the intercept of p_i , respectively. Initially, we take $y_0 = z_0 = 0$. Now we are able to convert n into a graphical curve.

Table 3 Values of the slope, intercept, and linear equation of 20 amino acids

Amino acid	Slope	Intercept	Linear equation
A (alanine)	0.05	-2.46	$y = 0.05x + (-2.46)$
R (arginine)	-0.05	-0.4	$y = -0.05x + (-0.4)$
N (asparagine)	-0.14	-1.23	$y = -0.14x + (-1.23)$
D (aspartic)	0.01	-1.35	$y = 0.01x + (-1.35)$
C (cysteine)	0.09	-5.44	$y = 0.09x + (-5.44)$
Q (glutamine)	-0.22	0.26	$y = -0.22x + (0.26)$
E (glutamic)	-0.19	1.0	$y = -0.19x + (1.0)$
G (glycine)	-0.3	-2.25	$y = -0.3x + (-2.25)$
H (Histidine)	-0.08	-1.28	$y = -0.08x + (-1.28)$
I (isoleucine)	0.32	-5.26	$y = 0.32x + (-5.26)$
L (Leucine)	0.23	-4.16	$y = 0.23x + (-4.16)$
K (lysine)	-0.19	1.33	$y = -0.19x + (1.33)$
M (methionine)	0.16	-3.4	$y = 0.16x + (-3.4)$
F(phenylalanine)	0.32	-4.61	$y = 0.32x + (-4.61)$
P (Proline)	0.02	-4.26	$y = 0.02x + (-4.26)$
S (serine)	-0.36	1.74	$y = -0.36x + (1.74)$
T (threonine)	0.05	-2.51	$y = 0.05x + (-2.51)$
W (tryptophan)	0.06	-3.65	$y = 0.06x + (-3.65)$
Y (tyrosine)	0.23	-3.11	$y = 0.23x + (-3.11)$
V (valine)	0.05	-3.09	$y = 0.05x + (-3.09)$

2.2 Numerical Characterization of Proteins

According to the graphical curves, we are able to find a particular vector against a protein sequence. Thanks to this vector, we are now able to analyze the similarity between two protein sequences. The main approach to finding the pairwise distance between two protein sequences is the characteristic vector. A good characteristic vector must avoid the concerns about the length and complexity of a sequence in its calculation [26].

We decided on using a two-tuple (\bar{l}_A, \bar{t}_A) of amino acid A for characterization. For protein sequence $P = p_1, p_2, p_3, \dots, p_n$ consists of n amino acids, and we can calculate the two-tuple by using Eq. 3:

$$\bar{l}_A = l_A \times \frac{N_A}{n} \text{ and } \bar{t}_A = t_A \times \frac{N_A}{n} \quad (3)$$

In Table 3, l_A and t_A represent the slope and the intercept of amino acid A, respectively. The count of amino acid A in a sequence is referred to as N_A . According to Eq. 3, N_A/n refers to the amino acid A ratio in its full sequence. With

every type of amino acid, we obtain a two-tuple for characterization. For all 20 amino acids, we have to create a 40-dimensional (40D) characteristic vector:

$$(\bar{l}_A, \bar{t}_A, \bar{l}_R, \bar{t}_R, \bar{l}_N, \bar{t}_N, \dots, \bar{l}_V, \bar{t}_V)$$

Using the sequence NNDDCDNNCC as an example, we can determine that the number of amino acids for N, D, and C in the sequence is 4, 3, and 3, respectively. Hence, we can obtain the 40D characteristic vector (0, 0, 0, 0, -0.56, -0.492, 0.003, -0.405, 0.027, -1.632, 0, 0, \dots, 0, 0) on the basis of using Table 3 and Eq. 3.

The similarity between two protein sequences can be represented by using similarity distance. Various approaches for finding similarity distances are available, such as Euclidean distance, Manhattan distance, and city-block distance. In this chapter, we used the Euclidean distance approach to find the similarity between two sequences. Here, a 40D characteristic vector helped us to compute the similarity distance. Two 40D characteristic vectors are represented by U and V , as follows:

$$U = (\bar{l}'_A, \bar{t}'_A, \bar{l}'_R, \bar{t}'_R, \bar{l}'_N, \bar{t}'_N, \dots, \bar{l}'_V, \bar{t}'_V)$$

$$V = (\bar{l}''_A, \bar{t}''_A, \bar{l}''_R, \bar{t}''_R, \bar{l}''_N, \bar{t}''_N, \dots, \bar{l}''_V, \bar{t}''_V)$$

Euclidean distance is measured by the following formula:

$$d(U, V) = \sqrt{\sum_{i=1}^{40} (U_i - V_i)^2} \quad (4)$$

As the value of distance d decreases, the similarity between two protein sequences increases.

3 Results and Discussion

3.1 Analysis of ND6 Protein Sequences of Eight Species

In this section, our novel proposed 3D graphical representation approach of protein sequences was applied to create the phylogenetic tree of eight ND6 proteins. Table 4 contains detailed information on ND6 proteins, along with their respective accession numbers. These sequences were obtained from the GenBank of the National Center for Biotechnology Information (NCBI).

By following our approach, for each sequence, we obtained a corresponding 20-dimensional (20D) vector. By calculating the similarity distance between various pairs of specific vectors, we obtained a similarity distance matrix, as shown in

Table 4 Eight ND6 proteins from NCBI dataset

Sr. number	Sequence name	Accession number
1	Human	AP_000650
2	Gorilla	NP_008223
3	Chimpanzee	NP_008197
4	Harbor seal	NP_006939
5	Gray seal	NP_007080
6	Rat	AP_004903
7	Mouse	NP_904339
8	Wallaroo	NP_007405

Table 5 Similarity matrix for eight ND6 protein sequences

Species	Chimpanzee	H. seal	Gorilla	G. seal	Mouse	Rat	Human	Wallaroo
Chimpanzee	0							
H. seal	0.04365	0						
Gorilla	0.01226	0.07119	0					
G. seal	0.06662	0.00071	0.05329	0				
Mouse	0.13501	0.10110	0.10584	0.11004	0			
Rat	0.20335	0.11893	0.12548	0.13501	0.02135	0		
Human	0.01036	0.08758	0.00442	0.07981	0.13664	0.183	0	
Wallaroo	0.93125	0.71146	0.79254	0.65428	0.68542	0.7514	0.91668	0

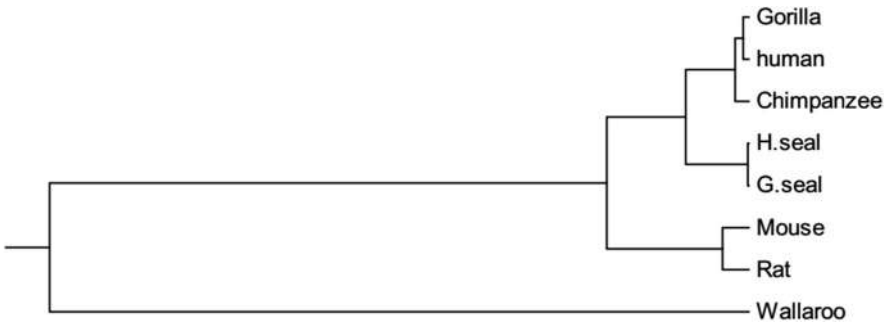


Fig. 1 The phylogenetic tree for eight ND6 protein sequences

Table 5. After analyzing Table 5, we found the smallest value to be that between the harbor seal and the gray seal, which indicates that this species pair has the highest similarity among all species pairs. In the same manner, chimpanzee, gorilla, and human are highly similar to one another. We constructed the phylogenetic tree of eight sequences by using the unweighted pair group method with an arithmetic mean (UPGMA) approach [27] and MEGA 10.1 software, which takes a similarity matrix as input (see Fig. 1). According to Fig. 1, gorillas, humans, and chimpanzees; H. seals and G. seals; and mice and rats are all closely related, where each group shares a common ancestor. As demonstrated in Fig. 1, the Wallaroo is most dissimilar to all the other various species.

Table 6 Comparison between our study and other studies on species similarity to humans

Study	Gorilla	Chimpanzee	H. seal	G. seal	Rat	Mouse	Wallaroo
This study	0.00442	0.01036	0.08758	0.07981	0.183	0.13664	0.91668
From Table 2 in [31]	0.00575	0.0125	0.0694	0.0736	0.196	0.12756	0.80224
From Table 3 in [30]	0.0094	0.0118	0.0247	0.0284	0.033	0.0262	0.0369
From Table 2 in [29]	0.0338	0.0979	0.1797	0.1487	0.2071	0.1472	0.278
From Table 4 in [28]	8.25	6.92	12.81	13.11	14.63	15.03	–

Table 7 Information on 22 ND5 protein sequences

Sequence name	Accession number	Sequence name	Accession number
Blue whale	NP_007066	Mouse	NP_904338
Bornean orangutan	NP_008235	Opossum	NP_007105
Cat	NP_008261	Bonobo	NP_008209
Chimpanzee	NP_008196	Platypus	NP_008053
Fin whale	NP_006899	Rat	AP_004902
Gibbon	NP_007832	Rhino	YP_002520019
Gorilla	NP_008222	Sumatran orangutan	NP_007845
Gray seal	NP_007079	Walleroo	NP_007404
Harbor seal	NP_006938	Tiger	ADK73290
Human	AP_000649	Korean bovine	YP_209215
Horse	ADQ55101	Spain bovine	AKK32014

3.2 Comparisons of Our Study with Previous Studies

We collected the results from several other studies that evaluated the rate of similarity between humans and various other species and compared those results to our study's results; that comparison is shown in Table 6. The conclusions of previous researchers were based on the same dataset as the one that we used. According to Table 6, (a) gorillas and chimpanzees have the most similarity to humans (these two columns are in boldface) and are in the first group; (b) harbor seals and gray seals are in the second group; (c) rats and mice are in the third group for their dissimilarity to humans; and (d) wallaroos are the most dissimilar to humans, according to this evolutionary relationship model using the datasets of various protein sequences.

3.3 Similarity Analysis of ND5 Protein Sequences of 22 Species

To validate our proposed method, we selected various datasets on different species and performed several experiments by using software like Octave 5.2.0 and MEGA 10.1. In this section, we applied our approach to test the protein sequences of 22 ND5 animal species, which were taken from the NCBI database and whose information is listed in Table 7.

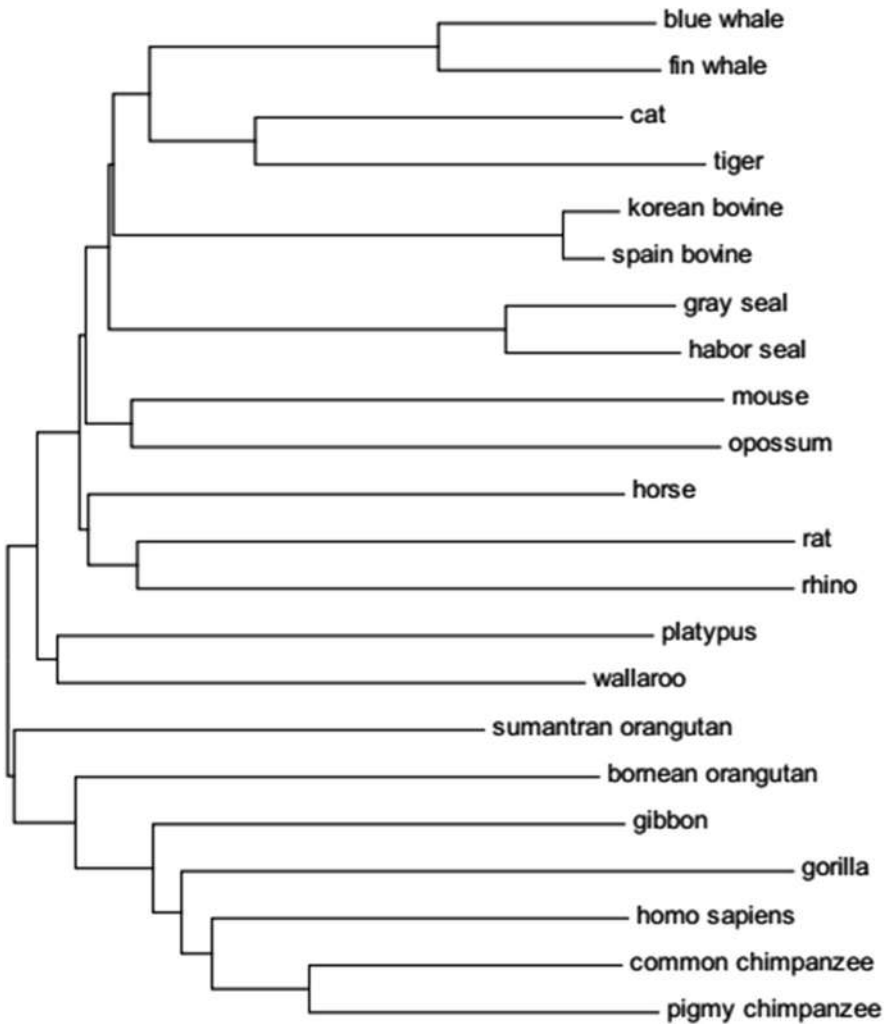


Fig. 2 The phylogenetic tree of the protein sequences of 22 ND5 species according to our method

The phylogenetic tree of the protein sequences of each of these 22 ND5 animal species is shown in Fig. 2. The pairs of Korean bovines and Spanish bovines; blue whales and fin whales; and chimpanzees and bonobos exhibit minimal differences according to the analysis. The established homology is in strong agreement with their respective evolutionary relationships. We also found that the pair of humans and gorillas and that of gibbons and orangutans have very close relationships with each other, according to this phylogenetic tree. All the classification that we have carried out is in agreement with the theory of classical evolution. For a comparison of our method with method in ref. [32], the outputs shown in Fig. 3 can be used.

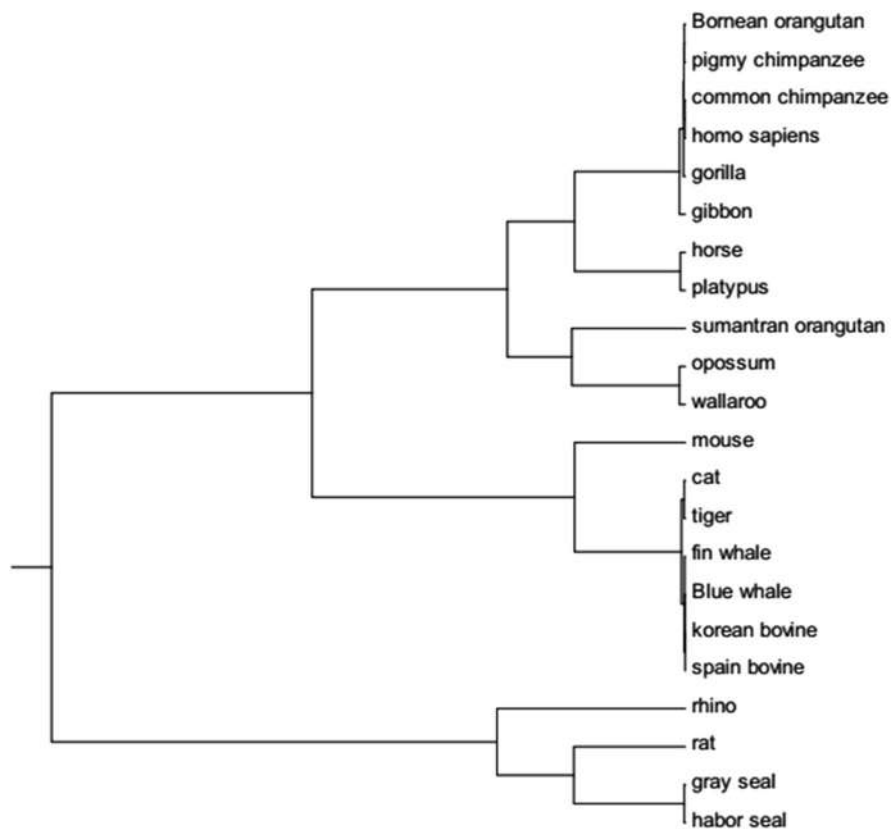


Fig. 3 The phylogenetic tree of the protein sequences of 22 ND5 species according to Chou's approach

The clusters in Figs 2 and 3 are similar, but some differences remain. In Fig. 3, the Sumatran orangutan and the Bornean orangutan exhibit a close relationship, but they are depicted in separate branches. In Fig. 2, we corrected the irregular allocations. Our analysis concludes that our proposed approach is also more effective than previous approaches in analyzing ND5 protein sequences.

4 Conclusion

In this study, a new 3D graphical representation technique was used to examine protein sequences. The method suggested in this chapter produced a useful outcome for analyzing protein sequences. Our approach was based on ten physicochemical properties and the BLOSUM62 matrix. We calculated specific vectors on the basis of graphical curves. To find similarity distances between various protein sequences,

we used these specific vectors. To validate the effectiveness of our proposed approach, we used two datasets: first, eight ND6 protein sequences of various species, and second, 22 ND5 animal protein sequences. Our results validated our approach to phylogeny and showed improvements over the results of other related studies.

References

1. C. Yu, M. Deng, S. S. T. Yau, DNA sequence comparison by a novel probabilistic method, *Inform. Sci.* 181 (2011) 1484–1492.
2. N. Jafarzadeh, A. Iranmanesh, A novel graphical and numerical representation for analyzing DNA sequences based on codons, *MATCH Commun. Math. Comput. Chem.* 68 (2012) 611–620.
3. J. Song, Analysis of similarity of DNA sequences based on a novel 3-D graphical representation, in: T. Chang (Ed.), *Advances in Biochemical Engineering, Inf. Engin. Res. Inst.*, Newark, 2012, pp. 29–36.
4. Bielińska-Waż D. Graphical and numerical representations of DNA sequences: statistical aspects of similarity. *J Math Chem.* 2011;49:2345–2407.
5. Bielińska-Waż D, W, aż P. Spectral-dynamic representation of DNA sequences, *J Biomed Inform.* 2017;72:1–7.
6. Wąż P, Bielińska-Waż D. 3D-dynamic representation of DNA sequences. *J Mol Model.* 2014;20:2141
7. H. J. Yu, D. S. Huang, Novel 20-D descriptors of protein sequences and it's applications in similarity analysis, *Chem. Phys. Lett.* 531 (2012) 261–266.
8. Z. H. Qi, J. Feng, X. Q. Qi, L. Li, Application of 2D graphic representation of protein sequence based on Huffman tree method, *Comput. Biol. Med.* 42 (2012) 556–563.
9. Randić M, Noviĉ M, Plavšić D. Milestones in graphical bioinformatics. *Int J Quant Chem.* 2013;113:2413–2446.
10. Wang L, Peng H, Zheng J. ADLD: a novel graphical representation of protein sequences and its application. *Comput Math Method Med.* 2014;2014:959753.
11. El-Lakkani A, Mahran H. An efficient numerical method for protein sequences similarity analysis based on a new two-dimensional graphical representation. *SAR QSAR Environ Res.* 2015;26:125–137.
12. Qi ZH, Jin MZ, Li SL, Feng J. A protein mapping method based on physicochemical properties and dimension reduction. *Comput Biol Med.* 2015;57:1–7.
13. Hu H. F-Curve, a graphical representation of protein sequences for similarity analysis based on physicochemical properties of amino acids. *MATCH Commun Math Co.* 2015;73:749–764.
14. Amorim RCD, Hennig C. Recovering the number of clusters in data sets with noise features using feature rescaling factors. *Inform Sci.* 2015;324: 126–145.
15. Qi ZH, Feng J, Liu CC. Evolution trends of the 2009 pandemic influenza A (H1N1) viruses in different continents from March 2009 to April 2012. *Biologia.* 2014;69:407–418.
16. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 2013;30: 2725–2729.
17. W. Hou, Q. Pan, M. He, A. Novel, 2D representation of genome sequence and its application, *J. Comput. Theor. Nanosci.* 11 (2014) 1745–1749.
18. C. Yin, S.S. Yau, An improved model for whole genome phylogenetic analysis by Fourier transform, *J. Theor. Biol.* 382 (2015) 99–110.
19. T. Hoang, C. Yin, H. Zheng, C. Yu, R. Lucy He, S.S. Yau, A new method to cluster DNA sequences using Fourier power spectrum, *J. Theor. Biol.* 372 (2015) 135–145.

20. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *P Natl Acad Sci.* 1992;89:10915–10919.
21. T. Ma, Y. Liu, Q. Dai, Y. Yao, P.-A. He, A graphical representation of protein based on a novel iterated function system, *Phys. A* 403 (2014) 21–28.
22. C. Yu, M. Deng, S.Y. Cheng, S.C. Yau, R.L. He, S.S. Yau, Protein space: a natural method for realizing the nature of protein universe, *J. Theor. Biol.* 318 (2013) 197–204.
23. S.S. Yau, W.G. Mao, M. Benson, R.L. He, Distinguishing proteins from arbitrary amino acid sequences, *Sci. Rep.* 5 (2015) 7972.
24. Y. Li, K. Tian, C. Yin, R.L. He, S.S. Yau, Virus classification in 60-dimensional protein space, *Mol. Phylogenet. Evol.* 99 (2016) 53–62.
25. P.A. He, S.N. Xu, Q. Dai, Y.H. Yao, A generalization of CGR representation for analyzing and comparing protein sequences, *Int. J. Quantum Chem.* 116 (2016) 476–482.
26. A. Czerniecka, D. Bielińska-Wąż, P. Wąż, T. Clark, 20D-dynamic representation of protein sequences, *Genomics* 107 (2016) 16–23.
27. H. Skutkova, M. Vitek, K. Sedlar, I. Provaznik, Progressive alignment of genomic signals by multiple dynamic time warping, *J. Theor. Biol.* 385 (2015) 20–30.
28. M. Randić, M. Novič, M. Vračko, On novel representation of proteins based on amino acid adjacency matrix, *SAR QSAR Environ. Res.* **19** (2008) 339–349.
29. L. Z. X. Xie, Y. Yu, L. Liang, M. Guo, J. Song, Z. Yuan, Protein sequence analysis based on hydropathy profile of amino acids, *J. Zhejiang Univ. Sci. B* **13** (2012) 152–158.
30. Y. H. Yao, Q. Dai, L. Li, X. Y. Nan, P. A. He, Y. Z. Zhang, Similarity/dissimilarity studies of protein sequences based on a new 2D graphical representation, *J. Comput. Chem.* **31** (2010) 1045–1052.
31. Manoj Kumar Gupta, Rajdeep Niyogi, Manoj Misra, A 2D graphical representation of protein sequence and their similarity analysis with probabilistic method, *Communications in mathematical and in computer chemistry* 72 (2014) 520–532.
32. M. K. Gupta, R. Niyogi, M. Misra, An alignment-free method to find similarity among protein sequences via the general form of Chou's pseudo amino acid composition, *SAR QSAR Environ. Res.* 24 (2013) 597–609.

Enhanced Security and Robustness of Data Using Steganography



P. LaxmiKanth , O. Sri Nagesh , V. S. S. P. L. Balaji Lanka ,
and P. Ramamohan Rao

Abstract Secure links are not used by the Internet. That is the main reason that the information may be vulnerable in transit to the interception. This is a major issue, while the data is to be transmitted via an unsecured link. One way we can use steganography as a solution. Steganography is a traditional technique and popular, which is anonymity in nature. This technique is the art of science used to hide the data between receiver and sender. Steganography is a popular technique to hide data from different files, they include images, video, and audio files. The requirements may vary from application to application. For Image steganography, there is a huge diversity of Stegano techniques. Some techniques are complex and some techniques are simple, and all techniques have weak and strong points. One of the techniques most popularly used is the Least Significant Bit (LSB) method. By randomly dispersing the message bits inside the image, this improved LSB technique makes it more difficult for unauthorised individuals to extract the original data from the stegano-image. This chapter is an outline of image steganography and its techniques. This chapter tells the importance of good steganography algorithms suitable for various applications.

Keywords Steganography · LSB · ISB · MLSB · DFrFT · PSNR · MSC · IWT

1 Introduction

Steganography is the practice of writing by concealing data or images. Steganos means shielded art of writing. It is used not only for hiding data but also for sending

P. LaxmiKanth (✉) · O. S. Nagesh · V. S. S. P. L. B. Lanka
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

P. R. Rao
Department of CSE, Sri Vasavi Engineering College (A), Tadepalligudem, India

confidential data secretly and securely through insecure channels. Only the recipient can access the original secret data when it is concealed using steganography. In the olden days, the messages are securely written on the back of the wax, on the tables, on rabbit's abdomens, or on the scalp of prisoners. Now people are sending messages or information through the Internet by video, audio, text, or images. In order to send data securely or confidentially, the data is placed inside the audio, video, or images. Steganography is termed as the study of communication that is not visible. Secrecy is achieved in image steganography by embedding messages into the image to generate a stegano image. This chapter includes steganography techniques like Least Significant Bit (LSB), MLSB, ISB, etc.

We use two techniques while using the Internet: cryptography and steganography. In cryptography, the messages are shared between sender and receiver by sharing common keys. Here the message is modified as an encrypted message at the sender end and decrypted at the receiver end. There will be no possibility of knowing the messages by the third party without knowing the decrypted key. But the attacker may intercept the data and decrypt it by any other malicious algorithms. To rectify the problems of cryptographic technique problems, steganography is used. So secret data is not revealed and misleads the attacker from secret data.

Figure 1 explains the process of embedding data inside the video, audio, or image of multimedia using steganography. In order to improve communication data confidentiality, both cryptography and steganography techniques are combined.

Steganographic Applications

- (i) Communication confidentiality
- (ii) Data variation protection
- (iii) Access control for digital content distribution
- (iv) Electronic-commerce
- (v) E-media
- (vi) Database Management System (DBMS)
- (vii) Watermarking
- (viii) Data storage secrecy



Fig. 1 Steganography architecture

2 Literature Survey

The Fractional Fourier Transform (FrFT)

The research on the fractional Fourier transform (FrFT), considered as a simplification of the conventional Fourier transform, was presented by David H. Bailey and Paul N. Swartztrauber [1] years ago in the math literature. The continuous fractional Fourier transform (FrFT), a unique type of FrFT that was founded on the existence of Discrete Fractional Fourier Transform (DFrFT), is described by Soo Chang et al. [2]. This training highlights the advantages of the unique fractional Fourier transform (DFrFT) for steganography in image processing when compared to other transforms. The outcome displays the identical peak signal-to-noise ratio (PSNR) in two areas (frequency and time) nonetheless DFrFT gives a benefit of an extra Segno key. The direction constraint of this change.

Least Significant Bit (LSB)

A distinction of the bare Least Significant Bit (LSB) approach was applied in 2013 by Akhtar et al. [3]. The bit-inversion technique has improved the quality of the stegano data or image. The stegano image of the PSNR increases thanks to the LSB method. The image will be generated perfectly by the bit pattern stores for which LSBs are reversed. In order to strengthen steganography by randomising the algorithm, RC4 is implemented for hiding messages. This data or message is converted into bits and scattered randomly onto the image instead of stored sequentially. The original message is very hard to extract by the attacker. Thus, the LSB technique is an improved technique for quality concerns and enhanced security concerns. This technique is very much suitable for medical-based image steganography considering patient information secrecy to be maintained. The protection and storage of digital images of medical can be done effectively by using this technique. The Integer Wavelet Transform technique is also a steganography technique used to defend the MRI image and convert it into a single container image. Arnold transform technique is applied to patients' diagnosis image and a secret image that is scrambled is obtained. As Inverse Integer Wavelet Transform (IWT) is used to obtain a fictitious hidden image, the quality of this image was good. The observation is quality concerns as the quality of the image improved drastically. Abdullah et al. did research on the rapid development of telemedicine [4]. Emam et al. talk about security issues when the data is available on servers [5]. Steganography plays a vital role in most of the security issues for audio, video, and text files. Peak signal-to-noise ratio (PSNR) and mean squared error (MSE) are used in Manser et al. [6]'s research for standard LSB and condition-based LSB. Mohammed et al. [7] did an investigation to improve the security level of the LSB method. Thenmozhi et al. [8] use the Set Partitioning in Hierarchical Trees (SPIHT) method to use lossless compression to convert the secret code into another image. Shabnam et al. [9] use MSE and PSNR for hiding information. Datta et al. [10] use binary addition for hiding the image. Shatnawi et al. [11] did research to improve image quality. Abhishek et al. [12] use deep learning techniques to improve hidden image quality. An overview of recent developments in LSB-based image steganography has been conducted by Prashanti et al. [13].

Multi Secure and Robustness of Medical Image

In 2013, Prabakaran et al. [14] devised that the medical reports of the patients are very sensitive and to be preserved securely robust and secure steganography techniques are proposed. These techniques became popular for the secure storage of medical images that are digitised. Some authors did research and proposed to use IWT in order to defend MRI image and place it as a single-container image. In this method, the transform technique of Arnold is given as an input to the patient's secret image in order to obtain a scrambled image that is kept in secrecy. Here, a secret image that had been scrambled had been inserted into a fake container image, and a fake secret image had been created using inverted IWT. The quality-related parameters are then enhanced with greater PSNR compared to earlier methods.

Cropping Function

In 2012, Thenmozhi et al. [15] invented a new structure that embeds messages or data in IWT coefficients in 8×8 blocks by utilising a cropping method on the image cover. After completely inserting the image, pixel change process is optimally done and applied. The writers hired the frequency sphere to rise the strength of our secret writing method by using steganography. Wavelet filter's floating point precision problems are avoided by IWT. Then capacity and signal-to-noise ratio are in the peak stage by using the IWT technique.

Huffman Encoding

In 2012, Das et al. [16] based on Huffman encoding for image steganography proposed a new technique. Two images of size $A \times B$ and $X \times Y$ 8-bit grayscale images are used as secret images and also the cover images, respectively. By changing the LSB of each pixel's intensity in the cover picture, each bit of the secret image or image's Huffman code is put inside the cover image during embedding. To make the Steg Image independent information, the Huffman encoded bit stream and Huffman Table is also included in the cover image. For the cover image to function as independent information, the Huffman encoded bit stream and Huffman Table are also incorporated.

Integer Wavelet Transform

IWT, which is utilised to hide the key, was created by Hemalatha et al. [17]. Because no one can see the concealed information and it cannot be lost owing to noise or signal processing techniques, it is therefore mostly safe and dependable. The peak signal-to-noise ratio is an excellent security indicator. By adjusting an integer wavelet coefficient in high-frequency sub-bands, the hidden information is hidden.

Secure Steganography Using Hybrid Domain Technique

It was reported in 2012 by Reddy et al. [18] that steganography was used to hide information. Secure Steganography Method Based on Hybrid Domains (SSHDT) is a secure steganography method based on hybrid domains. Different cover picture types and sizes are taken into consideration before being downsized to two-dimensional dimensions. In order to create four sub-bands from the cover image, the Daubechies Lifting Wavelet Transform (LWT) is applied. The XD, XH, XV, and

XA sub-bands are among them. The XD band, which is split into upper and lower equal blocks for the XD band, is used for payload embedding. A consideration of the payloads of different formats is made as well as resizing them to the power of two dimensions. In four equal segments, the payload is fragmented. Additional Stagno objects are scrambled using the Decision Factor-Based Manipulation to boost the payload's security (DFBM). The XD, XH, XV, and XA stegano objects are subjected to Daubechies Inverse LWT (ILWT) in order to produce segno images in the spatial domain. It has been noted that the proposed algorithm's PSNR and embedding capability are superior to those of the current algorithm.

Keeping Texture Distortion to a Minimum

Bingwen Feng et al. developed Secure Binary Image Steganography Based on Minimising Texture Distortion in 2015 [19] for securing binary image steganography. To reduce texture distortion, this method is suggested. In this steganography method, the binary picture is first used to extract the rotation, complement, and mirroring invariant texture patterns. They also suggested measuring, and this strategy was really put into practice using the recommended measurement. The recommended steganographic approach provides good statistical security, high stego picture quality, and high embedding capacity, according to real results.

Heuristic Genetic Algorithm

Shyla et al. [20] performed picture steganography using a genetic algorithm for embedding and selecting the cover image. By concentrating on the "before embedding hiding techniques," using this technique, the secret information is embedded in the cover image's most suitable locations. The image histogram undergoes the fewest possible changes as a result of the least amount of bit changes. The LSBs and secret messages are divided into a set of blocks using segmentation in this genetic process. Following the placement of the secret blocks in the proper places determined by this technique, it next builds the key file required for message extraction. According to experimental findings, this genetically based technique produces high steganographic quality images while being more effective than the standard LSB algorithm.

3 Image Steganography

3.1 Steganography Types

3.1.1 Steganography (Text Based)

It entails concealing data within text files. According to this strategy, every nth letter of each word in the text message conceals the secret data. There are numerous ways to conceal information or data in a text file.

These approaches are:

- (i) Method constructed on the format
- (ii) Statistical and random process
- (iii) Technique based on linguistics

The formatting of the text or specific properties of textual elements can be changed to create text steganography (e.g. characters). The aim of coding method design is to create modifications that are mostly undetectable to the reader while being consistently decoded (even noise is present). Designing document-marking systems is difficult because these requirements—reliable decoding and minimally noticeable change—are somewhat at odds. Instead of providing a full set of document-marking methods, the three coding strategies we suggest serve to demonstrate several approaches. The methods can be applied alone or collectively.

These are the following:

- (1) Coding by using Line-Shift: This technique involves vertically rearranging the positions of text lines in a document to uniquely encode it and may be used alone or in combination.
- (2) Coding by using Word-Shift: This technique modifies a document by horizontally moving words between text lines to uniquely encode the material.
- (3) Coding by using the Feature: This coding technique is used to encode a bitmap image of a document or a format file.

3.1.2 Image Steganography

Picture steganography is the practice of concealing data by using a cover item as the image. In image steganography, the data is concealed using pixel intensities. Images are frequently employed as cover sources in digital steganography since the digital representation of an image contains a number of bits.

3.1.3 Audio Steganography

When a concealed message is inserted into a digitalized audio signal in audio steganography, the matching audio file's binary sequence is illuminated. For audio steganography, numerous techniques are available. A quick introduction to some of them is what we'll do now. Data is concealed in audio files. Using this method, the contents in sound files such as WAV, AU, and MP3 are concealed. Audio steganography can be done in a variety of ways.

Some methods are

- (i) Encoding of low bit
- (ii) Phase coding
- (iii) Spread spectrum

3.1.4 Video Steganography

Any type of digital video file or data can be hidden using this method. Here, the info is concealed via video (a mixture of images). Discrete Cosine Transformations (DCT) often alter the values in each of the video images to disguise the data in a way that is undetectable to the human eye (for instance, from 8.667 to 9). The formats utilised for video steganography are H.264, Mp4, MPEG, and AVI.

The fundamental idea behind steganography is that, in all of these techniques, a secret message should be concealed within another cover object, which need not necessarily be significant, such that the cover data would be the only thing that the encrypted data eventually revealed. Therefore, unless appropriate decryption is applied, it cannot be quickly determined that it contains secret information.

4 Image Steganography

These days, it's common to practice concealing information within photographs. It is simple to share an image with a hidden message on the Internet or in newsgroups. German steganographic expert Niels Provos has studied the use of steganography in newsgroups and developed a scanning cluster that can find hidden messages in photographs that have been posted online. One million photos were checked, but no concealed messages were discovered, suggesting that steganography's practical applications are still quite limited. The process of concealing information or messages within an image so that the intended user cannot discover what is being concealed is known as image steganography. This process can be explained in Fig. 2.

The cover source can be updated in congested areas with lots of colour variations to conceal a message inside an image without changing its evident qualities and to make the changes less noticeable. The most popular methods for implementing these modifications include filtering, masking, and image alterations on the cover image, as well as the least significant bit, or LSB. These methods can be used on many kinds of image files with variable degrees of success. The primary objective of the study is to comprehend the many forms of steganography that are available. Images are encrypted using picture steganography, and the message or data image is recovered by also decrypting the relevant data. Image steganography is investigated because there are several ways to do this, and one of the methods is used to demonstrate it. The method of encrypting data, such as text, photos, or audio files, and hiding them within an image or video file is known as image steganography. The purpose of the current study is to combine two images using spatial domain steganography. This confidential information can only be acquired by employing the proper decoding techniques.

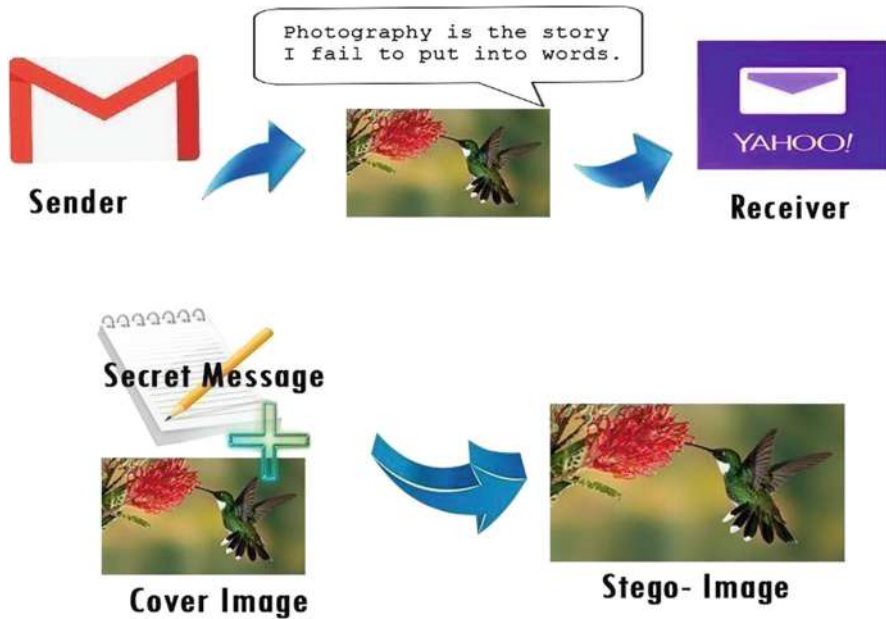


Fig. 2 Communication through steganography

5 Implementation

5.1 The Encoding Process

The LSB coding steganography method is employed. The header of the image contains the offset. To maintain the integrity of the header, that offset is left in place, and we begin the encoding process with the subsequent byte. Using a public key, the message is first converted to cipher text for encoding. Next, the data is embedded into the input carrier, which in this case is an image file, using the algorithm, and the carrier image is produced as a result. Creation of User Space.

5.2 The Decoding Process

The header of the image contains the offset. Use the same procedure as with the encoding to create the user space., utilising the Writable Raster and Data Buffer Byte classes' getRaster() and getDataBuffer() methods. A byte array is used to store the image's data. The bit stream of the original text file is reconfigured into another byte array using the byte array mentioned above. The decoded text file is subsequently

written with the aforementioned byte array, creating the encrypted text, which is then decrypted using the private key that was generated from the original message.

5.3 Least Significant Bit (LSB)

Image steganography can be done in two different ways:

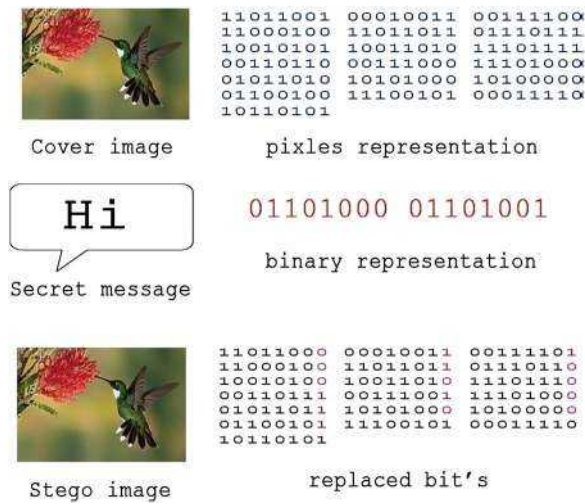
- (1) Spatial approaches
- (2) Transform approaches

But we are using Spatial Methods.

Spatial Technique

The LSB substitution technique is the utmost widely used technique in the spatial approach. A popular and straightforward way for implanting data in a cover file is the LSB method. The LSB substitution technique is employed in steganography, i.e. because there are three parts to every image Red, Green, Blue (RGB). This pixel data is encoded and fits into a single byte. The secret text can be stored by altering the initial bits of every pixel that carries this data. The text that will be stored must be stored at a size that is less than or equal to the image that will be used to cover the text. A spatial domain method is an LSB-based approach. However, this is susceptible to noise and cropping. The most common subgroup (MSB) in this approach is the Algorithm of the LSB method of steganography. The LSB approach may involve two distinct techniques for embedding and extracting data. Figure 3 explains the LSB technique. The following lists both phases’ algorithms:

Fig. 3 LSB operation



Embedding Phase Procedure:

- Step 1: Extract every pixel from the provided image and place it in an array named (image array).
- Step 2: Every character that can be extracted from the provided text file should be added to the array referenced to (message array).
- Step 3: It entails taking the typescripts out of the stegano key and putting them in a Key array instead. To prevent the embedded data from being found and/or recovered, the hiding process is managed by a segno-key.
- Step 4: The first component of the pixel should be filled with the first pixel and the characters from the Key array. Place the remaining characters in the first component of the next pixels if the Key array contains more characters than that.
- Step 5: Add some ending symbols to the key to denote its conclusion. In this procedure, the terminal symbol is set to 0.
- Step 6: Replace each component of the following pixels with a character from the message Array.
- Step 7: Continue performing step 6 until all characters have been implanted.
- Step 8: Once more, add some ending symbols to signify the data's end.
- Step 9: All of the characters that were entered will be hidden in the final image.

The message bits are deterministically embedded by the simplest steganography methods into the cover image's least significant bit plane. Due to the tiny amplitude of the change, modulating the least significant bit does not provide a noticeable effect on humans. A suitable cover image is required to conceal a concealed message inside of an image. A lossy compression format must be used because this technique uses bits from every pixel in the image; otherwise, the concealed information will be lost during the transformations of a lossy compression algorithm. Red, green, and blue colour components of a 24-bit colour image can each be represented by one bit, leaving three bits for storage.

```
(10100101 00101001 01001000)
(00100111 11001000 00101001)
(11001000 00100101 10101001)
```

The grid below is the result of the character A, whose binary value is 10000001:

```
(00100101 00101000 01001000)
(00100110 11001000 00101000)
(11001000 00100111 10101001)
```

To successfully insert the character in this instance, just three pieces need to be changed. Frequently, only half of a picture's pieces need to be changed, when using the maximum cover size in order to hide a concealed message. The information is effectively concealed because the least significant bit changes that are created as a result are too minute for the human visual system (HVS) to pick up on.

There have been no modifications to the third colour's least important component. It can be used to check the accuracy of the 8 bits that are stored within these 3 pixels. As a result, it might be a parity bit.

Filtering and Masking

An alternate method to hiding a message is taken using masking and filtering techniques, which are often limited to 24 bits or grayscale images. These techniques produce markings in an image that are practically identical to paper watermarks. For instance, this can be achieved by altering the intensity of image regions. While a mask may affect an image's evident features, it can also be carried out in a way that renders the modifications imperceptible to the human eye. Since masking utilises the visible portions of the image, it is more resistant to compression, cropping, and other methods of image processing than LSB alteration. Since the information is present in the visible portion of the image rather than being obscured by the noise level, it is more suitable than LSB.

6 Future Scope

But the concept of steganography is still relatively new. The world of computers is constantly evolving; thus, it stands to reason that steganography will also improve. There will probably soon be more effective and sophisticated Steganalysis methods. These are the potential future tasks for this project.

(1) Steganography in Image Files: How to Spot It.

Can steganography in image files be recognised? This is a challenging query. By examining the reduced bits of the image bits, it might be feasible to discover a straightforward steganographic technique. To tell whether a Steganalysis technique is more sophisticated and randomly disperses the embedded data throughout the image or encrypts the data before embedding can be quite difficult, though.

(2) Print media steganography:

Can information that has been printed, scanned, and stored in a file while embedded in an image be recovered?

This would necessitate using a unique type of steganography that might allow for errors in the printing and scanning machinery.

(3) The project can be expanded to the point where it can later be utilised for many image formats, such as .bmp, .jpeg, .tiff, etc.

(4) It is possible to increase the image-to-text compression ratio.

(5) The Least Significant Bit Algorithm provides reasonable security, but we can raise it to a certain point by employing alternative carriers and encryption and decryption methods explained in Figs. 4 and 5.

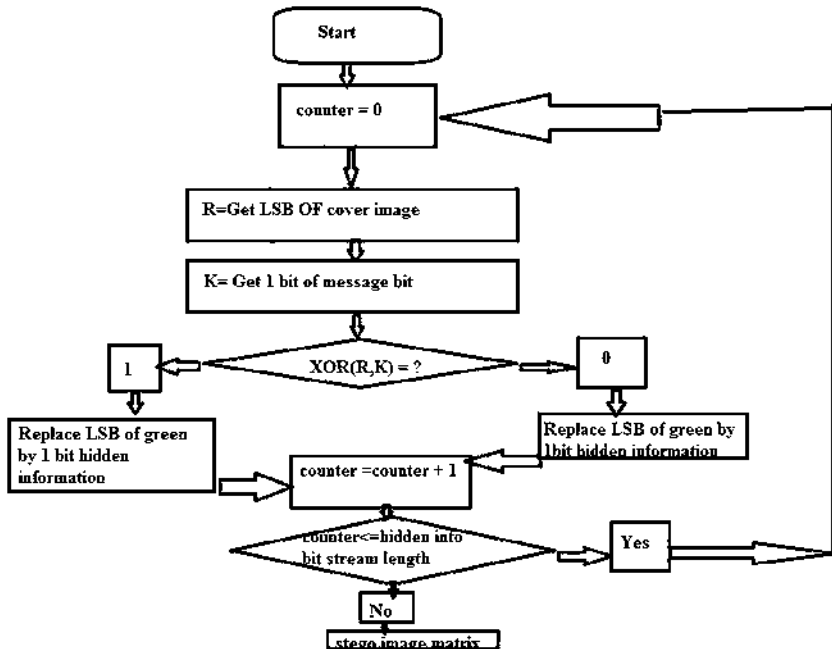


Fig. 4 Encryption methodology flow chart

7 Conclusion

It has been noted that the LSB Substitution Steganographic approach produces outstanding results in data concealment because it makes use of the straightforward fact that each image can be divided into separate bit planes, which each offers information on numerous levels. It should be emphasised that this method only functions with bitmap images because they require lossless compression techniques, as was previously stated. This technique can also be used with colour images; however, bit-plane slicing for each of the message image's top 4 bit planes for R, G, and B must be done independently.

Additionally, it is essential to discuss how, although steganography was formerly impenetrable, it is now possible to both identify its use and extract data using a variety of techniques. Simple approaches to determine whether an image file has been manipulated without the use of software or other sophisticated instruments for detection include, for instance, the following:

- (1) Image dimensions: A steganographic image takes up much more storage space than a conventional image of the same size. For instance, if the original image storage size was only a few KBs, the steganographic image could be many MBs in size. This again changes depending on the image type and resolution.

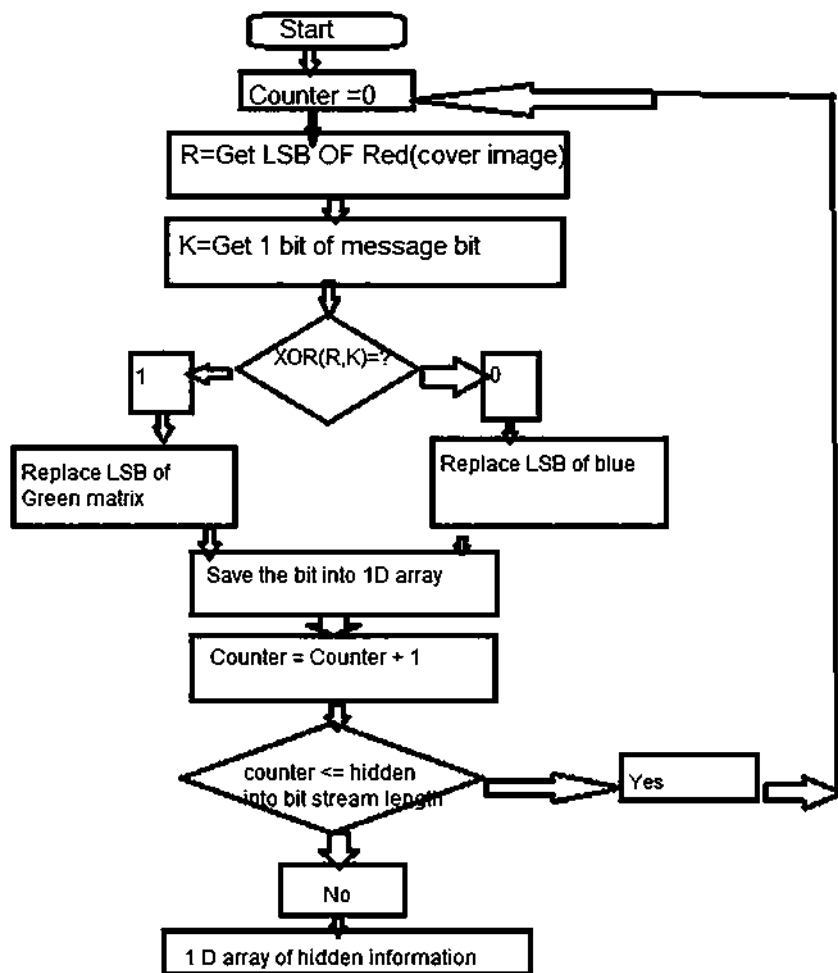


Fig. 5 Decryption methodology flow chart

- (2) Image noise: A steganographic image has noise when compared to a regular image. In order to make the steganographic image appear less noisy than the original cover image, a small amount of noise is first applied to the cover image.

Declaration

- (1) All authors do not have any conflict of interest.
- (2) This chapter does not contain any studies with human participants or animals performed by any of the authors.

References

1. David H. Bailey and paul N. Swarztrauber implemented “The Fractional Fourier Applications” in 1990.
2. Soo-Chang pe and Min-Hung yeh implemented “Discrete Fourier Transform” in the year 1996.
3. Nadeem Akhtar, Shahbaaz khan and Pragati johri implemented “An Improved Inverted LSB Image Steganography” in 2014.
4. Bushra Abdullah Shtayt, Nur Haryani Zakaria and Nor Hazlyna Harun made research in “A Comprehensive Review on Medical Image Steganography Based on LSB Technique and Potential Challenges” in 2021.
5. Emam, M. M., Aly, A. A., & Omara, F. A. An Improved Image Steganography Method Based on LSB Technique with Random Pixel Selection. *International Journal of Advanced Computer Science & Applications*, 1(7), pp. 361–366, (2016).
6. Saher Manaseer, Asmaa Aljawawdeh and Dua Alsoudi made research on “A New Image Steganography Depending On Reference & LSB” in 2017 Volume 12, Number 9 (2017) pp. 1950–1955.
7. Mohammed Abuzalata, Ziad Alqadi, Jamil Al-Azzeh and Qazem Jaber used for investigate a methodology to increase the security level of LSB method of data hiding Vol. 8, Issue. 2 pg. 93—103.
8. Thenmozhi, M. J., & Menakadevi, T. A New Secure Image Steganography Using Lsb And Spiht Based Compression Method. *International Journal of Engineering*, 16(17), (2016).
9. Shabnam, S., & Hemachandran, K. LSB based Steganography using Bit masking method on RGB planes. (IICSIT) *International Journal of Computer Science and Information Technologies*, 7 (3), pp. 1169–1173, (2016)
10. Datta, B., Mukherjee, U., & Bandyopadhyay, S. LSB Layer Independent Robust Steganography using Binary Addition. *International Conference on Computational Modeling and Security (CMS 2016)*, Elsevier Pub, (2016).
11. Al-Shatnawi, A. M. A new method in image steganography with improved image quality. *Applied Mathematical Sciences*, 6(79), 3907–3915, (2012).
12. Abhishek Das, Japsimar Singh Wahi, Mansi Anand, and Yugant Rana “Multi-Image Steganography Using Deep Neural Networks” 2021.
13. G. Prasanthi and K. Sandhyarani “A New Approach for Data Hiding with LSB Steganography” in a book chapter *Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2*.
14. G. Prabakaran, R. Bhavani and P. Rajeswari, “Multi secure and robustness for medical image-based steganography scheme” 2013 *International Conference on Circuits, Power and Computing Technologies (ICCPCT)*.
15. S. Thenmozhi, M. Chandrasekaran “Novel approach for image stenography based on integer wavelet transform” 2012 *IEEE International Conference on Computational Intelligence and Computing Research*.
16. Rig Das and Themrichon Tuithung “A novel steganography method for image based on Huffman Encoding” *National Conference on Emerging Trends and Applications in Computer Science (NCETACS)* 2012.
17. Hemalatha shivram, Dinesh U. Acharya, Renuka adage, and Priya Kamath “A secure image steganography technique using Integer Wavelet Transform” *Information and Communication Technologies (WICT)*, 2012 *World Congress*.
18. H.S. Manjunatha Reddy, N. Satisha, Annu Kumari and K B Raja “Secure steganography using hybrid domain technique” *International Conference on Computing and Networking Technology (ICCNT)* in 2012.
19. Bingwen feng, Wei Lu, and Wei Sun, “Secure Binary Image Steganography Based on Minimizing the Distortion on the Texture” *IEEE Transactions on Information Forensics and Security* in 2015.
20. M.K. Shyla, K.B. Siva Kumar, and Rajendra Kumar Das “Image steganography using genetic algorithm for cover image selection and embedding” in *Elsevier Vol 3 Dec 2021*.

A TTIG-Based Deep Convolution Combined GAN and CLS for Text-to-Image Synthesis



Raswitha Bandi, M. Sumanth, V. Lowkya, V. Manichandana, and K. Srinidhi

Abstract In recent years, text-to-image synthesis has drawn the attention of many researchers. Currently, GANs (generative adversarial networks) have been shown to deliver images from complex text. Though GANs provide reasonable images, obtaining accurate results is beyond the current artificial intelligence systems. On the other side, deep convolutional generative adversarial networks have been proven to compile images from specific types such as birds, cars, and dogs. In this work, we have combined two algorithms GAN-CLS from the authors Scott Reed and MS GAN (mode seeking generative adversarial networks) from the authors Qi Mao and Hisn-Yang Lee. One of the limitations of GAN-CLS is that while training the model is experiencing mode collapse problems. To overcome the mode collapse problem, we have combined MSGAN with GAN-CLS. We demonstrate the above-proposed algorithm by implementing it on the Birds Dataset. We have trained the model till 936 epochs.

Keywords Generative adversarial networks · Neural networks · Machine learning · GAN-CLS · MSGAN · Deep convolution generative adversarial networks

1 Introduction

In the era of rapid growth in technology, computers are becoming more powerful, which leads to output of numerous powerful applications. In the recent times, generative adversarial networks (GANs) have drawn attention of the researchers. Text-to-image synthesis is a complex task, researchers have also shown interest in

R. Bandi (✉)

Department of Information Technology, VNR VJIET, Hyderabad, Telangana, India
e-mail: raswitha_b@vnrvjiet.in

M. Sumanth · V. Lowkya · V. Manichandana · K. Srinidhi

UG Scholar Department of Information Technology, VNR VJIET, Hyderabad, Telangana, India

generation of images from visual description, but it is far away from the current systems. In this work, we are more focused on obtaining images from human sentence text. Obtaining attributes from the text often leads to a good outcome.

Often the problem of text-to-image synthesis can be divided into two sub-problems: one, learning the text attributes that are relevant to gather important visual details and the other are to use the gathered information for image synthesis. The power of deep convolutional and recurrent neural networks is that they can automatically learn high-text features from the words and characters. Google has its own word model called Word2Vec, which was trained on numerous huge amounts of text to obtain the attributes for the text. These attributes are helpful for the machine to understand better information about the text.

Deep learning solves the above-discussed problem by using neural networks. Generative adversarial networks have proved to generate more accurate images from the text. GANs also unlock the potential of the machines to learn and generate better images over a certain amount of training.

Consider the following the caption “This is a red small bird with a long-pointed beak.” A trained model directly generated the image from the above caption. But the problem of deep learning is that the conditional-based images suffer the image distribution problem because of the text descriptions are highly multimodal. The same also follows for the reverse translation method, i.e., image-to-text captioning.

Our goal in this work is to propose generative adversarial network Architecture that effectively compels text-to-image synthesis. The GAN-CLS architecture has suffered from mode collapse problem, i.e., the generator generates effective images for one specific discriminator, this causes the discriminator to never learn properly, and the outputs will be in one type with no or little diversity. This is when the GAN failure occurs leading to mode collapse problem [1].

To overcome the mode collapse problem, we have used the method from, i.e., mode seeking generative adversarial network (MSGAN) Architecture and implemented it on the GAN-CLS algorithm. These two algorithms will be discussed in further more detailed manner in the below sections. The dataset used in this project is Caltech- USCD Birds dataset, which contains around 11,788 images of various types of birds and has various bird part locations [2].

Although many researchers have used preprocessed image captions to train the model. Since, we wanted a real-time application, we went through the traditional method of preprocessing the given text and then converting them into machine-readable format. This is often achieved by training the neural network with text after certain amount of training the network gives the attributes for the word as a side result [2].

2 Method

Generative Adversarial Networks

Generative adversarial networks (GANs) consist of two components: one is Generator and the other is Discriminator. GANs’ principle is more like a min-max game

with two players. Or we can also say that it is more like a cop theft game. The Discriminator plays the role of cop the task of this is to identify the real images. On the other hand, the role of the generator is to generate images more real-like images. The discriminator distinguishes the images generated by the generator as real or fake. The generator tries to fool the discriminator by generating more real-like images. But the discriminator tries to improve its accuracy by identifying the images as fake/real.

The min-max equation for Generator $G(z)$ and Discriminator $D(x)$ is given below $V(D,G)$ [2, 3]:

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}} [\log D(x)] + E_{z \sim P_Z(z)} [\log (1 - D(G(z)))] \quad (1)$$

The above equation has global optimum when $p_g = p_{\text{data}}$ and in general p_g tends to be p_{data} .

At the start of the training, since the images generated by the generator make no sense, the discriminator easily rejects the images as fake images. But upon training for some time, the generator generates a reasonable image, which can fool the discriminator. Optimal and efficient results can be achieved by the generator if it focuses on maximizing the $[\log(D(G(z)))]$ instead of trying to minimize $[\log(1-D(G(z)))]$. GANs are really powerful when it comes to image generation and generally training till about 600–1200 epochs the generator generates more real-like images. But overfitting problem should be taken care of [2].

2.1 Preprocessing

Preprocessing is the method that is used to remove unwanted text, clean the text, and make text ready for the machine [4].

The following are the steps on how preprocessing is achieved:

1. Cleaning:

The stop words and punctuations are removed from the sentence. This includes removing unwanted words such as conjunctions and symbols.

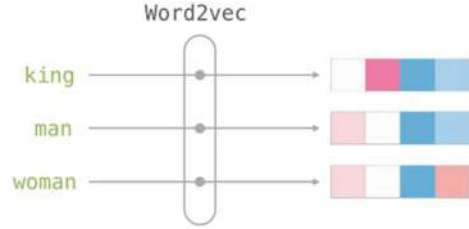
2. Tokenize:

The cleaned sentence is then made into smaller pieces known as tokens. Each individual word is treated as one token, tokens are made from space between the words.

3. Stemming:

Stemming is performed to remove extra content from the words. For example, consider the word “books” this can be replaced by the word “book.” Various algorithms are available for stemming one of them is porter’s stemming algorithm, and it works on the formula of $[C](VC)^M[V]$. Where V represents vowels, C represents consonants, and M represents number. Of times both V and C are occurring.

Fig. 1 Illustration of Word2Vec model



2.2 Word Embeddings

Machines cannot understand the words that humans understand. So, it is a good practice to make sure that we convert human-readable text into machine-readable format. Thanks to the concept of neural network the above said can be achieved. First, a neural network is defined with n number of neurons. Consider an example “A white bird with a black crown.”

First, the sentence is divided into tokens, and for each token the one-hot encoded vector is obtained. So, we input two words to the neural network expecting another word as output. After certain amount of training, we obtain the text attributes as a side result. Say if we input bird and black to the defined network it should output white. In the process of obtaining white as output, the neural network learns the text features [5].

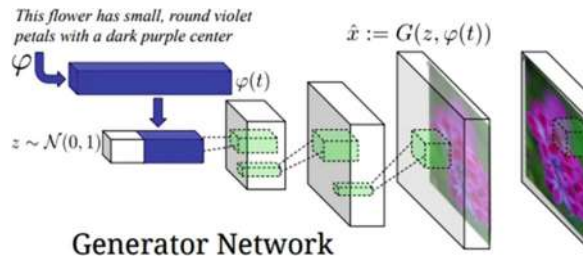
We have used Google’s Word2Vec Model as shown in Fig. 1 [2], it is a pretrained model which is used to obtain attributes of words in machine-understandable format. Word2Vec model is trained on huge amount of data, and it is convenient to use to obtain the text features.

These obtained text features are then given to the machine so that it can process and understand them.

The preprocessing needs to be done carefully because if our data contains unwanted features, the machine might learn that unwanted features and can make inaccurate decisions. It is always a good practice to preprocess the data while dealing with large dataset. The textual descriptions of the bird’s dataset are converted into machine understandable format using the word2Vec model and stored in a pickle file to load when training [6].

3 Literature Survey

Reed et al. [1] say that translating single sentence human description for the given text into the form of pixel image [7]. This uses birds used dataset and 102 flowers dataset by Oxford for model training and testing. The model uses text descriptions instead of class labels. Text-to-image translation can be obtained by using GANs

Fig. 2 GAN architecture

as shown in Fig. 2 [1, 8], which is a machine learning model used for generating images. This approach is to train a DCGAN on text that is encoded by a CNN. The network architecture has (generator, discriminator) for text encoding. Training a conditional GAN is to train the discriminator to judge the text and image pair as real or generated [9, 10].

The above image is a GAN. The image is generated via function G . Generator networks generate images, which will be judged by a network called discriminator, to see whether the generated image is valid or fake with pairs(text, images) [11]. This method uses the GAN-CLS approach to train the GAN, and the model distinguishes between realistic and unrealistic images of the incorrect class that do not adhere to requirements. As a third sort of input to GAN, the author provided actual photographs with text that were mismatched, which will aid the discriminator in identifying fake images. By interpolating caption embeddings from the training dataset, numerous more text embeddings can be produced.

Mao, Qi, et al. "Mode seeking generative adversarial networks for diverse image synthesis" [2] reported that MSGAN is a type of generative model that can be used for diverse image synthesis. They work by explicitly seeking out multiple modes or diverse image outputs, rather than simply generating a single output that represents the "average" of a given data distribution. The GAN objective function is modified to include a mode-seeking loss term, which pushes the generator to generate a variety of outputs that span a wider range of the data distribution. The authors use a number of benchmark datasets to show how successful MSGANs are, demonstrating that they can produce images of a much higher quality and diversity than conventional GAN models. In terms of generative modeling, MSGANs are a significant advancement since they offer a potent tool for synthesizing a wide range of picture sets that can be used for a number of tasks, such as data augmentation, image editing, and content creation [12]. They have used latent space to calculate the distance between the images to calculate the mode collapse problem. If the distance between the images is close, the model is said to have a mode collapse problem.

Sadia et al. [3] reported that when someone hears or reads a tale, their imaginations instantly conjure up images to help them visualize the content. A number of cognitive functions, including remembering, reasoning, and thinking, depend on visual mental picturing, sometimes referred to as "seeing with the mind's eye." The development of technology that can generate graphics that represent the substance

of textual descriptions and recognize the link between visual and verbal is a step toward empowering intelligence for the users [13]. To produce meaningful pictures from textual descriptions, CGAN were integrated with RNNs and CNN. The dataset utilized consisted of flower images and the appropriate linguistic descriptions. To produce realistic pictures from textual descriptions by using a GAN, preprocessing of the text and image scaling were done. They created a word list by preprocessing these caption phrases using the textual descriptions of the dataset. The matching IDs for these captions were then added to the list [14]. After loading, the photos were reduced in size to a certain size. The authors' suggested model is fed with the acquired data. RNN was used to describe the relationship between words at multiple timestamps in order to capture the context of text sequences. RNNs and CNNs were applied for text-to-image mapping. The CNN automatically identified the images' important details. The RNN produced 256-word embeddings of the textual descriptions in response to an input sequence. 512-dimensional noise vector was mixed with these word embeddings.

Sharma et al. [4] emphasized that by using skip thought vectors along with stack-based generative adversarial networks, they primarily concentrated on text-to-image synthesis. To produce fresh visuals, a GAN deep neural network is used. Its two networks are a generator network and a discriminator network [15]. Learn-phrase vectors incorporate skip-thought vectors. The encoder in skip thinking receives the training sentences and generates a vector, which is subsequently sent to the GAN network for additional processing. There are two text picture sub-methods. While the other entails creating a model that generates images using text representations, the first involves understanding description representations that correspond to the image provided in the text.

Flow diagram for stack-based GAN along with STV is shown in Fig. 3 [4].

In stage 1, the text description is formed into array-like vector, which is then given as the input to the generative adversarial network for image development and generation. Images created by the generator are given to the discriminator as an input, which uses pairs of images to determine whether they are real or phoney (text, image) [16].

The image is $64 * 64$ in size. Stack-based generative adversarial network along with skip thought vectors produces high-quality images by stacking multiple conditional generative adversarial networks given along with text embedding vector. In order to produce a good image with high-resolution image that is more realistic, the network creates a high-resolution image in stage 2 using the images produced by stage 1. Stage I produces photographs with a 64×64 resolution and little information, whereas Stage 2 produces images with a 256×256 resolution and extensive information [17].

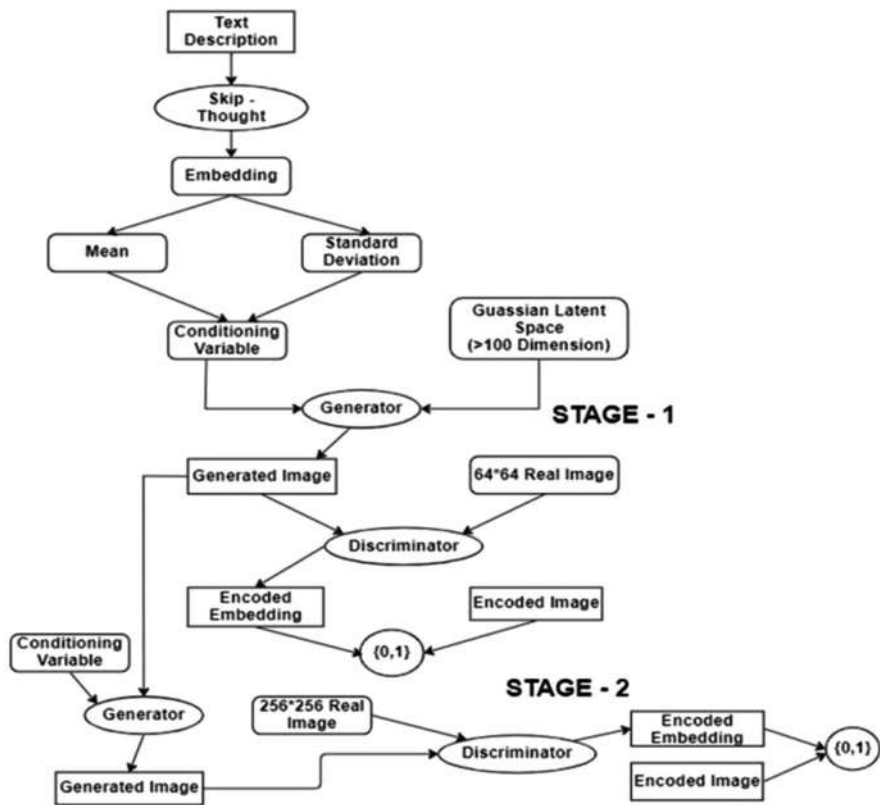


Fig. 3 C-GAN with STV

4 Algorithms

4.1 GAN-CLS

This is the algorithm shown in Fig. 4 [1] that was developed by the authors Scott Reed, Zeynep, Xincheng Yan, Lajanugen: pgeswaran and Honglak Lee [1]. The input to the generator is a joint pair of image and text. It is needed to see how well the generator performs in the conditional cases. At the beginning, the discriminator rejects the conditional information as it makes no sense but after the generator has learned to generate realistic images it must learn to align the output with conditional information and the discriminator should judge whether those conditional constraints are being met [18].

Fig. 4 GAN-CLS algorithm

```

1: Input: minibatch images  $x$ , matching text  $t$ , mis-
   matching  $\hat{t}$ , number of training batch steps  $S$ 
2: for  $n = 1$  to  $S$  do
3:    $h \leftarrow \varphi(t)$  {Encode matching text description}
4:    $\hat{h} \leftarrow \varphi(\hat{t})$  {Encode mis-matching text description}
5:    $z \sim \mathcal{N}(0, 1)^Z$  {Draw sample of random noise}
6:    $\hat{x} \leftarrow G(z, h)$  {Forward through generator}
7:    $s_r \leftarrow D(x, h)$  {real image, right text}
8:    $s_w \leftarrow D(x, \hat{h})$  {real image, wrong text}
9:    $s_f \leftarrow D(\hat{x}, h)$  {fake image, right text}
10:   $\mathcal{L}_D \leftarrow \log(s_r) + (\log(1 - s_w) + \log(1 - s_f))/2$ 
11:   $D \leftarrow D - \alpha \partial \mathcal{L}_D / \partial D$  {Update discriminator}
12:   $\mathcal{L}_G \leftarrow \log(s_f)$ 
13:   $G \leftarrow G - \alpha \partial \mathcal{L}_G / \partial G$  {Update generator}
14: end for

```

First, the text descriptions are encoded and a random noise is given to the network. The joint pair of text and noise is given as an input to the generator. The output and text are given as a pair as input to the discriminator, and the discriminator's loss and generator's loss are calculated and the weights of the network are reassigned. Here alpha refers to step size [18]. Discriminator learns to judge the image generator to be fake, whereas the generator learns to fool the discriminator by generating more realistic images. This process is repeated around 600–800 epochs [1, 19].

4.2 MSGAN

MSGAN stands for mode-seeking generative adversarial networks, and it is proposed by the authors Qi Mao, Hsin-Ying Tseng. The algorithm mainly focuses on the distance between the latent codes; if the distance between the latent codes is small, the images are likely to be in the same mode. Let z be the latent vector mapped to the image space I . Then two latent vectors are said to be in the same mode if the distance between the I_1 and I_2 is minimum.

To reduce the mode collapse issues, a new regularization term was derived by maximizing the ratio of distance between $I_1 = G(c, z_1)$ and $I_2 = G(c, z_2)$ with respect to distance between z_1 and z_2 .

The same can be written in the equation as [1, 20]:

$$\mathcal{L}_{ms} = \max_G (d_1(G(c, z_1), G(c, z_2)) / d_z(z_1, z_2)) \quad (2)$$

Here d^* stands for distance metrics.

The distance for a random generated is given in Table 1 [2].

Table 1 Distance for random modes

Mode	Distance
Mode—1	0.68
Mode—2	0.58
Mode—3	0.62
Mode—4	0.17

As we can see, the distance between the images generated in Mode 4 is very close to each other leading to Mode collapse problem.

Distance can be calculated using the formula [1, 21]:

$$d_I(I_a, I_b) / d_z(z_a, z_b) \quad (3)$$

Using the regularization term, the generator has a better chance to explore the latent space for generation of minor mode collapse, and the discriminator should be careful to evaluate the minor mode collapse images [2, 22].

4.3 Proposed Algorithm

Our proposed algorithm is to combine both MSGAN and GAN-CLS to overcome the problem of mode collapse in GAN-CLS.

The MSGAN algorithm can be combined with conditional GANs by simply adding the product of regularization term with regularization weights [23].

The new weights of the GAN-CLS then become:

$$L_{\text{new}} = L_{\text{old}} + \lambda_{\text{ms}} L_{\text{ms}} \quad [2] \quad (4)$$

where L_{new} is the new weight of the GAN-CLS, L_{old} is the old weight of the GAN-CLS, and L_{ms} is the regularization term of the MSGAN.

The equation of the generation for conditional text becomes [24]:

$$\mathcal{L}_{\text{ori}} = E_{c,y} [\log D(c, y)] + E_{c,z} [\log (1 - D(c, G(c, z)))] \quad [2] \quad (5)$$

where c , y , and z are class labels, images, and noise vectors.

This simply means that we are applying mode-seeking generative adversarial network from Eq. (3) on the GAN-CLS Algorithm from Fig. 2.

We trained the algorithm till 936 epochs on the Caltech-USCD Birds dataset which contains around 11,788 images of various types of birds and has various bird part locations [25].

The following is our proposed algorithm on GAN-CLS:

Algorithm [2]:

1. **Input:** images, right text, random text and no of steps S
2. **for** n=1 **to** S **do**
 - 1: h = Encode Real Text
 - 2: ~h = Encode Random Text
 - 3: z = Random Noise (0,1)
 - 4: sr = input real image and right text
(h) to Discriminator
 - 5: sw = input real image and random text (~h) to Discriminator
 - 6: $L_d = \log(sr) + (\log(1-sw) + \log(1-sf))/2$ Loss of Discriminator
 - 7: Update Discriminator with $L_d + \lambda_{ms} L_{ms}$
 - 8: $L_g = \log(sf)$ Loss of generator
 - 9: Update Generator with $L_g + \lambda_{ms} L_{ms}$
- End for**

4.4 Visualizing the Proposed Model

The above image contains the proposed model of our model is shown in Fig. 5 [3], it has dense layer, Re-Lu, Batch Normalization, Conv2D Transpose, and Leaky Re-Lu layers [26].

4.5 GAN Architecture

The following is the GAN architecture:

Discriminator Network

The first layer is input layer with the shape of 300 and adding a dense layer with a number of neurons 12,288 connected to layer 1.

The second layer is an input layer with the shape (64, 64, 3) and takes input as image. Then a concatenate layer is added which is connected to both second input layers [27].



Fig. 5 GAN model

Table 2 Parameters for GAN

Parameters	
Total Params	39,587,841
Trainable Params	39,585,153
Non-trainable	2688

Table 3 Parameters for generator

Parameters	
Total	6, 577, 923
Trainable	6,577,539
Non-trainable	384

Then Conv_2d layers is added to the network with (64, 64, 64, 6). A leaky Re-Lu is then added along with the Gaussian Noise layer.

Then six sets of Conv2D Batch Normalization and leaky Re-Lu layers are added to the network with the increasing filter size.

Finally, we add Flatten layer to convert it into array and add Dense layer with 1024 neurons.

The binary cross-entropy is used to calculate the discriminator loss function.

The following are the statistics of the Discriminator network shown in Table 2.

First, we start with 64 filters and go up to 512 filters with a kernel size of 3, 3 [28].

4.5.1 Generator Network

The generator consists of two input layers one for text and the other for noise. A dense layer is added to it with 8192 neurons. Four sets of resnet networks along with Conv2DTranspose and Leaky Re-Lu are added to the network and are shown in Table 3 [29].

4.5.2 Work Flow Diagram: Shown in Fig. 6 [3]

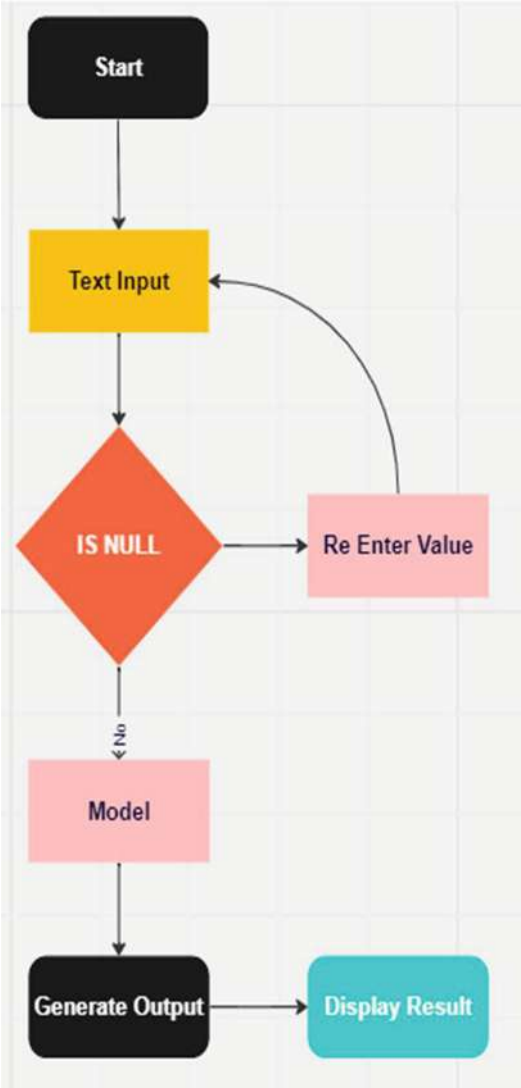
4.5.3 Results

We have trained the model for 960 epochs with the Adam optimizer for the discriminator and generator with a learning rate of 0.000035 and with beta value of 0.5.

Most of the synthesized images do depict plausible colors and shapes of bids, and there does seem to be a lot of diversity. There were some minor mode collapse problems when generating images based on made-up captions.

Figure 7 is obtained by randomly selecting the captions from the dataset and calculating the distance between them. As we can see only minor mode collapse issues were found on this algorithm.

Fig. 6 Workflow diagram



We have integrated the model with a web application based on flask and deployed it on a personal server [30].

Comparison with GAN-CLS

The main objective of our model is to reduce the mode collapse problem. So, we can use regularization term to update the weights of the old GAN-CLS architecture to the calculated weights.

Figure 8 is the result from GAN-CLS algorithm on the text caption “this small bird has a yellow breast, brown crown, and black superciliary.” we could say that



Fig. 7 Results from mode seeking GANCLS



Fig. 8 Gan-cls mode collapse problem

the model has experienced mode collapse problem since the similarity between the images is very high [1, 31].

The above Fig. 9 is the result for our proposed algorithm.

5 Web Application

We have used technologies like Flask, Python, Html and CSS to build a real-life application using the trained model.

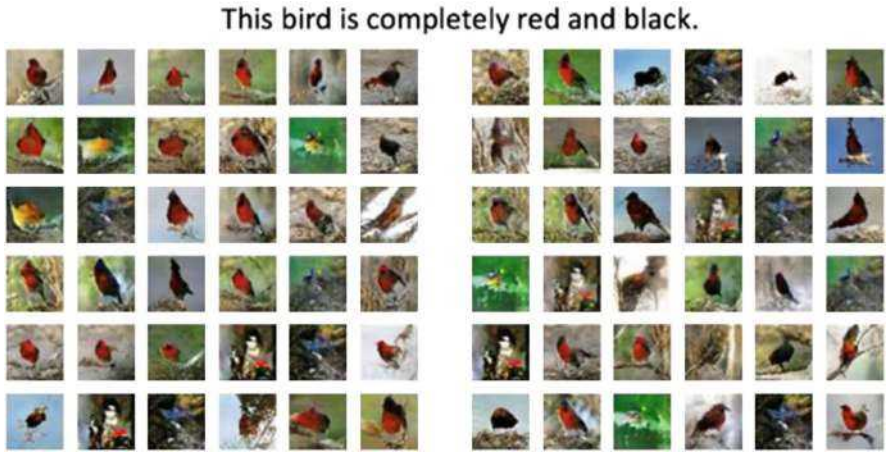


Fig. 9 Result for our proposed algorithm



Fig. 10 Home screen

User Interface

Figure 10 shows the main page of our project where the user can give the text. Figure 11 is the result page of our project, this is the page where the user can see the results and download the images [32].

5.1 Result Analysis

Figure 12 shows the images present in the dataset that contains images of around 200 species and 11,788 bird.



Fig. 11 Result page



Fig. 12 Bird images from dataset

Figure 13 shows the accuracy graph for bird’s dataset over a period of epochs.
Figure 14 shows the loss graph for the bird’s model.
Figure 15 is the output for the text input, there are three texts as seen in the above image.

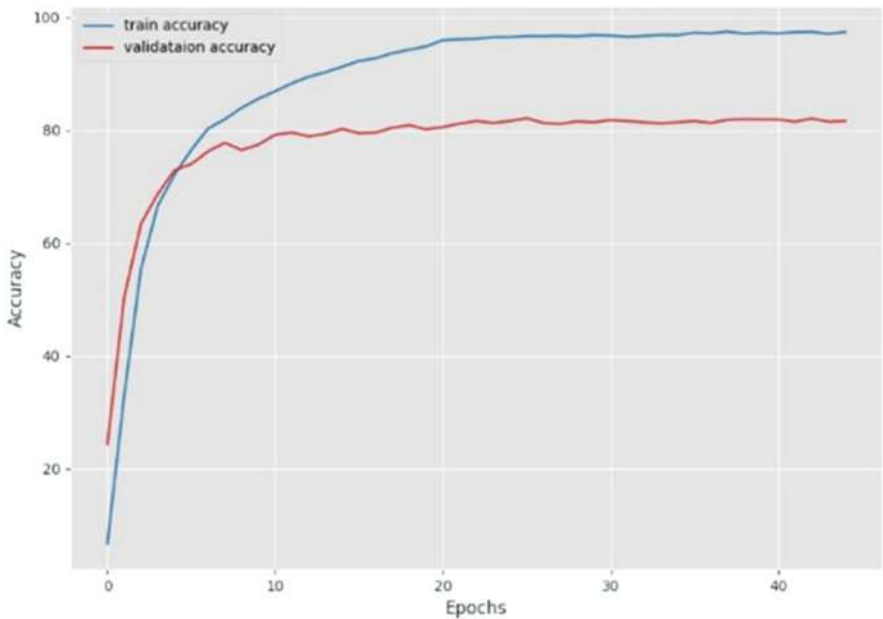
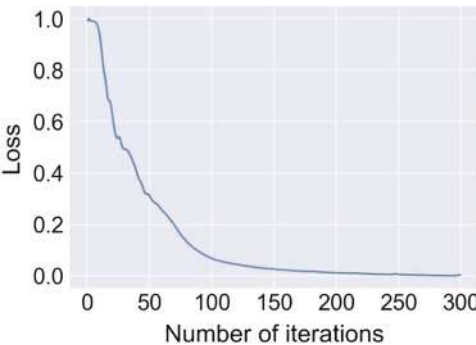


Fig. 13 Accuracy graph for bird’s label

Fig. 14 Loss graph for the model



Inspection Score Is Shown in Table 4

The main purpose of this project is to generate the images from text. Although many researchers have trained their model based on the preprocessed embeddings as we wanted a real-world application, we have followed the traditional methods of preprocessing to manually obtain the embeddings for the given text. After preprocessing we have used Word2Vec model to generate the embeddings and we have stored them in a pickle file for training. As for the user text, we dynamically process the text input and convert it into embeddings.

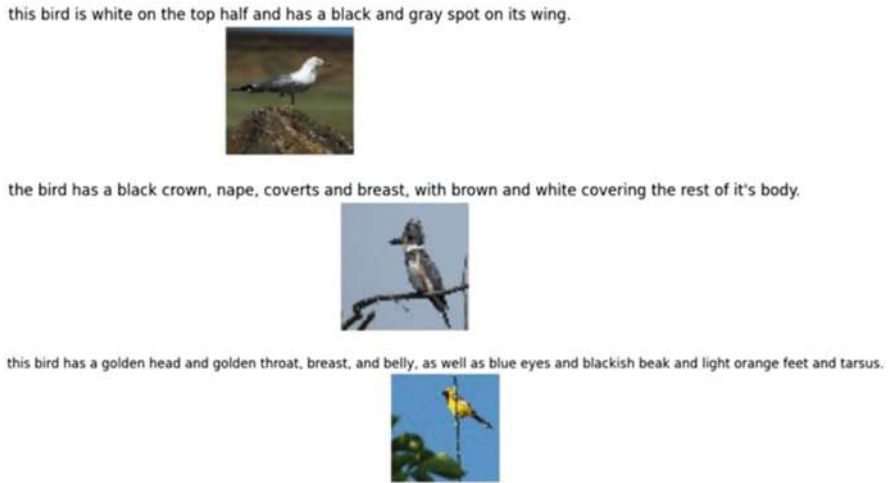


Fig. 15 Bird prediction for given text input

Table 4 Inspection score comparison

Dataset—BIRDS	
Algorithms	Inspection score
GAN-CLS	49 ± 01
MSGAN	54 ± 01
PROPOSEDALGORITHM	55 ± 02

And also, we were interested in reducing the mode collapse problem on the GAN- CLS algorithm so we have combined the method of regularization term from MSGAN algorithm and have proposed a new architecture for the GAN [33].

We trained the model for 960 epochs, and it sometimes had small mode collapse issues.

6 Conclusion

We have trained the model for 960 epochs with the Adam optimizer for the discriminator and generator with a learning rate of 0.000035 and with beta value of 0.5. Most of the synthesized images do depict plausible colors and shapes of bids, and there does seem to be a lot of diversity. There were some minor mode collapse problems when generating images based on made up captions.

References

1. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016, June). Generative adversarial text to image synthesis. In *International conference on machine learning* (pp. 1060–1069). PMLR.
2. Qi Mao, Hsin-Ying Lee, Hung-Yu Tseng, Siwei Ma, and Ming-Hsuan Yang. Mode Seeking Generative Adversarial Networks for Diverse Image Synthesis, *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
3. Akhtar, N., Rahman, S., Sadia, H., & Perwej, Y. (2021). A holistic analysis of Medical Internet of Things (MIoT). *Journal of Information and Computational Science*, 11(4), 209–222.
4. Sharma, S., Suhubdy, D., Michalski, V., Kahou, S. E., & Bengio, Y. (2018). Chatpainter: Improving text to image generation using dialogue. *arXiv preprint arXiv:1802.08216*.
5. Zhang, Jiong, et al. “Research on an interactive question answering system of artificial intelligence customer service based on Word2Vec.” *International Journal of e-Collaboration (IJeC)* 18.2 (2022): 1–12.
6. Akanksha Singh, Sonam Anekar, Ritika Shenoy, Sainath Patil, 2021, Text to Image using Deep Learning, *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)* Volume 10, Issue 04 (April 2021)
7. Zhang, C., & Peng, Y. (2018, September). Stacking vae and gan for context-aware text-toimage generation. In *2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM)* (pp. 1–5).
8. T. Hinz, S. Heinrich, and S. Wermter. Semantic object accuracy for generative text-to-image synthesis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (01):1–1, Sep 5555.
9. Jifeng Wang, Xiang Li, and Jian Yang. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1788–1797, 2018.
10. Tao Xu, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. Attngan: Fine- grained text to image generation with attentional generative adversarial networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1316–1324, 2018.
11. Lee, Minhyeok, and Junhee Seok. “Regularization methods for generative adversarial networks: An overview of recent studies.” *arXiv preprint arXiv:2005.09165* (2020).
12. Zhang, Jiong, et al. “Research on an interactive question answering system of artificial intelligence customer service based on Word2Vec.” *International Journal of e-Collaboration (IJeC)* 18.2 (2022): 1-12.
13. Zhang, Zijun. “Improved adam optimizer for deep neural networks.” *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)*. IEEE, 2018.
14. Gong, Fuzhou, and Zigeng Xia. “Generate the corresponding Image from Text Description using Modified GAN- CLS Algorithm.” *arXiv preprint arXiv:1806.11302* (2018).
15. Hong, Seunghoon, et al. “Inferring semantic layout for hierarchical text-to- image synthesis.” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
16. Zhang, Han, et al. “Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks.” *Proceedings of the IEEE international conference on computer vision*. 2017.
17. Chauhan, Rahul, Kamal Kumar Ghanshala, and R. C. Joshi. “Convolutional neural network (CNN) for image detection and recognition.” *2018 first international conference on secure cyber computing and communication (ICSCCC)*. IEEE, 2018.
18. Zhu, Minfeng, et al. “Dm-gan: Dynamic memory generative adversarial networks for text-to-image synthesis.” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
19. Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. “Understanding of a convolutional neural network.” *2017 international conference on engineering and technology (ICET)*. IEEE, 2017.

20. Fan, Engui. "Extended tanh-function method and its applications to nonlinear equations." *Physics Letters A* 277.4–5 (2000): 212–218.
21. Sperl, Philip, et al. "DLA: dense-layer- analysis for adversarial example detection." 2020 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, 2020.
22. Shrestha, Ajay, and Ausif Mahmood. "Review of deep learning algorithms and architectures." *IEEE access* 7 (2019): 53040-53065.
23. Miyato, Takeru, and Masanori Koyama. "cGANs with projection discriminator." *arXiv preprint arXiv:1802.05637* (2018).
24. Dittmer, Sören, Emily J. King, and Peter Maass. "Singular values for ReLU layers." *IEEE transactions on neural networks and learning systems* 31.9 (2019): 3594–3605.
25. Wang, Yaohui, et al. "Latent image animator: Learning to animate images via latent space navigation." *arXiv preprint arXiv:2203.09043* (2022).
26. Kukačka, Jan, Vladimir Golkov, and Daniel Cremers. "Regularization for deep learning: A taxonomy." *arXiv preprint arXiv:1710.10686* (2017).
27. Agarwal, Mohit, Suneet Gupta, and K. K. Biswas. "A new Conv2D model with modified ReLU activation function for identification of disease type and severity in cucumber plant." *Sustainable Computing: Informatics and Systems* 30 (2021): 100473.
28. Murray, Naila, and Florent Perronnin. "Generalized max pooling." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
29. Jeczmiónek, Ernest, and Piotr A. Kowalski. "Flattening layer pruning in convolutional neural networks." *Symmetry* 13.7 (2021): 1147.
30. Zeiler, Matthew D. "Adadelta: an adaptive learning rate method." *arXiv preprint arXiv:1212.5701* (2012).
31. Wang, Yan, et al. "3D conditional generative adversarial networks for high- quality PET image estimation at low dose." *Neuroimage* 174 (2018): 550–562.
32. Ryan Kiros, Yukun Zhu, R. Salakhutdinov, R. Zemel, R. Urtasun, A. Torralba, and S. Fidler. Skip-thought vectors. In *NIPS*, 2015.

Smart Agricultural Greenhouse System: A Context-Aware Application



Sujata Swain and Rajdeep Niyogi

Abstract Smart applications are getting more powerful and cheaper cost due to the advancement in sensor technology. In this chapter, we have considered a smart greenhouse application. The important parameters of the smart greenhouse are to control light, temperature, humidity, and water. Greenhouse can provide a favourable growth environment for plants. This smart greenhouse system is a context-aware application. When the context changes in the environment, the application automatically adapts it and does the changes seamlessly without user intervention. In this chapter, we propose a model for the smart greenhouse system. The proposed model can adapt to context changes such as temperature rises, and it automatically turns on the AC seamlessly. The proposed model is based on the machine learning approach. The model maintains the greenhouse temperature, humidity, and sunlight duration in the suitable ranges of the plant.

Keywords Greenhouse · Machine learning · Sensors · NLP · Web service · Context

1 Introduction

A greenhouse is a structure, typically made of glass or transparent material, designed to create an environment suitable for the cultivation of plants. It is used to extend the growing season, protect plants from harsh weather conditions, and provide optimal conditions for plant growth. They can be used for a wide range of purposes, including growing flowers, vegetables, fruits, and even exotic plants that require

S. Swain (✉)

School of Computer Engineering, Kalinga Institute of Industrial Technology, Bhubaneswar, India
e-mail: sujata.swainfcs@kiit.ac.in

R. Niyogi

Department of Computer Science, IIT Roorkee, Roorkee, India
e-mail: rajdpfec@iitr.ac.in

specific climatic conditions. The controlled environment in a greenhouse allows for year-round production of crops, regardless of the external weather conditions.

Context refers to some information about a user and/or the environment. According to [1], “Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including location, time, activities, and the preferences of each entity”. When the context changes in the environment, the application automatically adapts it and does the changes seamlessly without user intervention.

Consider a situation in the greenhouse system where the temperature rises. To cool it down, the application automatically turns on the AC. Here, the context is temperature, and when the context changes, the devices (AC) will on or off accordingly.

In the world, India is the second most populous country, and agriculture in India is the largest economic sector. It plays a very important role in the overall socio-economic fabric of India. Due to this, the agriculture industry pursues a common goal, i.e. to maximize the agricultural output while maintaining quality. Agriculture is also at high risk due to the unpredictable environment. Farmers also commit suicide due to some conflicting reasons, such as high debt burdens, monsoon failure, and low yields. In India, 11.2% of all the suicide cases are of farmers [2]. To improve this situation, there is a need for a system in which farmers will get information related to agriculture and climate a priori, which means yield will be higher, farmers will get high profit, and crop production will be maximum.

Computer technology advances agriculture by using the Internet of Things (IoT), Web service [3–5], and image analysis. Agri-Info [6], Agri-IoT [7], and AgroMobile [8] systems are developed to assist farmers. In Agri-Info [6], the system is based on cloud and service technologies, and it delivers information related to agriculture. In Agri-IoT [7], the system is based on the Internet of Things (IoT), and it provides the reasoning for the various sensor data streams. In AgroMobile [8], the system is based on mobile cloud computing, and it assists farmers in better cultivation and marketing.

The climate greatly affects agricultural production and also depends upon the geographical location. Different types of plants require different amounts of temperature, humidity, and sunlight wavelength. Due to lack of awareness, this information can cause plants to die. However, with the emergence of greenhouses, farmers can cultivate different types of plants in different climates, and they need not depend upon the geographical location. Due to the closed structure of the greenhouse, insects, and pests cannot enter inside or minimally enter, thereby eliminating the requirement for insecticides and pesticides. Farmers can adjust temperature, humidity, and sunlight wavelength corresponding to the requirements of the type of plant. The greenhouse is a solution to cultivate the crops anytime, anywhere [9]. To control the temperature, air conditioner (AC) and heater will automatically turn on/off. To control the humidity, humidifier and dehumidifier will automatically turn on/off. To give sunlight to the plants, the bulb will turn on/off. However, controlling the climate of an agricultural greenhouse corresponding to the

given suitable range of the parameters of different crops and maintaining suitable range parameters has become a difficulty. Various monitoring systems are applied to the agricultural greenhouse, due to the advancement of computer technology and sensor technology. According to [10], monitoring systems are mostly wired where drawbacks occur such as inconvenient post-maintenance, massive direct investment, and heavy duty. There is a need for a context-aware system which can adjust the greenhouse parameters when the context changes. Here, contexts are the sensor inputs which are temperature, humidity, and sunlight duration in the greenhouse environment. Context changes mean the temperature rises and falls, humidity rises and falls, and the sunlight duration. The context-aware system collects the sensors' generated data. It analyses, interprets, and understands these data [11].

In this chapter, we propose a model for a smart greenhouse system as a context-aware application. When context changes, the proposed model controls the devices associated with it. This model is based on the machine learning approach. It controls the environmental parameters.

The organization of this chapter is as follows: Sect. 2 discusses related works. In Sect. 3, proposed methodology is discussed where their experimental results are shown in Sect. 4. At the end, we conclude this chapter in Sect. 5.

2 Related Work

In this section, we present the related works of the existing agriculture systems.

In [12], the proposed system collects the greenhouse environmental information as well as the humidity and temperature of the leaves of the crop. The system examines the relationship between crop growth status, environmental information, and crop disease. Crop disease depends upon the humidity lasting time on leaves and the temperature of leaves. The system prevents crop disease. In [13], a context-aware middleware is proposed based on wireless sensor network, which efficiently processes data collected from greenhouse. The proposed middleware integrates different forms of data collected from heterogeneous sensors and provides intelligent context aware, event service to maximize the operability and scalability of the middleware. In [14], a ubiquitous sensor network middleware for the agricultural environment is proposed. It collects information related to the greenhouse's environment. The proposed middleware is composed of a context manager, sensor manager, and control manager. The context manager processes the collected heterogeneous sensor data into information which is suitable to the user's demand. The sensor manager sends it to the control manager. The control manager provides different services. It makes decisions which are suitable to the conditions. In [15], a wireless sensor network middleware is proposed to overcome the limits of wire connection. It processes the collected data, analyses the data, and provides intelligent diagnosis service and event service. It guides the agriculture production method, such as irrigation management and malady hindrance.

In [16], a monitoring system is proposed. The proposed system consists of the control centre, acquisition and transmission system, and mobile control terminal. It collects environmental information in real time. It transmits the information to the control centre, and it provides the remote control through the mobile device. In [17], a smart greenhouse helps farmers, which provide information related to the climate automatically. It saves the use of manual inspection. In [18], a farm management system architecture is proposed which focuses on the process of farming and the process of exchanging information between stakeholders. In [19], a Web-based agricultural support system is proposed which identifies functionalities such as decision support or expert system, collaborative work support, and information support. The basic characteristics of the proposed system are research-education subsystem, production subsystem, and agricultural management subsystem. In [20], a farm management information system is studied. It analyses the farmers' dynamic needs and improves the decision process. In [21], a cloud-based weather forecasting system is proposed, which collects and analyses the information related to the weather and identifies the farming needs of various seasons.

Many research works[22–25] have been published for automated disease detection in plant leaves based on image processing and machine learning such as using pattern recognition techniques[26], support vector machine [27], digital image processing techniques [28], and computer vision [29]. In [30], a deep learning method is used to develop a plant disease identification system. They used the Convolutional Neural Network (CNN) method to train their model. The model achieved an accuracy of 99.35%. However, the model performed poorly when it was tested on a set of images taken under different conditions. In [31], the authors proposed a model to detect the disease in leaves. They considered the meta-architectures which are the combination of R-CNN, R-FCN, SSD faster region-based convolutional neural network (faster R-CNN), region-based fully convolutional network (R-FCN), and single shot multibox detector (SSD). They detected the nine diseases effectively. However, the accuracy is only 83%. In [32], they proposed a model to identify the plant leaf disease. However, the proposed model provides reliable results in the context of limited data.

3 Proposed Approach

In this section, we present the proposed smart greenhouse system. The algorithm for the context changes in smart greenhouse system is given in Sect. 3.1. The service provided by the smart greenhouse system is described in Sect. 3.2. A brief description of data preprocessing is given in Sect. 3.3. The flowchart of the proposed system is given in Sect. 3.4.

The proposed system is basically developed to automatically adjust the greenhouse environment parameters. Greenhouse is a closed structure, and it provides an artificial environment for the plant, which helps farmers to cultivate the plant at anytime and anywhere. The important parameters of the greenhouse environment

are temperature, humidity, and sunlight duration. Different species of plants require a different range of temperature, humidity, and sunlight. Sunlight is required for photosynthesis. The different hardware components are required for adjusting temperature, humidity, and sunlight duration. Air conditioners and heaters are used in the greenhouse for adjusting the temperature, and humidifier and dehumidifier are used to adjust humidity, and the bulb is used to provide sunlight.

When the temperature rises from the suitable ranges, the AC automatically turns on and the heater turns off. When the temperature falls from the suitable ranges, the AC automatically turns off and the heater turns on. When the humidity rises from the suitable ranges, the dehumidifier automatically turns on and the humidifier turns off. When the humidity falls from the suitable ranges, the dehumidifier automatically turns off and the humidifier turns on. The plant requires a certain duration of sunlight. The bulb turns on during that duration, otherwise, the bulb turns off automatically.

3.1 Algorithm

When the context changes in the smart greenhouse system, the system sends commands for the device, which is given in Algorithm 1.

Algorithm 1 Algorithm for context change in smart greenhouse application

```

while Context Changes do
  if Temp > Suitable_Range then
    Command: AC will On for a Period
  else if Temp < Suitable_Range then
    Command: Heater will On for a Period
  end if
  if Humidity > Suitable_Range then
    Command: Dehumidifier will On for a Period
  else if Humidity < Suitable_Range then
    Command: Humidifier will On for a Period
  end if
  if Time lies in Daytime then
    Command: Bulb will On.
  else if Time lies in Nighttime then
    Command: Bulb will Off
  end if
end while

```

3.2 Smart Greenhouse System

The services provided by the smart greenhouse system are as follows:

1. Recommendation of fertilizers: To cure disease or to kill insects, fertilizers are recommended by the system.
2. Replacement of fertilizer using functionally equivalence method: If the required fertilizer is not available, then the functionally equivalent fertilizers are provided.
3. Weather forecast: Farmers get the weather information a prior.
4. Machinery information: Farmers get the machinery information (when and how to use it) in the field.
5. Government facility regarding loans: Farmers get the bank offers for the loan.
6. Leaf disease detection: Leaf disease identification using image processing.
7. Harmful insect detection: Insect identification using image processing.
8. Emergency case scenario like fire, animals nearby field: If the animals are present nearby field or fire is in a nearby field, then it gives an alarm or notification to the system.

3.3 Data Preprocessing

The greenhouse environment parameter data is collected from different sensors presented in the greenhouse. These data may contain different types of noise, i.e. unwanted signals, repeated signals, incomplete signals, or system failure. Due to this, the preprocessing method is required to find out the noise. The greenhouse environment parameter data is one dimensional. Therefore, one-dimensional data preprocessing method and applying different types of clustering techniques to remove outliers are required.

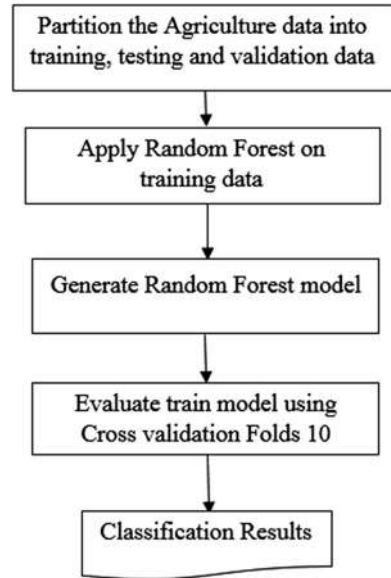
3.4 System Flowchart

The proposed smart greenhouse system is based on the machine learning approach. The flowchart of the system is shown in Fig. 1. The modules used in smart greenhouse system are described below.

1. Partition the data into training, testing, and validation data.
2. Apply all the learning methods to the training data.
3. Generate and evaluate the learning model one by one.
4. Select the best learning model.

The dataset is divided into three subsets, as shown in Fig. 2: training dataset, testing dataset, and validation dataset. Training and testing datasets are also divided into subsets. Each learning algorithm is applied to the training dataset. A learning model is generated after providing the testing dataset. The learning model provides correct classification of the dataset. The validation dataset is applied to the learning model, to find out true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

Fig. 1 System flowchart
block diagram



We have applied six types of learning algorithms. These are random tree (RT), random forest (RF), decision tree (DT), K-NN, support vector machine (SVM), and naive bayes (NB).

4 Experiment and Results

We implement the proposed smart greenhouse system. Experiments are run on the Intel Core i5 with 2.53 GHz machine with 4 GB of RAM. The system is developed with the help of Weka.¹

4.1 Simulation

This section presents the simulation of the proposed system. We have simulated the system on the different plants, like rice, wheat, and apple. Here, we are going to discuss the system on the rice plant.

¹ <https://www.cs.waikato.ac.nz/~ml/weka/>

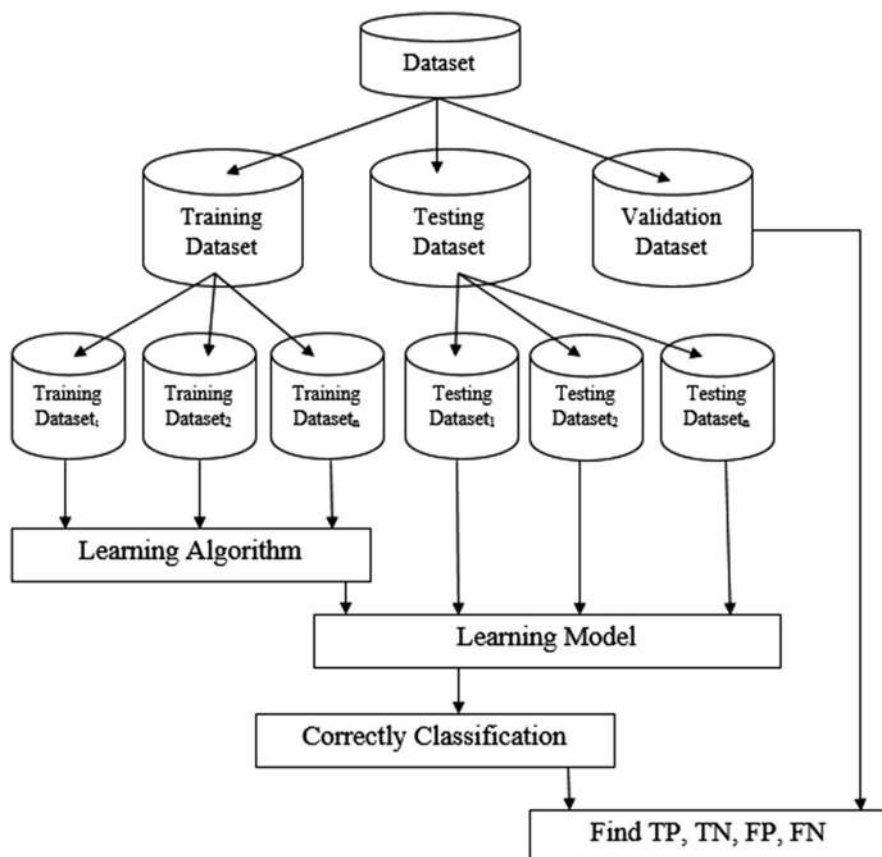


Fig. 2 Machine learning approach

Rice :

Suitable temperature : 26–32 degree Celsius

Suitable humidity : 65\%–70\%

Suitable light : 8000 LUX and

duration of light : 12hrs

duration of darkness : 12hrs .

We have applied six types of machine learning approaches. These are random tree (RT), random forest (RF), decision tree (DT), K-NN, support vector machine (SVM), and naive bayes (NB). Decision tree gives the best result over other classifiers.

5 Conclusion

This chapter proposed a smart greenhouse system as a context-aware application that controls the environmental parameter of an agriculture greenhouse based on machine learning approach. Decision tree gives better classification result over the other machine learning approaches. The proposed system collects the environmental parameters of the greenhouse. It controls the temperature using an air conditioner and heater. It controls the humidity using humidifier and dehumidifier.

References

1. A. K. Dey, "Understanding and Using Context." IN Personal and ubiquitous computing Vol.5(1), pp.4–7, 2001
2. Agriculture in India https://en.wikipedia.org/wiki/Agriculture_in_India
3. S. Swain and R. Niyogi, "Functionally Equivalent Service Composition", in Internet of Things Engineering Cyber Physical Human Systems, Vol. 9 pp:1–10, 2020.
4. S. Swain and R. Niyogi, "ContextAware Service Composition with Functionally Equivalent Services for Complex User Requests", in 34th International Conference on Advanced Information Networking and Applications (AINA 20), Caserta, Italy, April 15–17, 2020.
5. S. Swain and R. Niyogi, "Service Composition in a ContextAware Setting with Functionally Equivalent Services", in 20th International Conference on Computational Science and Its Applications (ICCSA-20), Cagliari, Italy, July 1–4, 2020.
6. S. Singh, I. Chana, and R. Buyya. "Agri-Info: cloud based autonomic system for delivering agriculture as a service." arXiv preprint arXiv:1511.08986 (2015).
7. A. Kamilaris, G. Feng, F. X. Prenafeta-Boldu, and M. I. Ali. "Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications." In IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 442–447, 2016.
8. S. Prasad, S. K. Peddoju, and D. Ghosh. "AgroMobile: a cloud-based framework for agriculturists on mobile platform." In International Journal of advanced science and technology, vol. 5, no. 9, pp. 41–52, 2013.
9. X. Li, Y. Deng, and L. Ding. "Study on precision agriculture monitoring framework based on WSN." In 2nd IEEE International Conference on Anti-counterfeiting, Security and Identification (ASID'08), pp. 182–185, 2008.
10. P. P. Ray. "Internet of things for smart agriculture: Technologies, practices and future direction." Journal of Ambient Intelligence and Smart Environments, vol. 9, no. 4, pp. 395–420, 2017.
11. C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey." in IEEE communications surveys and tutorials, vol. 16, no. 1, pp. 414–454, 2014.
12. D. Park, B. Kang, K. Cho, C. Shin, S. Cho, J. Park, and W. Yang. "A study on greenhouse automatic control system based on wireless sensor network." Wireless Personal Communications, vol.56, no. 1, pp. 117–130, 2011.

13. J. Hwang, and H. Yoe. "Study on the context-aware middleware for ubiquitous greenhouses using wireless sensor networks." *Sensors*, vol. 11, no. 5, pp. 4539–4561, 2011.
14. J. Lee, J. Hwang, and H. Yoe. "Design and Implementation of Middleware for GreenHouse Based on Ubiquitous Sensor Network." *Future Generation Information Technology*, pp. 708–718, 2010.
15. L. Zhao, L. He, X. Jin, and W. Yu. "Design of wireless sensor network middleware for agricultural applications." In *6th Computer and Computing Technologies in Agriculture (CCTA)*, no. Part II, pp. 270–279, Springer, 2012.
16. R. Li, X. Sha, and K. Lin. "Smart greenhouse: A real-time mobile intelligent monitoring system based on WSN." In *IEEE International Conference Wireless Communications and Mobile Computing (IWCMC'14)*, pp. 1152–1156, 2014.
17. R. K. Kodali, V. Jain, and S. Karagwal. "IoT based smart greenhouse." In *IEEE Region 10 Conference on Humanitarian Technology (R10-HTC)*, pp. 1–6, 2016.
18. A. Kaloxylos, R. Eigenmann, F. Teye, Z. Politopoulou, S. Wolfert, C. Shrank, M. Dillinger. "Farm management systems and the Future Internet era." *Computers and electronics in agriculture* vol. 89, pp. 130–144, 2012.
19. Y. Hu, Z. Quan, and Y. Yao. "Web-based Agricultural Support Systems." In *Workshop on Web-based Support Systems*, pp. 75–80, 2004.
20. C. G. Sorensen, S. Fountas, E. Nash, Liisa Pesonen, D. Bochtis, S. M. Pedersen, B. Basso, and S. B. Blackmore. "Conceptual model of a future farm management information system." *Computers and electronics in agriculture* vol. 72, no. 1, pp. 37–47, 2010.
21. R. Pawlak, S. Lefebvre, Z. Kazi-Aoul, and R. Chiky. "Cloud elasticity for implementing an agricultural weather service." In *IEEE 11th Annual International Conference on New Technologies of Distributed Systems (NOTERE'11)*, pp. 1–8, 2011.
22. H. A. Hiary, S. B. Ahmad, M. Reyilat, M. Braik, Z. A. Rahamneh. "Fast and accurate detection and classification of plant diseases." In *International Journal of Computer Application* Vol.17 no.1, pp:31–38, 2011.
23. S. Bashir, N. Sharma. "Remote area plant disease detection using image processing". In *IOSR Journal of Electron. Commun. Eng.* Vol.2 no.6, pp:31–34, 2012
24. A. H. Kulkarni, A. Patil. "Applying image processing technique to detect plant diseases." In *Int. J. Modern Eng. Res.* Vol.2 no.5, pp:3661–3664, 2012.
25. C. E. Too, Y. L. Chebet, N. Sam and Y. Liu. "A comparative study of fine-tuning deep learning models for plant disease identification." *Computers and Electronics in Agriculture* Vol. 161, pp: 272–279, 2019.
26. S. Phadikar, J. Sil. "Rice Disease identification using pattern recognition techniques." In *IEEE 11th International Conference on Computer and Information Technology*, pp. 420–423, 2008.
27. A. Shmilovici. "Support Vector Machines." In: Maimon, O., Rokach, L. (eds) *Data Mining and Knowledge Discovery Handbook*. Springer, Boston, MA.
28. J.G.A Barbedo, L. V. Koenigkan, T.I. Santos. "Identifying multiple plant diseases using digital image processing.", In *Biosyst. Eng.* 147, 104116, 2016
29. A. Asfarian, Y. Herdiyeni, A. Rauf, K. H. Mutaqin. "A computer vision for rice disease identification to support Integrated Pest Management." In *Crop prot.* Vol. 61, pp:103–104, 2013
30. S. P. Mohanty, H. P. David and S. Marcel. "Using deep learning for image-based plant disease detection." in *Frontiers in plant science* Vol. 7, pp: 1419, 2016.
31. A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park. "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition." *Sensors* Vol.17, no. 9, pp: 2022, 2017.
32. J.G.A. Barbedo. "Plant disease identification from individual lesions and spots using deep learning." *Biosystems Engineering* Vol.180, pp: 96–107, 2019

Improving Performance of Plant Disease Detection Using YOLOv7 and YOLOv8



Lakshmi Narayana Chintala , K. Sreerama Murthy,
and Venkata Ramana Kondapalli

Abstract Plant diseases are currently one of the most significant issues influencing modern agricultural production. One of the most important measures for monitoring, interpreting, and predicting how plant diseases may affect crop yield is the disease severity index. The essential plant disease treatment may be prescribed depending on the value of the disease severity, which is critical in minimizing additional yield loss. Traditional methods for diagnosing the severity of plant diseases often involve a skilled expert conducting visual analysis of specimens or plant tissue. However, this procedure is time-consuming, costly, and inefficient. As a result, now is the time to develop disease evaluation methods that will be beneficial in current agricultural production. In this study, YOLOv7 and YOLOv8 object detection methods were utilized to perform multi-class classification of various plant diseases using the PlantDoc dataset, with the goal of reducing losses caused by disease outbreaks. The YOLOv7 model demonstrated superior detection performance and exhibited strong generalization capabilities in complex scenes when evaluated on the PlantDoc dataset. It achieved high scores in metrics such as mean average precision (mAP), F1-score, precision, and recall of 60.99, 57.63, 52.83, and 63.38, respectively. The model's effectiveness and efficiency are better than other object detection models.

Keywords YOLOv7 · YOLOv8 · PlantDoc · Object detection · Plant disease

1 Introduction

Every year, the population of the Earth increases, and with it, the need for all types of plant products increases [1, 16]. To fulfill the rising demand for food quality

L. N. Chintala (✉) · V. R. Kondapalli
Department of CS & SE, AUCE, Andhra University, Visakhapatnam, India
e-mail: dr.kondapallivr@andhrauniversity.edu.in

K. Sreerama Murthy
Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad, India

and quantity, crops must be protected against plant diseases [17]. According to the ICAR, over 35% of agricultural productivity is lost annually owing to pests and diseases [15]. An alarming rise in the number of insect and crop disease outbreaks puts the safety of our food in danger. Aside from endangering food security, these plant diseases also have significant effects on the economy, society, and environment.

Farmers still struggle to recognize plant diseases with accuracy. Besides talking to other farmers or calling the Kisan helpline, they do not have many choices. The ability to recognize diseased leaves requires knowledge of plant diseases. In addition, in most circumstances, a lab infrastructure is required to diagnose a plant disease.

In this study, we examine the viability of scalable and economical plant disease diagnosis using computer vision. Deep convolutional neural networks have made a great development in recent years in computer vision. Even while it can take a long time to train large neural networks, once they are trained, the models can categorize pictures relatively rapidly, making them suited for customers' mobile applications. Using computer vision to identify plant diseases opens up new ways to apply deep learning methods to agricultural problems, which allows improvements in understanding of agriculture, crop yield, and disease management.

Most vision-based systems currently available require high-resolution photos captured against a plain background. However, this approach may not be suitable for the majority of Indian farmers who often rely on low-end mobile devices with natural lighting and varied backdrop conditions. Recognizing the importance of real-time environmental conditions and the presence of significant background noise, our system prioritizes the analysis of images captured under such circumstances. We have specifically developed our solution to address the unique challenges faced by farmers and provide them with accurate and reliable information about plants and agriculture. By focusing on real-time images and catering to the specific needs of farmers, our system aims to offer the highest level of support and assistance in agricultural decision-making processes. In this kind of background, we want to emphasize our two primary contributions: PlantDoc was created, consisting of 2569 photos from 13 different plant species and 27 different classes, benchmarking the dataset that has been carefully selected and demonstrating its value in identifying diseases in an uncontrolled environment.

In this paper, object detection-based methods YOLOv7 and YOLOv8 are proposed to address the challenge of plant disease detection in complex conditions. The YOLOv7 and YOLOv8 object detection techniques were used to analyze the PlantDoc dataset, which was developed in an uncontrolled environment. The performance of these models is analyzed in terms of mean average precision (mAP), precision, recall, and F1-score to choose the best.

For this work, the TensorFlow and Keras libraries are used for training object detection models. Colab provides a notebook service similar to Jupyter, along with parallel computing resources dedicated to training machine learning (ML) models. Google Colab offers free access to Graphics Processing Units (GPUs) for accelerated computation. The GPU model used in Colab is the Tesla-K80, which offers 12.68GB of memory and 78.19GB of disk space.

2 Literature Review

Recent advancements in precision agricultural technologies [3] have resulted in a significant rise in crop production; however, with this gain in crop yield comes the concern of worsening product quality [4]. The quality of agricultural products is adversely affected by plant diseases [5]. Traditional approaches [6] entail a detailed examination of the plant species, but these procedures are time and money-consuming. With the advancement of IoT and machine learning, there is an increasing need to digitize the detection of plant diseases [7].

In [1], they propose a novel dataset (PlantDoc), which is a non-lab dataset that can support deep learning approaches for the detection of plant diseases because no other dataset is currently available that provides images captured in real-world environments. They have also examined the dataset's efficiency in comparison to several pre-trained models. In [8], they used YOLO to identify plant diseases. They also investigated several approaches for improving accuracy and precision.

The PlantDoc dataset [1] and the PlantVillage dataset (PVD) [2] are the only publicly available databases for identifying plant diseases to the best of our knowledge. The PlantVillage dataset curators developed an automated approach for disease identification that used GoogleNet and AlexNet and achieved an accuracy of 99.35%. However, the images in the PlantVillage dataset were collected in laboratory settings, potentially limiting their effectiveness in real-world agricultural scenarios. In contrast, the PlantDoc dataset consists of real-time photos of both diseased and healthy plants. The PlantDoc dataset, on the other hand, contains real-time photos of diseased and healthy plants.

In [9], they developed a plant disease diagnosis strategy that involves training the deep learning model on photos of disease patches irrespective of the crop being diagnosed. In comparison to the conventional crop-disease pair-based strategy, they showed superior performance. In [18], the YOLOv5 deep learning model is employed for the detection of rice leaf diseases in this study. The model is trained on a dataset consisting of images related to four different rice leaf diseases. The training is conducted over 100 epochs, and the model achieves mAP, recall, and precision scores of 0.62, 0.94, and 0.83, respectively.

YOLOv7 and YOLOv8 are the latest detectors in object detection models. These networks are built using trainable bag-of-freebies, allowing real-time detectors to enhance accuracy without raising inference costs. Additionally, extend and compound scaling are used by the target detector to significantly increase detection time by efficiently reducing the number of parameters and calculations [12]. At present, YOLOv7 and YOLOv8 are the latest detectors and have not been used to detect plant diseases. Therefore, in the present work, YOLOv7 and YOLOv8 were used to detect plant diseases, which gives the best and most efficient results of accuracy in the detection that is not been done before.

3 Methodology

To create a framework for detecting plant diseases, we should complete data collecting, training model, and multi-class categorization of plant species. The dataset used in this study consists of images with different resolutions, which were standardized to the same dimensions using the input layer of the deep learning model. Once the dataset was prepared, the data was split into a training set comprising 90% of the samples and a testing set containing the remaining 10%. Then, using that data, YOLOv7 and YOLOv8 are trained for plant disease detection. The trained model’s performance is assessed based on various parameters. The flowchart of training and detection of disease is shown in Fig. 1.

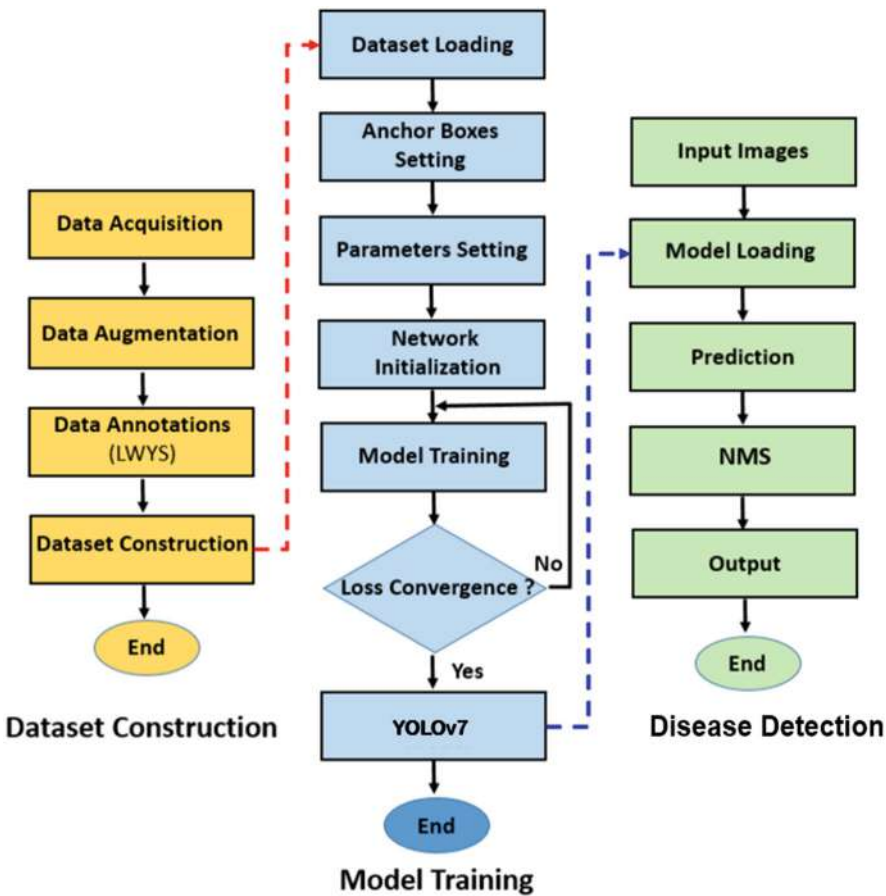


Fig. 1 Flowchart for plant disease detection

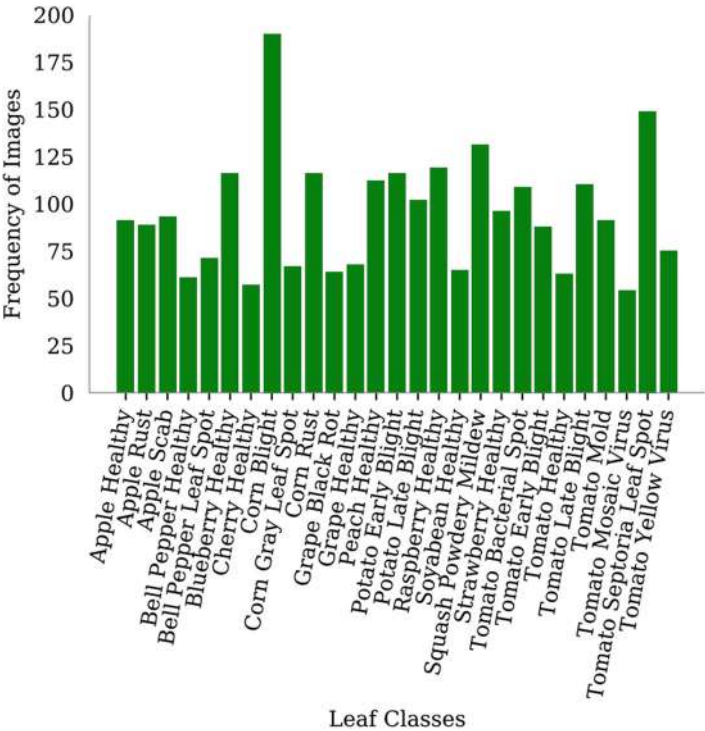


Fig. 2 Distribution of PlantDoc images

3.1 PlantDoc Dataset

This PlantDoc dataset [1] has 2569 images from 13 plant species, and it is divided into 27 classes (healthy and diseased) and contains 8851 labels for object detection. There are several bacterial, fungal, and viral illnesses in the diseased classes that affect food crops including cherry, strawberry, tomato, bell pepper, grape, etc. as shown in Fig. 2.

3.2 You Only Look Once (YOLOv7)

The application of YOLOv7 in this study enabled faster and more accurate object detection through a neural network-based technique. YOLOv7 divides the image into multiple grids of predefined sizes, with each grid responsible for detecting objects within its boundaries. Predicted bounding boxes (PBB) and confidence scores are assigned to each grid cell in the YOLO model. These confidence scores indicate the model’s certainty in detecting an object and provide an estimate of

detection accuracy [10]. Non-maximal suppression is incorporated in YOLOv7 to address redundant predictions by disregarding bounding boxes with lower probability values. The technique selects the bounding box with the highest probability as the reference and suppresses overlapping bounding boxes based on the intersection over union (IoU) metric. The process continues until a final set of accurate bounding boxes for object detection is obtained [11, 12]. YOLOv7 combines several bag-of-freebies techniques to enhance detection accuracy while reducing computational resources. The architecture of YOLOv7 is highly efficient and utilizes the efficient layer aggregation network (E-ELAN) to enhance learning capacity [13]. The paper introduces a compound scaling technique that improves average precision (AP) while reducing parameter count and enhancing computational efficiency. However, limited research has been conducted on the specific application of the YOLOv7 model for disease detection. The architecture of YOLOv7 is shown in Fig. 3.

3.3 *You Only Look Once (YOLOv8)*

YOLOv8 [19] is the most recent detection model in the YOLO family, which is noted for its object-detecting capabilities. The architecture is similar to YOLOv7 in that it has a backbone, head, and neck with improved convolution layers and detection head, which makes this an ideal choice for real-time plant disease detection. The most recent computer vision algorithms, including instance segmentation, which enables multiple-item detection in an image or video, are also supported by YOLOv8. The backbone network used by the model is Darknet-53, which is quicker and more accurate than the prior YOLOv7 network. A detecting head without anchors is used by YOLOv8 to predict bounding boxes. The model's upgraded convolutional network and larger feature map, which increase its precision and speed, make it more efficient than prior iterations. Pyramid feature networks are another addition to YOLOv8 that help it recognize objects of various sizes. In addition to having a user-friendly API, the model is easy to integrate into a variety of applications. The architecture of YOLOv8 is shown in Fig. 4.

4 Results and Discussion

The predictions generated by the classification methods are summarized in the confusion matrix. The classification technique's confusion matrix displays the values for true negatives (TN), true positives (TP), false positives (FP), and false negatives (FN).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

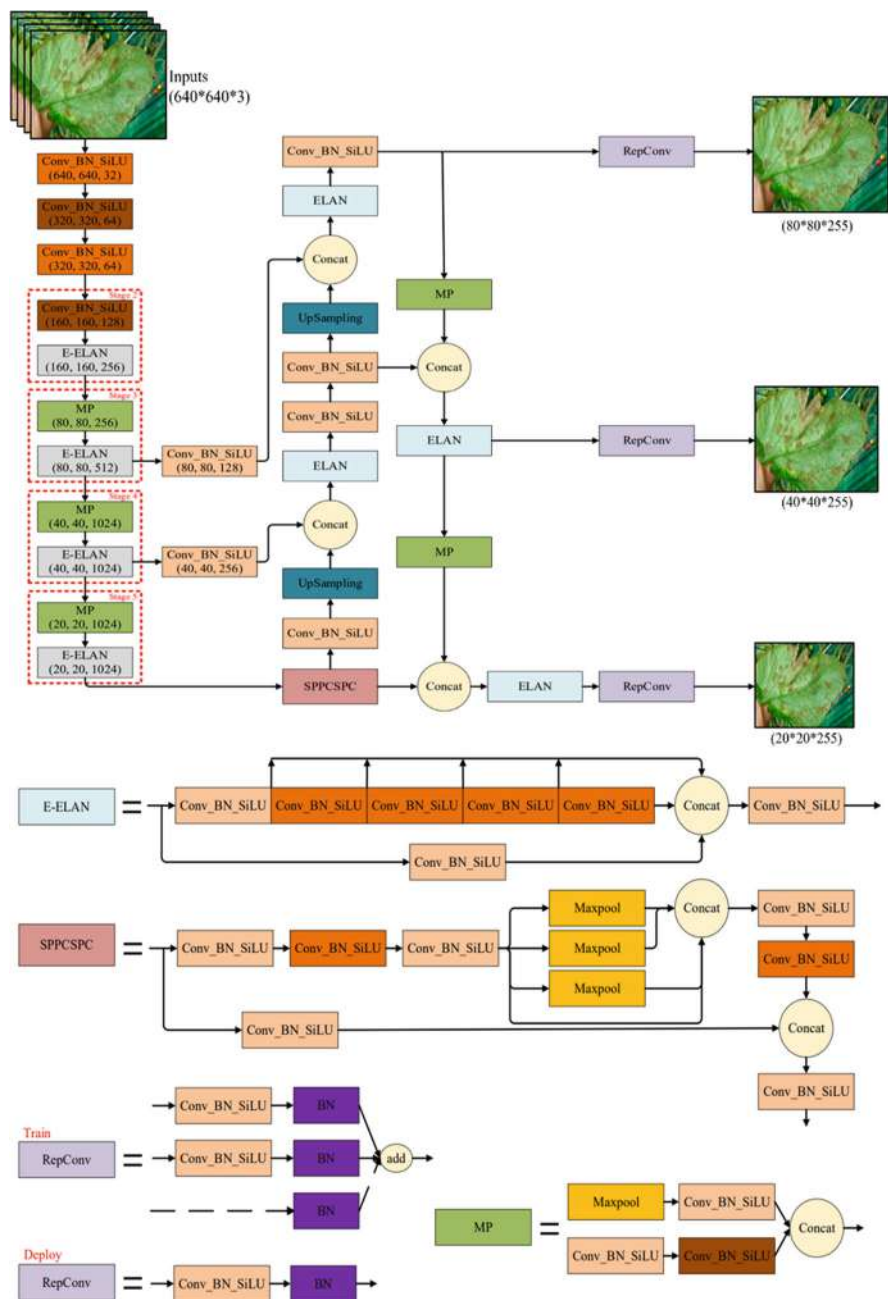


Fig. 3 Architecture of YOLOv7

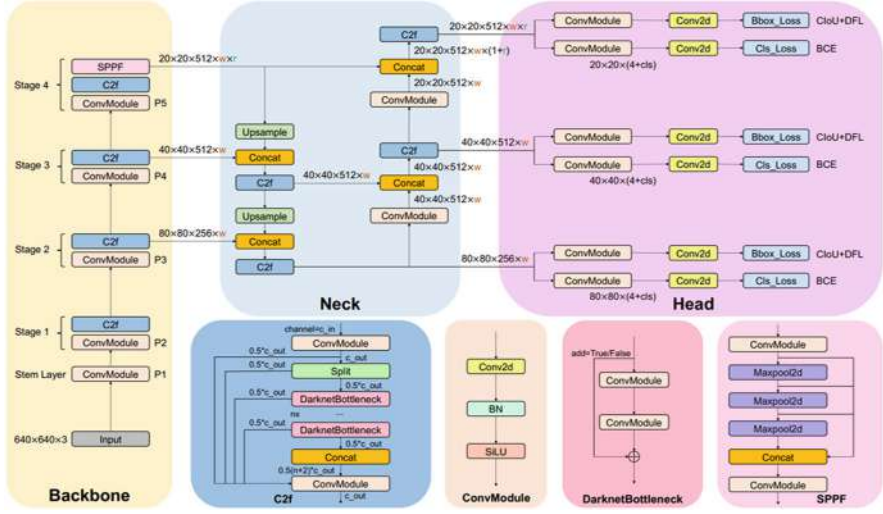


Fig. 4 Architecture of YOLOv8

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{F1Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

The prediction of the bounding box's accuracy is evaluated by intersection over union (IoU). The equation below calculates the intersection over union (IoU), which represents the proportion of the area shared by the predicted bounding box (PBB) and the ground truth bounding box (TBB) in the sample images.

$$\text{IoU} = \frac{\text{area}(\text{PBB} \cap \text{TBB})}{\text{area}(\text{PBB} \cup \text{TBB})} \quad (4)$$

The average precision at IoU (AP^{IoU}) is determined by calculating the area under the precision-recall curve for a specific IoU threshold [14]. This metric serves as a performance indicator for a particular class or category. To assess the overall performance across all detection classes, the mean average precision at a threshold IoU (mAP^{IoU}) is computed and represented as follows:

$$\text{mAP}^{\text{IoU}} = \frac{1}{N} \sum \text{AP}_n^{\text{IoU}}; n \in \{\text{class1}, \text{class2}, \dots, \text{classN}\} \quad (5)$$

The effectiveness of the YOLOv7 and YOLOv8 models is assessed using the PlantDoc dataset. Results indicate that YOLOv7 achieves a higher mean average precision (mAP) compared to previously trained models, including YOLOv8.

Table 1 Comparison of leaf detection mAP

Model	Dataset	mAP (at 50% IoU)	Inference speed
MobileNet [1]	PlantDoc	32.8	–
Faster-rcnn-inception-Resnet [1]	PlantDoc	38.9	–
Proposed YOLOv7	PlantDoc	60.99	8.6 ms
Proposed YOLOv8	PlantDoc	57.00	7.1 ms

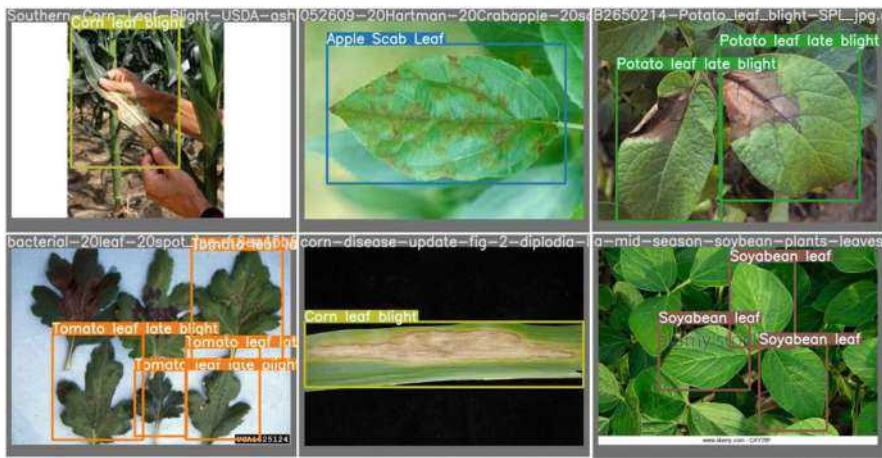


Fig. 5 Detection of a specific class on test images

Furthermore, YOLOv8 demonstrates faster inference speed than YOLOv7, as illustrated in Table 1.

The YOLOv7 and YOLOv8 models were evaluated using randomly selected testing images, and the outcomes are presented in Fig. 5. The experimental findings demonstrate the successful identification of plant diseases by the models.

After training for 50 epochs, YOLOv7 achieved an mAP of 60.99, an F1-score of 57.63, a precision of 52.83, and a recall of 63.38. Similarly, YOLOv8 achieved an mAP of 57.00, an F1-score of 57.65, a precision of 63.20, and a recall of 53.00. The graphs of these metrics are shown in Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15.

5 Conclusion

The process of detecting and classifying plant diseases using digital imagery is difficult. As a result, rapid detection of plant diseases is critical for farmers and plant pathologists to take the appropriate measures. The proposed approach uses a PlantDoc dataset of 27 distinct types of plant leaves and disease photos to detect plant diseases. The object detection models YOLOv7 and YOLOv8 are trained using images from the PlantDoc dataset. The proposed YOLOv7 model achieved

Fig. 6 mAP for YOLOv7

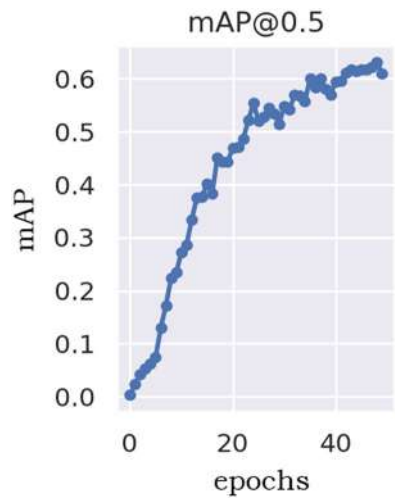
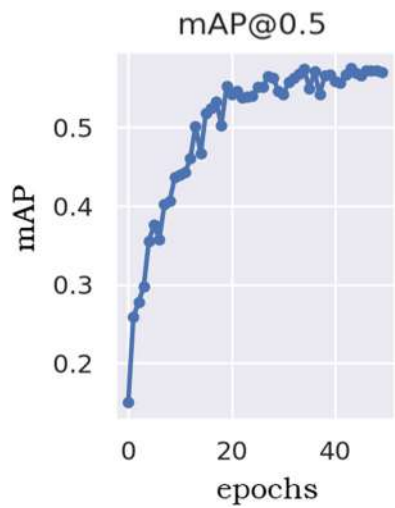


Fig. 7 mAP for YOLOv8



the highest mAP of 60.99% and has demonstrated its efficacy in detecting plant diseases and classifying them. It can be used for early disease prediction, reducing crop failure, and increasing agricultural industry revenues.

In future work, the plant leaf dataset will be diversified by including additional plant leaf images. It will strengthen the ability of models to make more accurate predictions in challenging situations. Future research in this area aims to focus on the advancement of enhanced algorithms for the efficient and rapid detection of diseased leaves.

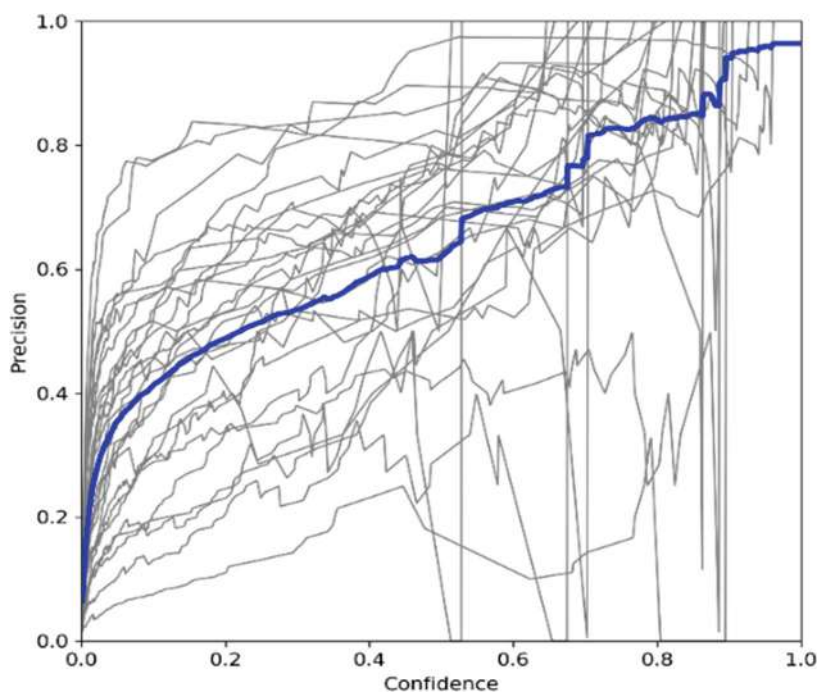


Fig. 8 Precision for YOLOv7

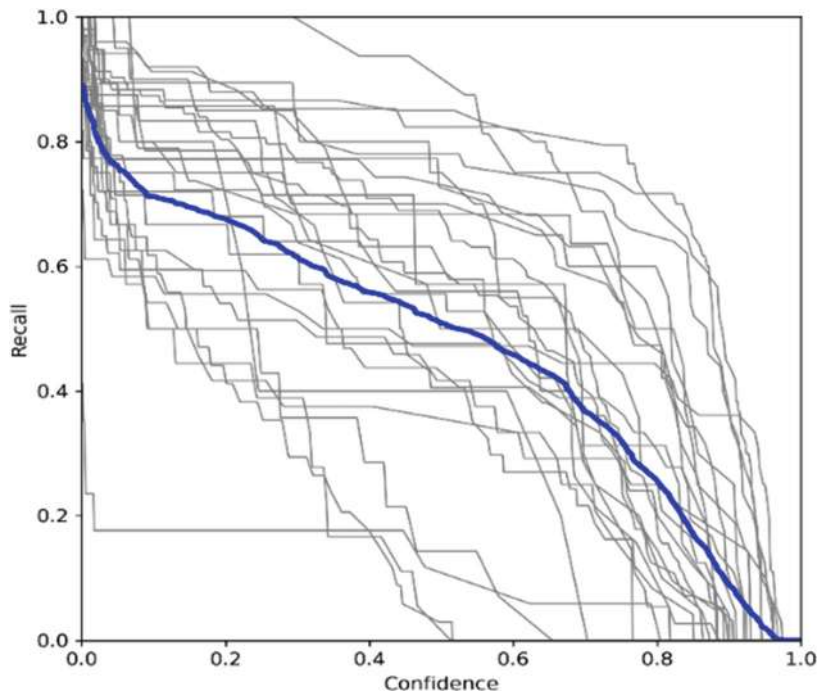


Fig. 9 Recall for YOLOv7

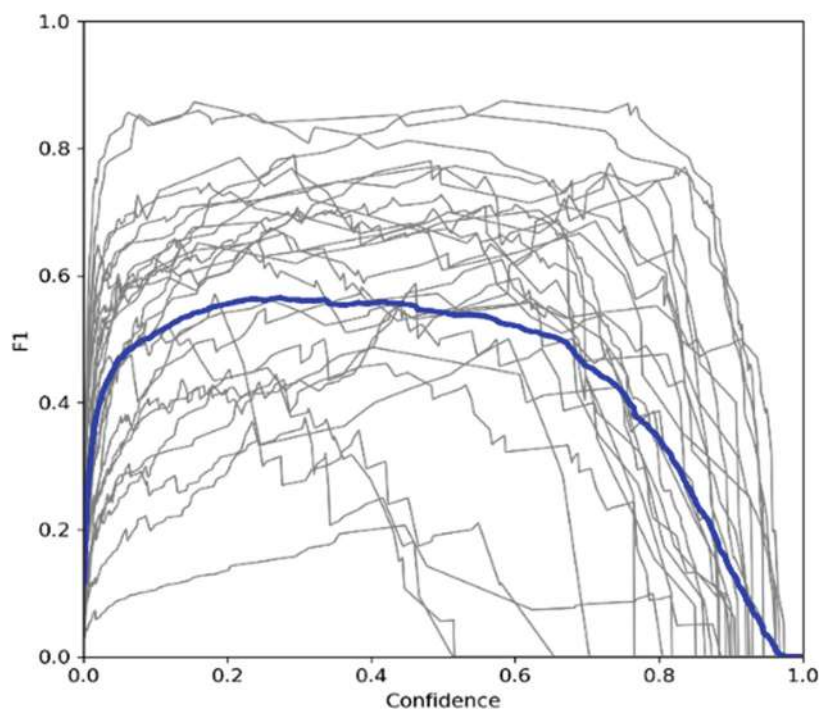


Fig. 10 F1-score for YOLOv7

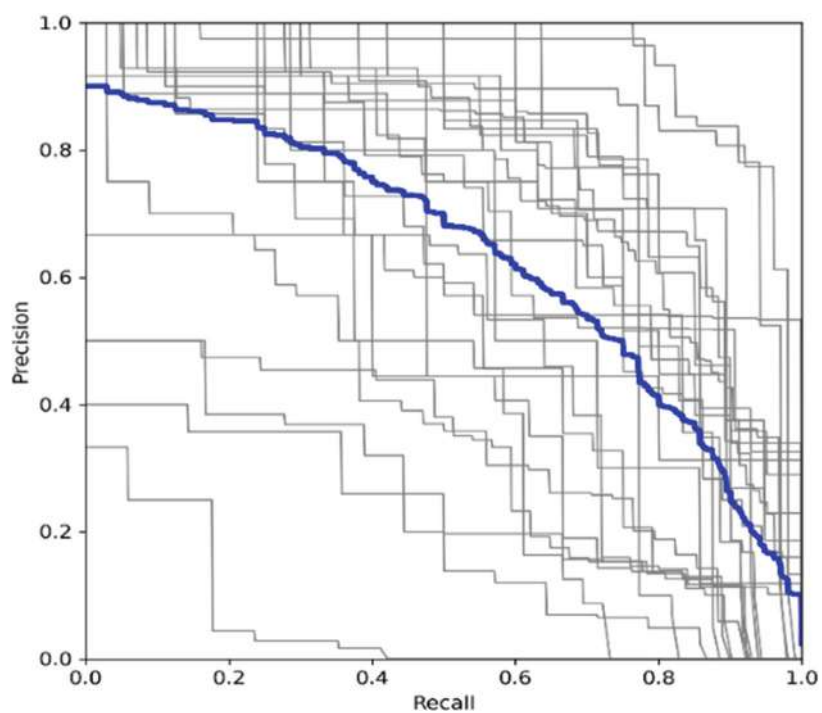


Fig. 11 Precision-recall for YOLOv7

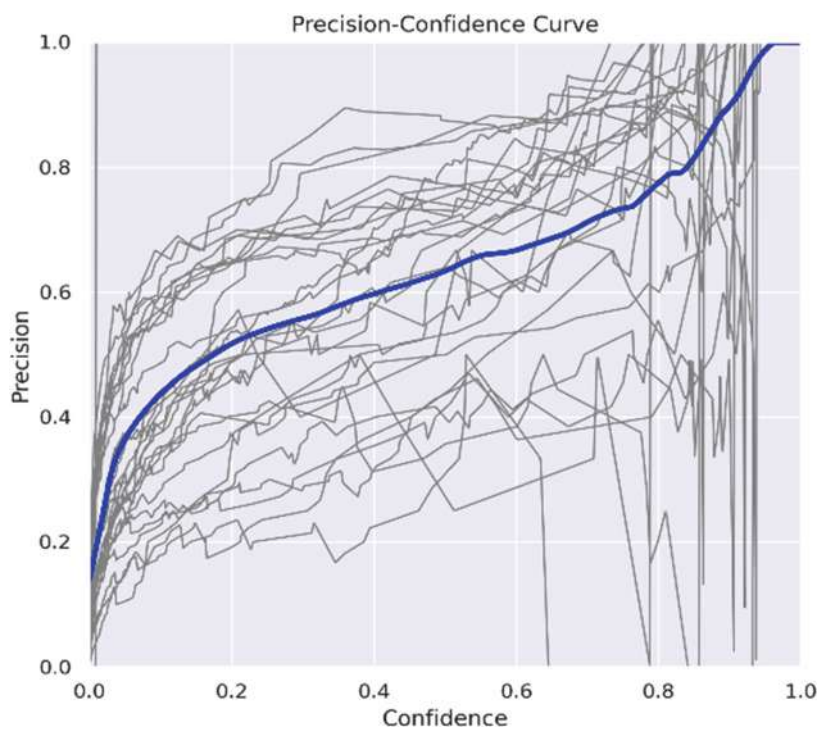


Fig. 12 Precision for YOLOv8

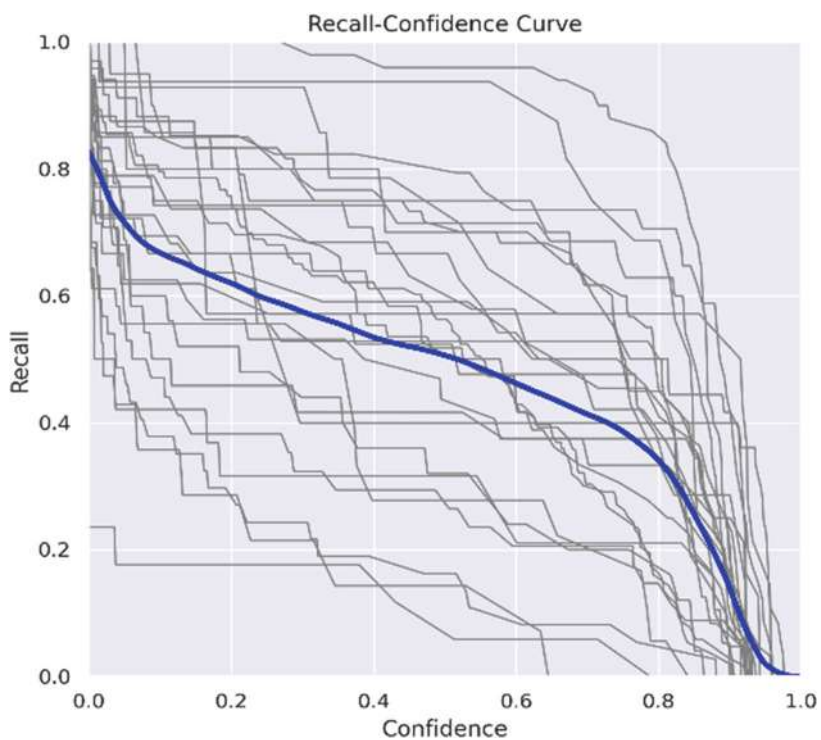


Fig. 13 Recall for YOLOv8

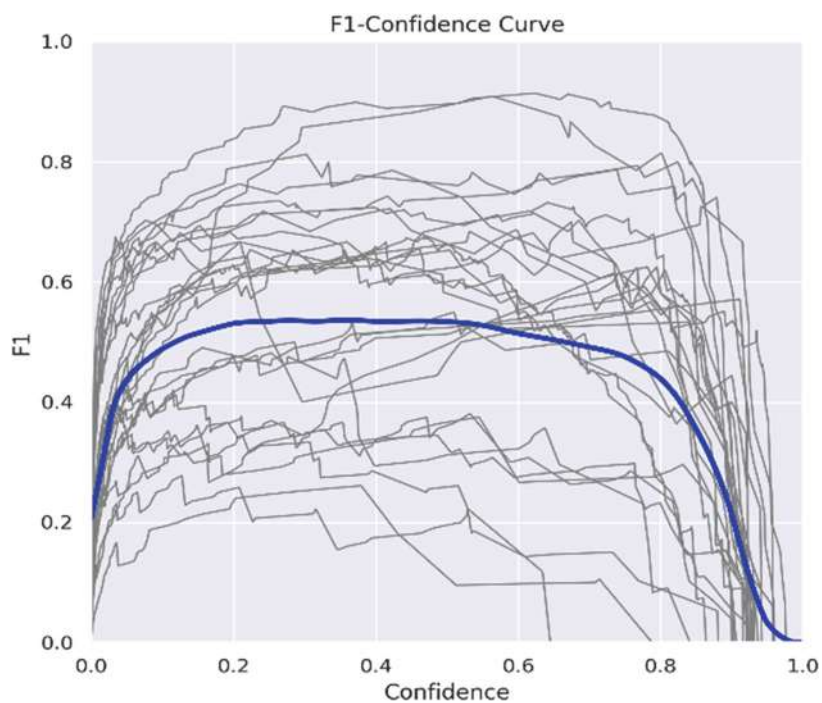


Fig. 14 F1-score for YOLOv8

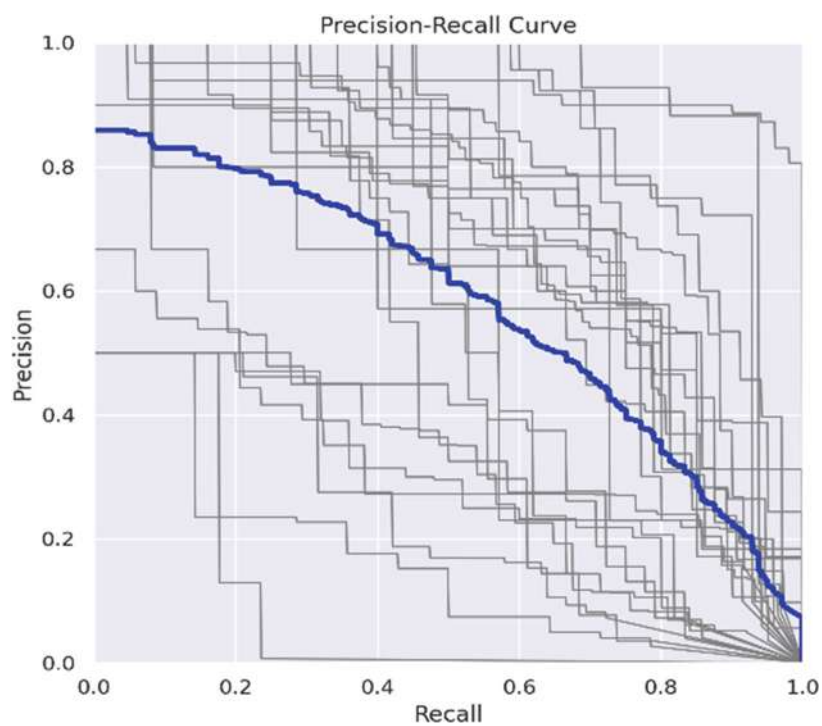


Fig. 15 Precision-Recall for YOLOv8

References

1. Singh D, Jain N, Jain P, Kayal P, Kumawat S, Batra N. PlantDoc: a dataset for visual plant disease detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD 2020 Jan 5 (pp. 249–253).
2. Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. *Frontiers in plant science*. 2016 Sep 22; 7:1419.
3. Kale SS, Panzade KP, Chavan NR. Modern farming methods: An initiative towards increasing the food productivity. *Food and Scientific Reports*. 2020 Apr;1(4):34.
4. Kasso M, Bekele A. Post-harvest loss and quality deterioration of horticultural crops in Dire Dawa Region, Ethiopia. *Journal of the Saudi Society of Agricultural Sciences*. 2018 Jan 1;17(1):88–96.
5. Almeida RP. Emerging plant disease epidemics: Biological research is key but not enough. *PLoS biology*. 2018 Aug 22;16(8): e2007020.
6. Fang Y, Ramasamy RP. Current and prospective methods for plant disease detection. *Biosensors*. 2015 Aug 6;5(3):537–61.
7. Sujatha R, Chatterjee JM, Jhanjhi NZ, Brohi SN. Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocessors and Microsystems*. 2021 Feb 1; 80:103615.
8. Morbekar A, Parihar A, Jadhav R. Crop disease detection using YOLO. In 2020 International Conference for Emerging Technology (INCET) 2020 Jun 5 (pp. 1–5). IEEE.
9. Lee SH, Goëau H, Bonnet P, Joly A. New perspectives on plant disease characterization based on deep learning. *Computers and Electronics in Agriculture*. 2020 Mar 1; 170:105220.
10. Shinde S, Kothari A, Gupta V. YOLO based human action recognition and localization. *Procedia computer science*. 2018 Jan 1; 133:831–8.
11. Yang F, Zhang X, Liu B. Video object tracking based on YOLOv7 and DeepSORT. *arXiv preprint arXiv:2207.12202*. 2022 Jul 25.
12. Wang CY, Bochkovskiy A, Liao HY. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*. 2022 Jul 6.
13. Patel D, Patel S, Patel M. Application of image-to-image translation in improving pedestrian detection. *arXiv preprint arXiv:2209.03625*. 2022 Sep 8.
14. Sanchez PR, Zhang H, Ho SS, De Padua E. Comparison of one-stage object detection models for weed detection in mulched onions. In 2021 IEEE International Conference on Imaging Systems and Techniques (IST) 2021 Aug 24 (pp. 1–6). IEEE.
15. T Mohapatra. 2018. ICAR News July–September 2018. Published in the monthly newsletter, <https://www.icar.org.in/sites/default/files/ICARNewsJulySeptember2018.pdf>
16. Oerke EC, Dehne HW, Schönbeck F, Weber A. Crop production and crop protection: estimated losses in major food and cash crops. Elsevier; 2012 Dec 2.
17. Strange RN, Scott PR. Plant disease: a threat to global food security. *Annual review of phytopathology*. 2005 Jul 28;43(1):83–116.
18. Jhatial MJ, Shaikh RA, Shaikh NA, Rajper S, Arain RH, Chandio GH, Bhangwar AQ, Shaikh H, Shaikh KH. Deep Learning-Based Rice Leaf Diseases Detection Using YOLOv5. *Sukkur IBA Journal of Computing and Mathematical Sciences*. 2022 Jul 21;6(1):49–61.
19. Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Computer software]. <https://github.com/ultralytics/ultralytics>. (n.d.).

Detection of Congenital Heart Disease from Heart Sounds Using 2D CNN-BiLSTM with Attention Mechanism



Ann Nita Netto, Lizy Abraham, and Saji Philip

Abstract The birth prevalence of congenital heart disease (CHD) is 9/1000, and it is always challenging for paediatricians to diagnose. We propose a novel lightweight approach for CHD detection using a combined 2D convolutional neural network (CNN), bidirectional long short-term memory (Bi-LSTM), and attention mechanism. The objective is to classify paediatric phonocardiogram (PCG) recordings into five distinct classes: normal (N), functional (F), pathological (P), ventricular septal defect (VSD), and atrial septal defect (ASD). The work also investigates two-category classification into N and P and three-category classification into N, F, and P. The model achieved 95.6%, 93.2%, and 83.7% for two, three, and five categories. The model complexity and number of trainable parameters are much lower than the existing technologies, and the performance is comparable to the previous works. This is the first-ever deep learning-based work to classify paediatric heart sounds into more than four classes. The work was done on a dataset collected by the researchers. The robustness of the model was evaluated by testing against the publically available five-category adult PCG dataset. The model's performance surpasses existing methods, highlighting its potential as a reliable tool for CHD and other cardiovascular disease detection.

Keywords Attention · ASD · Auscultation · BiLSTM · CNN · Congenital heart disease · Functional murmur · Phonocardiogram · VSD · Phonocardiogram · Softmax

A. N. Netto (✉) · L. Abraham

Department of Electronics and Communication Engineering, LBS Institute of Technology for Women, APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India
e-mail: lizyabraham@lbsitw.ac.in

S. Philip

Department of Cardiology, Thiruvalla Medical Mission Hospital, Thiruvalla, Kerala, India

1 Introduction

Congenital heart defects (CHD) are malformations that occur due to abnormal development of the heart. These diseases affect approximately 1% of newborns and account for 3% of all infant deaths. Therefore, CHD is one of the most frequent causes of infant mortality [1]. Early recognition and proper treatment of cardiac diseases, especially in newborns and infants, are crucial to prevent increased mortality rates worldwide. Therefore, prioritising the early detection of heart disease is of utmost importance.

A newborn's heart continues to develop within the first days or weeks of life; therefore, there exist extra sounds called innocent murmurs or functional murmurs. Also, the additional sound may be due to a heart problem that a baby is born with, called a pathological murmur. Pathological murmurs are due to a congenital heart defect. Heart murmurs are very common and occur in up to 80% of children at some time or another. It is essential to distinguish between innocent and pathological murmurs as the delayed diagnosis or referrals to the centres with expertise results in death or a truncated life. The auscultation of the heart can provide helpful information that can contribute to diagnosing CHD. However, the limitation is that paediatricians possess varying levels of auditory abilities, and it is always challenging to identify the different murmurs from the typical heart sounds. Artificial intelligence incorporated with a heart-sound differentiating system solves the problem.

The number of studies focusing on paediatric heart sound classification is limited due to the scarcity of datasets uncorrupted by noise. Also, most work focuses on two-class categories, distinguishing between healthy individuals and those with CHD. The recent works include both machine learning and deep learning algorithms. Various studies have been conducted to classify paediatric PCG into normal and abnormal. For instance, Amir et al. [2] used an Arash band and fixed split S2 along with a support vector machine (SVM) to achieve an accuracy of 87%. Meanwhile, Amir M et al. [3] utilised SVM with Gabor wavelet and zero crossing to gain an accuracy of 92%. Xiao et al. [4] developed the 1D-CNN model and gained 93% accuracy. The work proposed a lightweight architecture with outstanding performance. Jiaming et al. [5] explored the possibilities of PCG decomposition and recomposition with the Shannon energy envelope and Artificial Neural Network. However, the results are unreliable due to the limited dataset of 86 subjects. Sergi Gómez et al. [6] created a model using XBoost with 67 features and could only achieve 77% accuracy. Ding Chen et al. [7] developed another lightweight model using CNN-LSTM on raw data and obtained a comparable result of 85%. Weize et al. [8] presented a majority voting ensemble model backed by segmentation and feature extraction, achieving 95% accuracy with complex architecture. Sukryool Kang et al. [9] classified it into a functional and pathological murmur, whereas Raj Shekhar et al. [10] categorised into functional murmur and no murmur. Amir A. Sepheri et al. [11] introduced a model to specifically detect atrioventricular

diseases from other heart diseases. Sumair Aziz et al. [12] developed a model to classify paediatric PCG into N, ASD, and VSD using features obtained from Empirical Mode Decomposition, Mel- Frequency Cepstral Coefficients (MFCC), and 1-D Local Ternary Pattern applied to SVM. The limitation of this model is that it considered only 56 subjects for the study. To our knowledge, paediatric heart sound classification into normal and three distinct CHD classes (4 class classification) or 4 CHD classes (5 class classification) has not been reported.

The work proposes a five-class CHD detection model from phonocardiograms using a serial 2D convolutional neural network (CNN) and bidirectional long short-term memory with an attention mechanism. This model can classify the paediatric PCG into five classes: atrial septal defect (ASD), ventricular septal defects (VSD), functional murmur (F), normal (N), and pathological (P). The researchers collected the dataset from Thiruvalla Medical Mission Hospital under the supervision of an expert paediatric cardiologist. The study also evaluates the model's performance over two (normal/pathological) and three (normal, pathological, functional) classes.

The contribution of the study includes the following:

1. Development of a novel method for detecting and classifying CHD using a lightweight deep-learning model that combines 2D-CNN and BiLSTM with an attention mechanism.
2. Collected and organised a new paediatric PCG dataset for the study.
3. The model was tested for two, three, and five-class classifications and produced excellent results.
4. Validated the model's effectiveness and transferability by applying it to a publically available adult PCG dataset with five classes, achieving an average class accuracy of 99.4%.
5. The study confirmed that incorporating the attention mechanism reduced computational complexity and training parameters while improving accuracy.

2 Methodology

This study aims to develop a 2D CNN and BiLSTM model with an attention mechanism to analyse and classify paediatric PCG data. The 1D PCG signal must be represented in 2D space, which is achieved by converting the signal into MFCC. The MFCC is then applied to a 2D CNN for temporal and spatial feature extraction. The correlation among and within the PCG signals is further analysed using BiLSTM. Finally, the attention mechanism extracts relevant information from the extracted features. Figure 1 gives the visual representation of the model. The PCGs are collected using an electronic stethoscope by connecting them to a laptop or mobile phone via Bluetooth. Thus, obtained PCGs are given to 2D CNN-BiLSTM with attention mechanism for early disease detection.

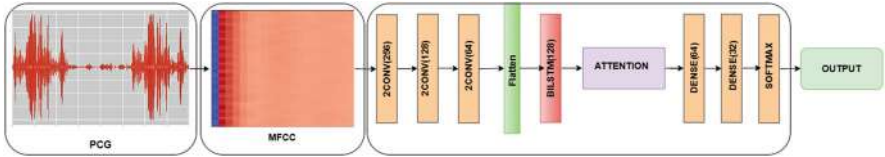


Fig. 1 2DCNN-BiLSTM with attention model

Table 1 Dataset

Category	No. of samples
ASD	313
Functional	333
Normal	851
Pathological	853
VSD	365

2.1 Dataset

The paediatric dataset was collected at Thiruvalla Medical Mission Hospital following the approval of the local institutional ethics committee. The auscultation was performed by a paediatric cardiologist within the clinical setting. The heart sounds were collected from 400 paediatric subjects who attended the cardiology OP from 20 August 2022 to 1 June 2023 using a Littman Eko Core electronic stethoscope. The study’s participants’ age range spans from newborns up to 13 years old. The duration of data varies between 3 and 30 s. Data was collected from five auscultating points for paediatric patients, while for neonates and infants, from four points. PCG signal was truncated to 3 s to include all the available data during the training process. The details of the dataset are given in Table 1.

The second dataset involved in the study is the publically available adult dataset used in the paper [13]. 1000 PCG signals are equally divided among five disease classes: Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), Mitral Valve Prolapse (MVP), and Normal (N).

2.2 2D Signal Representation

MFCC [14] is the methodology selected for feature representation as it is one of the predominant techniques for representing signals at the lower end of the frequency spectrum. The MFCC is derived using the Python Librosa package, and the selected tuning parameters are given in Table 2. After resampling to 8000 Hz, the raw data is directly applied to MFCC for feature representation.

MFCC features effectively reduce the impact of noise and high-energy components and highlight the essential frequency elements that help in classification. The steps involved in MFCC feature representation are Mel-frequency wrapping,

Table 2 MFCC tuning parameters

Parameters	Value	Remark
No. of coefficient	20	The shape of the spectrum considered
FFT window length	2048	Wide window provides good frequency resolution
Hop length	512	One-fourth of the window size

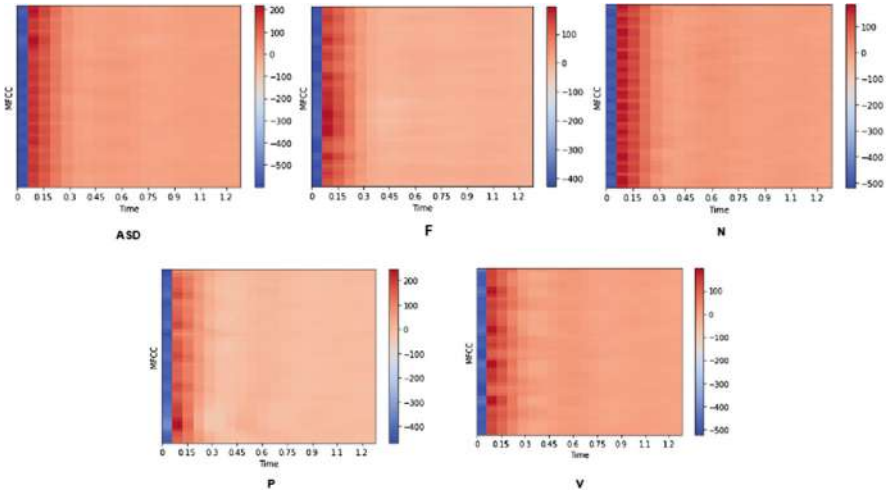


Fig. 2 MFCC representation of paediatric PCG signals

logarithmic compression, and cepstral representation. The MFCC representation for each class is shown in Fig. 2. The resultant MFCC is applied to the CNN to extract high-level features.

2.3 CNN-BiLSTM Attention Model

The high-level features that aid the classification can be extracted using the 2D CNN-BiLSTM layers. The CNN provides the local, spatial, and hierarchical features, whereas the sequential information, the correlation with other PCG signals and within the signal, is derived by BiLSTM. The CNN module comprises three Convolution (Conv) layers with 256,128 and 64 units. MaxPooling (2 × 2) follows each Conv layer. The output of the third MaxPooling is flattened and applied to BiLSTM. Only one Bilstm layer is involved in this model. The number of units associated with the BiLSTM layer is 128. Both CNN and BiLSTM activation function applied is ReLU. The output from the BiLSTM is the input to the Attention module.

The attention module [15] first generates a weight matrix W_t with glorot uniform as the initialiser. This step is followed by initialising self-bias vector S_b with zero

initialisers. These parameters are used to generate the attention weights that decide the relevant information. The weighted input is given by Eq. 1, and attention weight is described by Eq. 2.

$$W_i = \tan h (X * W_t + S_b) \quad (1)$$

$$A_w = \text{softmax} (W_i) \quad (2)$$

where X represents the 2D input vector sequence from the BiLSTM. The generated attention weights are applied to the input sequence to create a weighted sequence. Element-wise multiplication is performed, and it scales the input vector, helping the model to concentrate on the relevant information.

$$O = X * (A_w) \quad (3)$$

O represents the weighted output sequence, and finally, the summation is performed on the output to obtain the weighted sequence of the input sequence X .

$$\text{output} = \text{sum} (O, \text{axis} = 1) \quad (4)$$

The output from the attention module is applied to a dense layer of 64 units, followed by another dense layer of 32 units. The last dense layer predicts the output class with the help of the softmax activation function. The optimiser, loss function, learning rate, and batch size involved in this study are ‘Adam’, ‘sparse categorical cross entropy’, ‘0.0001’, and ‘32’ based on the investigation done in the paper [16]. The proposed model was developed using Keras and TensorFlow libraries in Python 3.7 and executed on a system equipped with an Intel Core i7 processor and Nvidia GeForce GTX 1060 graphics card with 6GB of RAM. Out of 2715 samples, 10% was dedicated to testing (272 samples), 10% of the remaining data was assigned for validation (245 samples), and 2198 samples were for training. The model achieved an accuracy of 86%, and to validate the robustness of the model, 10 cross-validations were performed.

2.4 Evaluation Metrics

The evaluation metrics showcase how well a model performs. The metrics involved in this study are accuracy, specificity, recall, precision, Area Under the Receiver Operating Characteristics (AUC-ROC), and Precision-Recall score. This study also plotted the AUC-ROC Curve to visualise the performance of the proposed model.

$$\text{Accuracy} = \frac{\text{True positive} + \text{True Negative}}{\text{Total Samples}} \quad (5)$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

(6)

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False Positive}}$$

(7)

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False Negative}}$$

(8)

3 Results and Discussions

The 2DCNN-BiLSTM with attention model is a lightweight model that predicts paediatric heart anomalies with high accuracy and precision. This work explores the ability of the proposed model to classify PCG into two, three, and five classes. For 2 class classification, the model achieved 98%, and 3 class exhibited 96% model accuracy. The accuracy of the 5 class classification is 86%, which is lower compared to the 2 and 3 class’ performance but higher than the cardiologist’s accuracy (<80%) [17]. After ten-fold cross-validation, the obtained accuracies are 83.7%, 93.2%, and 95.7% for five, three, and two class predictions, respectively.

For binary classification, the normal and functional murmur PCG data combined to form the normal category, as functional murmur is typical in newborns and does not represent any disease. ASD, VSD, and other diseases combined to get the pathology category. The normal category comprises 1184 samples, and the pathology category has 1498 samples. The proposed methodology can classify PCG into healthy or congenital heart conditions with high precision and accuracy. Table 3 presents the results of the model’s performance analysis.

Table 3 Binary class evaluation results

Fold number	Precision	Recall	F1-Score	Specificity	Model accuracy	ROC-AUC score	PR score
1	83	82.5	82.5	82.5	83.1	90.83	90.6
2	97.5	98	97.5	97.9	98	99.4	99.3
3	92	92	91.5	90.8	92	97.9	97.85
4	97	97	97	97.15	97	99.6	99.6
5	97.5	97	97.5	97.40	97.4	99.2	99
6	97.5	98	97.5	97.80	98	99.9	99.9
7	97.5	98	98	98	98	99.7	99.75
8	98.5	98	98	98.15	98	99.4	99
9	98	98	98	97.70	98	99.8	99.8
10	97	97.5	97.5	97.45	97	99.5	99.15
Average	95.5	95.6	95.5	95.5	95.7	98.52	98.4

Table 4 Three class evaluation results

Fold number	Precision	Recall	F1-Score	Specificity	Model accuracy	ROC-AUC score	PR score
1	78.33	73	74	83	75	90.5	85.8
2	86.67	92.33	89	94.2	89	97.4	96.2
3	96.33	96.6	96.3	97.83	96	99.3	99.3
4	92	94.3	92.3	95.7	92	98.96	98.4
5	97	98	97	98.30	97	99.2	98.4
6	96.3	97.6	96.6	97.63	97	99.6	99.3
7	95.3	96.3	96	98	96	99.6	99
8	97.3	98.3	97	98	97	99.1	97
9	96.3	97.3	97.6	98.23	97	99.4	99.1
10	95	97.3	96	97.90	96	99.6	99.4
Average	93	94	93.2	95.9	93.2	98.2	97.2

The labels in the three-label classification are functional, normal, and pathological. It is always challenging for paediatricians to distinguish functional murmurs from pathological murmurs. This model exhibited a significant ability to differentiate between functional and abnormal murmur. For the proposed model, functional murmur prediction accuracy is 98.7%. Three class classification results are shown in Table 4.

The five category classification results of CNN-BILSTM with Attention are impressive, achieving an accuracy of 83.7% and a precision of 85.6% on a raw signal without any preprocessing, denoising, or segmentation. In folds 7 and 8, the model even surpassed 90% accuracy. The model’s ability to predict functional and normal categories with 97.3% and 93.8% sensitivity is exceptional, given the limited dataset. Since some diseases in the pathology category (aortic stenosis, pulmonary stenosis, and mitral regurgitation), VSD, and ASD belong to the systolic group of murmurs; the model finds it challenging to differentiate among themselves. This limitation can be mitigated by increasing the number of samples. Table 5 demonstrates the five-category results.

Figures 3 and 4 showcase the confusion matrix and AUC-ROC curve for the top-performing folds in three categories (2, 3, and 5 classes). The superior AUC-ROC score (>0.9) demonstrates the model’s ability to distinguish all distinct classes effectively.

3.1 Comparison with Existing Works

Many existing methods for paediatric PCG classification rely on denoising techniques, segmentation, and manual feature extraction [5, 8, 10, 12, 17]. However, the highest reported accuracy for 2-class classification using hand-crafted data is 95% [8], and raw data is 85% [7]. For 3-class classification, the top accuracy is 97%

Table 5 Five category results

Fold number	Precision	Recall	F1-Score	Specificity	Model accuracy	ROC-AUC score	PR score
1	68.4	58.8	60.8	89.8	63	86.3	69.16
2	75.4	72.4	71.4	90.14	74	92.28	81.72
3	92.4	83.6	84.6	96.74	89	97.14	91.6
4	92.6	83.6	82.4	96.66	88	96.32	90.24
5	90	84.8	84	96.72	88	95.56	89.08
6	83.8	81	80	96	86	94.26	85.96
7	91.2	83.8	83.8	98.32	90	95.68	89.78
8	92.4	87.8	88.6	97.66	92	97.32	93.46
9	84	79.6	76.6	95.56	84	94.42	83.62
10	85.6	77.8	78	94.86	83	94.66	86.1
Average	85.6	79.32	79.1	95.23	83.7	94.4	86.07

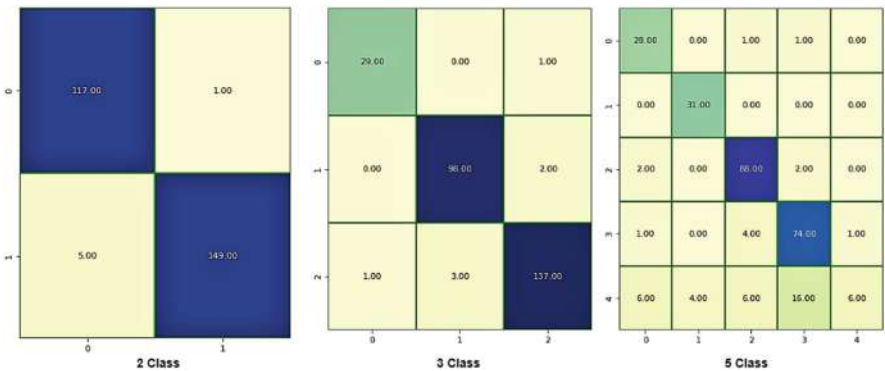


Fig. 3 Confusion matrix for highest scoring folds in two, three, and five class categories

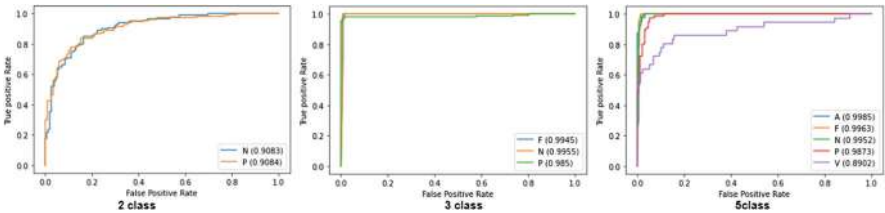


Fig. 4 AUC-ROC Curves of the highest scoring folds in two, three, and five class categories

[9], but with a very complex model architecture. These complex models cannot fit into low-processing devices like mobile phones or tablets. To our knowledge, no work has been reported for 5-class paediatric PCG classification using deep learning algorithms, and 2DCNN-BiLSTM with attention model achieved 83.7% accuracy on raw data. While a direct comparison with other works is not applicable due to different datasets, our proposed model acquired comparable results using

Table 6 Comparative analysis with existing pediatric models

Author	Acc	Sens	Spec	Precision	F1-score	AUC-ROC	Classes	Validation
Sumair Aziz et al. [12]	95.24	95.24	95.24	–	–	–	3	10 fold
Amir A. Sepheri [11]	91.6	88.4	–	–	–	–	2	5 fold
Jiaming Wang et al. [5]	93	93.5	91.7	–	–	–	2	–
Sergi Gómez et al. [17]	–	–	–	–	–	78	2	5 fold
Ding Chen et al. [7]	85	87	82	–	–	99.8	2	–
Weize et al. [8]	95	94.6	99.2	–	95.3	–	2	5 fold
Raj Shekhar et al. [10]		90	98	–	–	94	2	5 fold
This study	95.7	95.6	95.5	95.5	95.5	98.52	2	10 fold
	93.2	94	95.9	93	93.2	98.2	3	10 fold
	83.7	79.32	95.23	85.6	79.1	94.4	5	10 fold

a less complex architecture and training parameters. The pre-training overhead of data preparation is minimal in this proposed work. Comparison with the existing pediatric PCG classification is demonstrated in Table 6. It is evident from the table that the proposed model excels in all evaluation metrics.

We tested the effectiveness of our model for 5-class classification using the publicly available adult dataset. This dataset is the only five-class dataset publically available. The results were impressive, achieving excellent accuracy and specificity of 99.4% and 99.6%. A previous research paper [18] reported an accuracy of 99.6% with a total of 0.67 M trainable parameters, while our proposed model achieved an average class accuracy of 99.4% with only 0.58 M parameters. The model took only 45 s to train and test. A comparison of the proposed work with existing methodologies has been done and shown in Table 7. The comparison criteria are the data set and ten-fold cross-validation. The same pediatric model was implemented, except for the number of MFCC coefficients engaged (13 instead of 20).

The results and analysis show that the model performed well for all three types (two, three, and five) of classifications and can be used for adult and paediatric PCG classification.

4 Conclusion and Future Scope

Congenital heart defects (CHDs) are common congenital disabilities that can be life-threatening if not diagnosed and treated early. The presence of functional murmur in neonates makes the diagnosis challenging. The proposed model uses a 2D CNN-BiLSTM with an attention mechanism to overcome this challenge and to identify CHDs. The CNN-BiLSTM gathers all possible features (spatial, temporal, sequential) from the data, while the attention mechanism selects the most relevant features to train the model. The CNN-BiLSTM Attention combination has resulted in a highly efficient, lightweight model with high accuracy, parameter efficiency,

Table 7 Comparative analysis with existing models (GitHub Dataset)

Author	Methodology	Acc	Sens	Spec	Precision	F1- score	AUC-ROC	AUPRC	Validation
Yaseen et al. [13]	MFCC + DWT + KNN	97.4	97.6	98.8	–	–	–	–	5 fold
	MFCC + DWT + DNN	92.1	94.5	98.2	–	–	–	–	5 fold
	MFCC + DWT + SVM	97.9	98.2	99.4	–	–	–	–	5 fold
Jahmunah et al. [19]	WaveNet	97	92.5	98.1	–	–	–	–	10 fold
Neeraj et al. [20]	Mel + CNN	98.6	–	–	–	–	–	–	10 fold
Samiul et al. [18]	CardioXnet	99.6	98.6	–	98.6	98.5	–	–	10 fold
Alkhodari et al. [21]	MODWT+CNN-BiLSTM	99.32	98.30	99.58	98.50	98.32	99.8	–	10 fold
J S Khan et.al [22]	Cardi-Net	98.87	97.3	99.2	96.7	96.7	–	–	10 fold
Al-Issa et al. [23]	FFT+CNN-LSTM	98.5	–	–	–	98.5	99.78	–	10 fold
Pichori et al. 2022 [24]	FBSE-EWT +PSR	98.5	96.3	99.05	–	96.29	99.74	–	10 fold
Qaisar et al. 2022 [25]	CVT-Trans	99	96.7	96.3	96.3	96.3	–	–	10 fold
This study	CNN+BiLSTM+ATTENTION	99.4	98.5	99.6	98.5	98.4	99.9	99.9	10 fold

and computational efficiency. This model can identify normal, functional, ASD, VSD, and other pathologies accurately and precisely. The innocent murmur, often misdiagnosed, was detected with 98% accuracy and 97% sensitivity for five category classifications. This model also applies to adult PCG classification, as it performed excellently on an adult dataset. Since it is lightweight, the model can be used to develop mobile or web applications to assist the paediatrician/cardiologist during diagnosis.

In the future, we aim to create a substantial dataset for PCG that includes numerous heart diseases. Data augmentation techniques can be explored to increase the diversity and quantity of data. Another potential scope is using transfer learning to leverage pre-trained models trained on large-scale audio datasets. Additionally, we plan to design a mobile application that medical professionals can use to diagnose patients with a user-friendly interface.

Acknowledgement The first author would like to acknowledge the CERD of A.P.J. Abdul Kalam Technological University, Kerala, India, for providing PhD fellowship.

Funding This research is supported by the Centre for Engineering Research and Development (CERD), A.P.J Abdul Kalam Technological University, Kerala, India.

Compliance with Ethical Standards

Conflict of Interest The authors declare no conflict of interest.

Ethical Approval This study was performed in line with the principles of the Declaration of Helsinki. Ethical committee approval has been obtained from the local ethical committee ECR/975/Inst/KL/2017 with study reference number 05/08/2022.

Consent to Participate Informed parental consent was obtained from all participants included in the study.

Consent to Publish Any personal information or photographs are not used in this study, hence does not require consent to publish.

Date or Code Availability The code will be made public upon publication.

Acknowledgement The first author would like to acknowledge CERD, Kerala, India, for supporting the work.

References

1. D. Van Der Linde *et al.*, "Birth prevalence of congenital heart disease worldwide: A systematic review and meta-analysis," *J. Am. Coll. Cardiol.*, vol. 58, no. 21, pp. 2241–2247, Nov. 2011, <https://doi.org/10.1016/J.JACC.2011.08.025>.

2. A. A. Sepehri, A. Kocharian, A. Janani, and A. Gharehbaghi, "An Intelligent Phonocardiography for Automated Screening of Pediatric Heart Diseases," *J. Med. Syst.*, vol. 40, no. 1, pp. 1–10, 2016, <https://doi.org/10.1007/s10916-015-0359-3>.
3. A. M. Amiri, M. Abtahi, N. Constant, and K. Mankodiya, "Mobile phonocardiogram diagnosis in newborns using support vector machine," *Healthc.*, vol. 5, no. 1, pp. 1–10, 2017, <https://doi.org/10.3390/healthcare5010016>.
4. B. Xiao *et al.*, "Follow the Sound of Children's Heart: A Deep-Learning-Based Computer-Aided Pediatric CHDs Diagnosis System," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1994–2004, 2020, <https://doi.org/10.1109/IJOT.2019.2961132>.
5. J. Wang *et al.*, "Intelligent Diagnosis of Heart Murmurs in Children with Congenital Heart Disease," *J. Healthc. Eng.*, vol. 2020, 2020, <https://doi.org/10.1155/2020/9640821>.
6. S. Gómez-Quintana *et al.*, "A framework for ai-assisted detection of patent ductus arteriosus from neonatal phonocardiogram," *Healthc.*, vol. 9, no. 2, pp. 1–20, 2021, <https://doi.org/10.3390/healthcare9020169>.
7. D. Chen *et al.*, "Automatic Classification of Normal–Abnormal Heart Sounds Using Convolution Neural Network and Long-Short Term Memory," *Electronics*, vol. 11, no. 8, p. 1246, 2022, <https://doi.org/10.3390/electronics11081246>.
8. W. Xu *et al.*, "Automatic pediatric congenital heart disease classification based on heart sound signal," *Artif. Intell. Med.*, vol. 126, p. 102257, Apr. 2022, <https://doi.org/10.1016/J.ARTMED.2022.102257>.
9. S. Kang, R. Doroshov, J. McConaughy, and R. Shekhar, "Automated identification of innocent Still's murmur in children," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 6, pp. 1326–1334, 2017, <https://doi.org/10.1109/TBME.2016.2603787>.
10. R. Shekhar, G. Vanama, T. John, J. Issac, Y. Arjoune, and R. W. Doroshov, "Automated identification of innocent Still's murmur using a convolutional neural network," *Front. Pediatr.*, vol. 10, 2022, <https://doi.org/10.3389/fped.2022.923956>.
11. A. Gharehbaghi, A. A. Sepehri, and A. Babic, "Distinguishing septal heart defects from the valvular regurgitation using intelligent phonocardiography," *Stud. Health Technol. Inform.*, vol. 270, pp. 178–182, 2020, <https://doi.org/10.3233/SHTI200146>.
12. S. Aziz, M. U. Khan, M. Alhaisoni, T. Akram, and M. Altaf, "Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features," *Sensors (Switzerland)*, vol. 20, no. 13, pp. 1–20, 2020, <https://doi.org/10.3390/s20133790>.
13. Yaseen, G., Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features," *Appl. Sci.*, vol. 8, no. 12, 2018, <https://doi.org/10.3390/app8122344>.
14. E. Delgado-Trejos, A. F. Quiceno-Manrique, J. I. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, "Digital Auscultation Analysis for Heart Murmur Detection," 2008, <https://doi.org/10.1007/s10439-008-9611-z>.
15. N. Peng *et al.*, "Environment sound classification based on visual multi-feature fusion and GRU-AWS," *IEEE Access*, vol. 8, pp. 191100–191114, 2020, <https://doi.org/10.1109/ACCESS.2020.3032226>.
16. A. N. Netto and L. Abraham, "Detection and Classification of Cardiovascular Disease from Phonocardiogram using Deep Learning Models," *Proc. 2nd Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2021*, pp. 1646–1651, Aug. 2021, <https://doi.org/10.1109/ICESC51422.2021.9532766>.
17. S. Gómez-Quintana *et al.*, "A framework for ai-assisted detection of patent ductus arteriosus from neonatal phonocardiogram," *Healthc.*, vol. 9, no. 2, pp. 1–19, 2021, <https://doi.org/10.3390/healthcare9020169>.
18. S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, and A. Gumaei, "CardioXNet: A Novel Lightweight Deep Learning Framework for Cardiovascular Disease Classification Using Heart Sound Recordings," *IEEE Access*, vol. 9, pp. 36955–36967, 2021, <https://doi.org/10.1109/ACCESS.2021.3063129>.
19. S. L. Oh *et al.*, "Classification of heart sound signals using a novel deep WaveNet model," *Comput. Methods Programs Biomed.*, vol. 196, p. 105604, 2020, <https://doi.org/10.1016/j.cmpb.2020.105604>.

20. N. Baghel, M. K. Dutta, and R. Burget, "Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 197, 2020, <https://doi.org/10.1016/j.cmpb.2020.105750>.
21. M. Alkhodari and L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Comput. Methods Programs Biomed.*, vol. 200, p. 105940, Mar. 2021, <https://doi.org/10.1016/J.CMPB.2021.105940>.
22. J. S. Khan, M. Kaushik, A. Chaurasia, M. K. Dutta, and R. Burget, "Cardi-Net: A deep neural network for classification of cardiac disease using phonocardiogram signal," *Comput. Methods Programs Biomed.*, vol. 219, p. 106727, Jun. 2022, <https://doi.org/10.1016/J.CMPB.2022.106727>.
23. Y. Al-Issa and A. M. Alqudah, "A lightweight hybrid deep learning system for cardiac valvular disease classification," *Sci. Reports 2022 121*, vol. 12, no. 1, pp. 1–20, Aug. 2022, <https://doi.org/10.1038/s41598-022-18293-7>.
24. S. I. Khan, S. M. Qaisar, and R. B. Pachori, "Automated classification of valvular heart diseases using FBSE-EWT and PSR based geometrical features," *Biomed. Signal Process. Control*, vol. 73, p. 103445, Mar. 2022, <https://doi.org/10.1016/J.BSPC.2021.103445>.
25. Q. Abbas, A. Hussain, and A. R. Baig, "Automatic Detection and Classification of Cardiovascular Disorders Using Phonocardiogram and Convolutional Vision Transformers," *Diagnostics*, vol. 12, no. 12, 2022, <https://doi.org/10.3390/diagnostics12123109>.

Automated Reviewer Assignment Process Using Machine Learning Technique



Sovan Bhattacharya, Arkaprava Mazumder, Ayan Banerjee,
Chandan Bandyopadhyay, and Subrata Nandi

Abstract The development of technology has surged people's curiosity about new technology, thus improving research works. This led to an increase in the number of article submissions. Hence, an already difficult task gets more difficult with an increased workload. This chapter implements an automated reviewer assignment on top of machine learning techniques. Here we collect ICBIM 2016 (<http://icbim2016.nitdgp.ac.in/>) conference data along with the corresponding data required. Five machine learning techniques have been applied over the collected data set, among which logistic regression and support vector machine (SVM) perform better. Lastly, it is compared with the “automated conflict of interest-based greedy approach” over the same data set, and the model is improved by 8%.

These authors contributed equally to this work.

S. Bhattacharya (✉)

Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India

Department of CSE, National Institute of Technology, Durgapur, West Bengal, India
e-mail: sovan.bhattacharya@bcrec.ac.in

A. Mazumder · A. Banerjee

Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India

e-mail: Ayan.B@labvantage.com

C. Bandyopadhyay

Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India

Department of CSE, University of Bremen, Bremen, Germany

e-mail: chandan.bandyopadhyay@bcrec.ac.in

S. Nandi

Department of CSE, National Institute of Technology, Durgapur, West Bengal, India

e-mail: snandi.cse@nitdgp.ac.in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_7

Keywords Reviewer assignment problem · Machine learning · Digital library · Citation analysis · Author name disambiguation · Text analytics · Bibliometric analysis

1 Introduction

A prevalent issue is the automated reviewer assignment problem (RAP), also known as the reviewer recommendation problem. Here, there are many inputs and assignment restrictions, and a group of submitted research papers or articles are assigned to one or more relevant reviewers in order to properly conduct the review process. RAP must be automated in order to become an automated reviewer assignment system, which will decrease time requirements, eliminate manual involvement, and improve reviewer assignment accuracy. It is conceivable for a reviewer to only have experience in one field of study or to have experience in others. The reviewers are assigned to papers based on their areas of expertise. The efficiency and calibre of the review process can be increased by using machine learning (ML) to automatically match reviewers with papers based on their experience and prior performance.

Motivation and Objective

A lot of researchers have performed various methods to resolve the reviewer assignment problems, but still problems like biasness and manual assignment have not been attended to be solved. Hence, here we present a machine learning-based solution to the RAP problem taking more features that will assign the papers automatically and also select the reviewers according to their quality and performance in the research field. If the appropriate paper is assigned to the appropriate reviewer by removing the biasness, then the quality papers will only get accepted which will lead to an increase in the quality of research in the particular field.

Contribution and Challenges

The framework of our reviewer assignment system works by collecting the set of reviewers and papers to pre-process the data. Then clusters are formed to extract out required features forming the final data set. Now the ML models are trained and implemented on a new data set for testing purposes, and an out-sampling test is carried out. Finally, we get the mapping of the submitted papers with appropriate reviewers. Also, we perform a baseline comparison with greedy algorithm. SVM and logistic regression perform the best with 95% accuracy, and the baseline comparison shows that our model is better by 8%.

2 Research Review

Many researchers have tried several approaches to normalize the reviewer assignment problems. In this section, we will be mentioning about our study of the previous state of the arts.

In Zhao et al. [1], they proposed a paper-reviewer recommendation problem as a classification issue. Similarly, they calculated the keyword similarity to solve the problem. Adebisi et al. [2] proposed a comparative study among Term Frequency-Inverse Document Frequency (TF-IDF), Latent Semantic Indexing (LSI), and Latent Dirichlet Allocation (LDA) and took out a suitability score between a submitted paper and a reviewer's representation papers. Anjum et al. [3] proposed a paper-reviewer matching method with the help of common topic space to achieve improvement in reviewer allocation. Peng et al. [4] proposed a time-perception and topic-based reviewer assignment technique. Bhattacharya et al. [5] proposed a machine learning-based approach to predict the forthcoming future quality of a budding researcher in the field based on various metrics, collaboration, consistency, etc. Tomkins et al. [6] proposed a method of reviewer bias in single- and double-blind peer review. Boyack et al. [7] proposed characterizing in-text citations in scientific papers. They analysed in-citations in different ways, recorded many feature sets, and cited the papers at multiple field levels. Shah et al. [8] determined several aspects of the data collected during the review process, including an experiment investigating the efficacy of collecting ordinal rankings from reviewers. Zhang et al. [9] proposed a multi-label categorization technique using hierarchical and transparent representation and named it Hiepar-MLC. Pradhan et al. [10] proposed a multilayered method using a greedy-based approach where three conditions are applied, topical bases selection, minimizing the COI, and optimization-based load distribution. We have compared this technique with our model by considering it as a baseline comparative study. Yan et al. [11] analysed how a paper is unsuited to assign to a reviewer with user-defined parameters. Payan et al. [12] have proposed reviewer assignment as a condition for unbiased allocation problems. They introduced an extension of pre-existing round-Robin mechanism and named it reviewer round-Robin mechanism. Bhattacharya et al. [13] proposed a heuristic model to measure the impact of indexing on a researcher's career based on metrics and collaboration partners. Fiez et al. [14] determined the key challenge that is suitably formalizing the peer review bidding process, for which to the best of our knowledge there are no prior formulations. Yong et al. [15] proposed a reviewer recommendation system, which is designed based on a knowledge graph and rules matching.

3 Methodology

Here we explain our developed methodology comprising the six stages. The first two initial stages involve collecting data and pre-processing for removal of author name disambiguation. Thereafter, the next three stages are used for ML model cross-validation for the automated reviewer selection process in TPC, and the last stage is the deployment of our model. Figure 1 shows the framework we have built to solve the RAP.

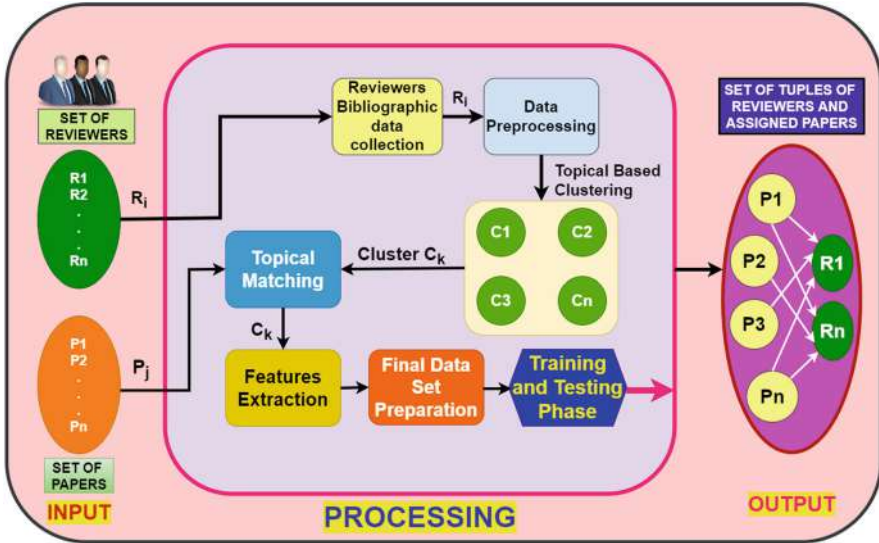


Fig. 1 Framework used to automate the RAP problem

- **Data Collection**

In this chapter, two different data sets are used. The first data set is taken from ICBIM 2016 conference data set. A total of 63 articles were received for reviews from different continents. Some of the papers were out of the conference scope. We also found that a total of 34 reviewers were assigned to this conference. The second data set consists of the reviewers' bibliographic data sets collected from Google Scholar, like publication title, abstract, publication year, co-author name with details, and citation.

- **Data Pre-processing**

In the data pre-processing step, a variety of errors are detected, and subsequently they are eliminated. The most common types of errors like handling the missing data values by removing or replacing them with a threshold and using a referential naming convention to assign names to the authors or reviewers in a particular format and author name ambiguity are covered.

- **Different Supervised ML Technique**

In this section, we have taken into account five supervised machine learning algorithms to get a better result in terms of accuracy, precision, recall, and F1-score with minimization of error. They are *naïve Bayes*, *decision tree*, *support vector machine (SVM)*, *logistic regression*, and *decision tree*. The reasons for adopting these techniques are explained below.

The computationally effective **naïve Bayes** algorithm performs well when there are a lot of irrelevant features. Both numerical and categorical data may be handled by **decision trees**, and their hierarchical structure makes it simple to analyse and comprehend how decisions are made. **SVM** is capable of managing big feature spaces and complex decision boundaries. The implementation of

logistic regression is comparatively straightforward, interpretable, and effective for both numerical and categorical variables. It is appropriate for problems involving reviewer assignment since it may produce rules that assign reviewers to papers in accordance with particular criteria. The robustness of **random forest** and its aptitude for handling noisy and multidimensional data are well recognized. By taking into account the outputs of various decision trees, it can offer a holistic view of the reviewer assignment problem.

4 Experimental Setup

We have inserted ML techniques in the tradition reviewer assignment system to automate it. Eight features are considered for the preparation of the data set from the reviewer's bibliographic records.

- **Feature used for automated reviewer assignment machine learning system**

We will now explain our choice of features for the reviewer assignment problem. From the literature survey we conducted, we drew an idea about the types of features that can be used for solving the RAP. So finally, in this chapter, eight specific features are selected for the research, as shown in Table 1.

- **Feature Selection/Feature Analysis**

As seen from the heat map in Fig. 2, seven among the eight features have positive values. So, the “Probability of frequency-based ranking amongst the Group of Fields” feature is removed from the data set.

Table 1 List of the selected features used in this chapter

Feature	Description	Reason behind selection of feature
$D(R_i, C_{P_j})$	Collaborative Distance between Reviewers and their Co-author Relationship	COI optimization
$D_{I5}(R_i, C_{P_j})$	Collaborative Distance between Reviewers and their Co-author relationship in last 5 years	COI optimization
$FRGF(C_k)_{R_i}$	Probability of frequency-based ranking amongst the Group of Fields	Reviewer domain identification
$P_{H-index}(C_k)_{R_i}$	Probability of H-index-based ranking amongst the Group of Fields	Reviewer performance measure
$P_{Age}(C_k)_{R_i}$	Probability of Academic Age-based ranking amongst the Group of Fields	Reviewer expertise measure
$P_{Citation}(C_k)_{R_i}$	Probability of Citation Count-based ranking amongst the Group of Fields	Reviewer expertise and performance measure multi-domain issue
$P_{i10-index}(C_k)_{R_i}$	Probability of i-10 index-based ranking amongst the Group of Fields	Reviewer performance measure
$P_{rgroup}(C_k)_{R_i}$	Probability of Recent Group of Fields ranking	To know the current field of interest of reviewer

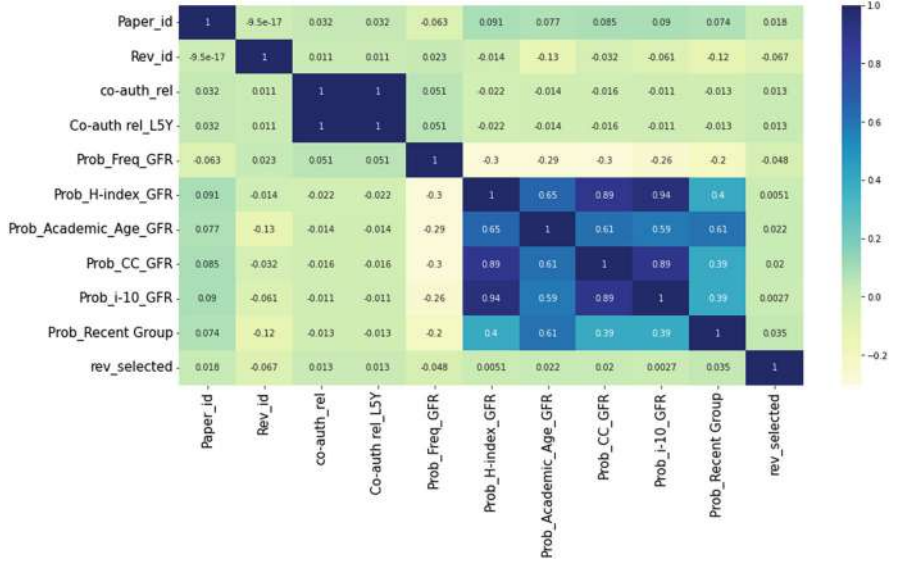


Fig. 2 Feature set heat map for RAP

• Evaluation Metrics

In the quality measurement, metrics like accuracy, precision, recall, and f1-score are used to assess the quality of models and compare various algorithms. Derivation of all the evaluation metrics are explained below:

$$\text{Accuracy}_{\text{avg}} = \frac{1}{N} \sum_{t=1}^N \frac{T_p^t + T_n^t}{T_p^t + T_n^t + F_p^t + F_n^t} \quad (1)$$

$$\text{Precision}_{\text{avg}}[P_{\text{avg}}] = \frac{1}{N} \sum_{t=1}^N \frac{T_p^t}{T_p^t + F_p^t} \quad (2)$$

$$\text{Recall}_{\text{avg}}[R_{\text{avg}}] = \frac{1}{N} \sum_{t=1}^N \frac{T_p^t}{T_p^t + F_n^t} \quad (3)$$

$$\text{F1-score}_{\text{avg}} = \frac{1}{N} \sum_{t=1}^N \frac{2 * P_{\text{avg}} * R_{\text{avg}}}{P_{\text{avg}} + R_{\text{avg}}} \quad (4)$$

5 Results and Discussion

Here, we discuss the results obtained from the two experiments and a baseline comparison we have performed.

• Variation of Training Ratio with Fixed Data Size

This experiment is performed by altering the training ratio as 50%, 60%, 70%, and 80% keeping the data size fixed at 59. From Fig. 3, which shows the plots of statistical score values of evaluation metrics against various training ratios, we conclude that SVM and logistic regression are the best options with an average accuracy of 95%. All the score values of this experiment are mentioned in Table 2.

• Variation of Data Size with Fixed Training Ratio

In this experiment, keeping the training ratio fixed at 80%, we vary the data size as 20, 30, 40, 50, and 59. From Fig. 4, which visualizes the graph plotted against the score values, we conclude that SVM and logistic regression perform the best with an average accuracy of 95%. For the score values of all the models at varying data sizes, refer to Table 3.

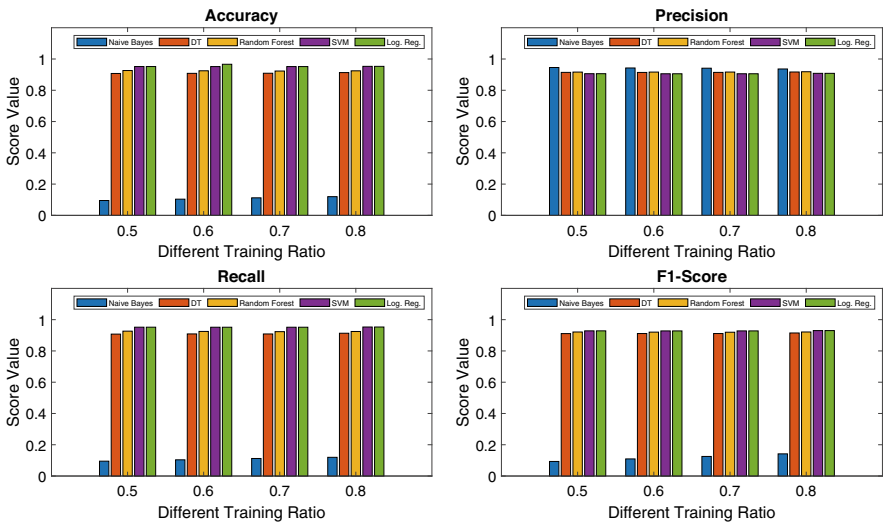


Fig. 3 All the model evaluation metrics score at different training ratios in fixed data size in RAP

Table 2 Score value of different train ratio with fixed data size

	Accuracy				Precision				Recall				F1-score			
	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8
Train ratio																
Naive Bayes	0.095	0.103	0.112	0.119	0.945	0.942	0.941	0.936	0.095	0.103	0.112	0.119	0.93	0.109	0.125	0.141
Decision tree	0.907	0.908	0.908	0.913	0.914	0.914	0.914	0.916	0.907	0.908	0.908	0.913	0.912	0.911	0.911	0.914
Random forest	0.926	0.924	0.923	0.924	0.916	0.916	0.916	0.918	0.926	0.924	0.923	0.924	0.921	0.920	0.919	0.921
SVM	0.951	0.951	0.951	0.953	0.906	0.905	0.905	0.908	0.951	0.951	0.951	0.953	0.928	0.928	0.928	0.930
Logistic regression	0.951	0.966	0.951	0.953	0.906	0.905	0.905	0.908	0.951	0.951	0.951	0.953	0.928	0.928	0.928	0.930

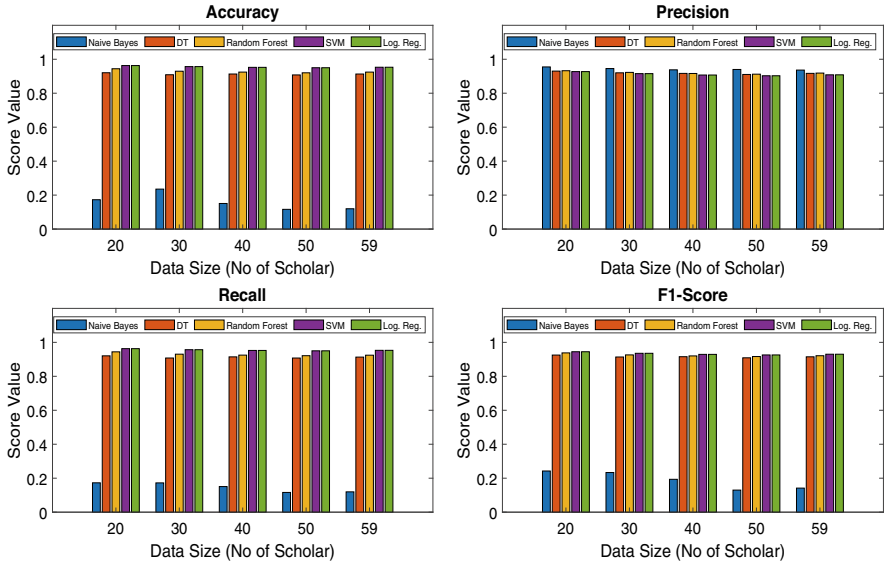


Fig. 4 All the models’ evaluation metrics score at different data sizes at 80% fixed training ratio in RAP

5.1 Baseline Comparison

We have compared the proposed method with the greedy-based method, based on two features, namely, topic extraction and similarity measure, and co-authorship distance measure, both implemented on our data set.

Figures 5 and 6 show comparison based on two features (co-authorship distance and topical similarity) and all eight features, respectively. In both cases, our model outperform by 8% and 5%, respectively. Table 4 shows the score values of all the models.

6 Conclusion and Future Scope

The automated reviewer assignment problem can be solved using machine learning (ML) techniques. Among the five ML techniques considered, logistic regression and SVM performed the best with an average accuracy of 95%. As for the baseline comparison, our model outperforms by an average of 5% against the state of the art. In future, automated load distribution can be introduced, and we can test our model on a bigger conference data set.

Table 3 Score value of different data size with 80% training ratio

	Accuracy					Precision					Recall					F1-score				
	59	50	40	30	20	59	50	40	30	20	59	50	40	30	20	59	50	40	30	20
Data size																				
Naive Bayes	0.119	0.116	0.150	0.235	0.172	0.936	0.939	0.937	0.945	0.954	0.119	0.116	0.150	0.172	0.172	0.141	0.129	0.193	0.233	0.242
Decision tree	0.913	0.907	0.913	0.908	0.920	0.916	0.910	0.916	0.920	0.930	0.913	0.907	0.914	0.907	0.920	0.914	0.908	0.915	0.913	0.925
Random forest	0.924	0.920	0.924	0.929	0.944	0.918	0.912	0.916	0.922	0.932	0.924	0.921	0.924	0.930	0.944	0.921	0.916	0.920	0.926	0.937
SVM	0.953	0.950	0.952	0.956	0.962	0.908	0.902	0.907	0.915	0.927	0.953	0.950	0.952	0.956	0.962	0.930	0.925	0.929	0.935	0.944
Logistic regression	0.953	0.950	0.952	0.956	0.962	0.908	0.902	0.907	0.915	0.927	0.953	0.950	0.952	0.956	0.962	0.930	0.925	0.929	0.935	0.944

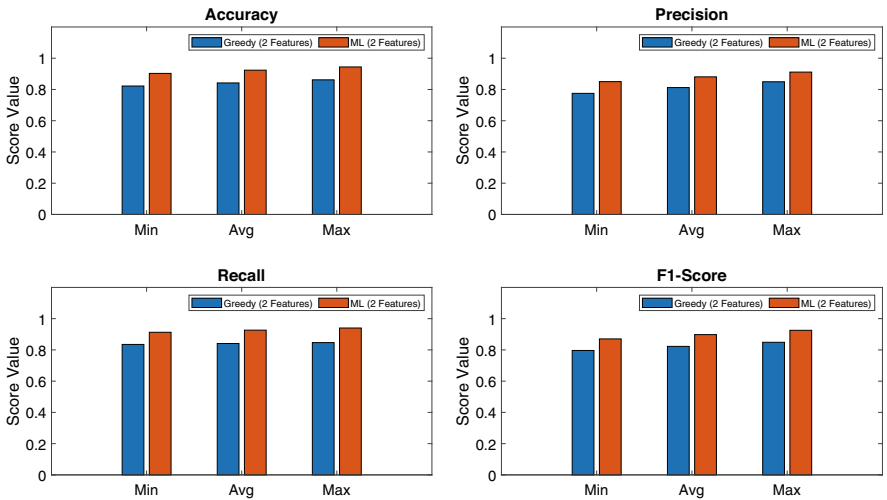


Fig. 5 Comparative study of greedy vs machine learning using minimum two features

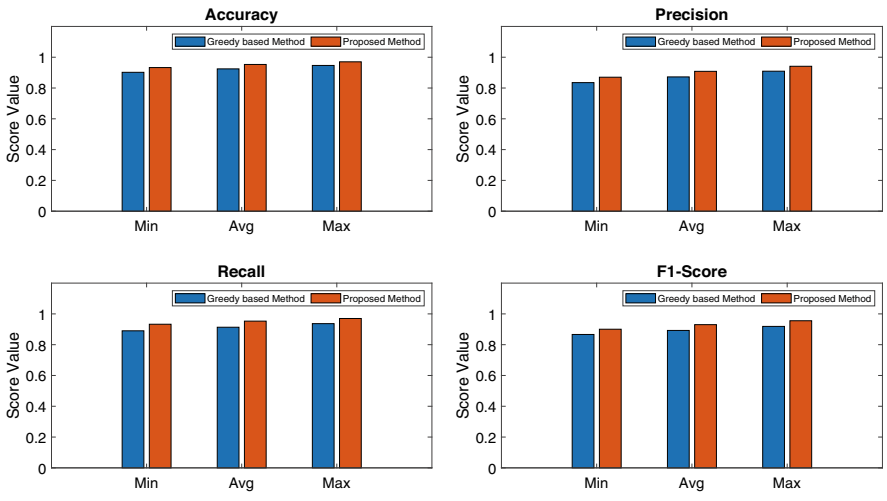


Fig. 6 Comparative study with the proposed method and greedy-based methods using the proposed eight features

Table 4 Score value of all the models

	Accuracy			Precision			Recall			F1-Score		
	MIN	AVG	MAX	MIN	AVG	MAX	MIN	AVG	MAX	MIN	AVG	MAX
Naive Bayes	0.077	0.119	0.191	0.865	0.936	0.971	0.077	0.119	0.191	0.066	0.141	0.244
Decision tree	0.885	0.913	0.937	0.883	0.916	0.946	0.880	0.913	0.937	0.888	0.914	0.940
Random forest	0.900	0.924	0.948	0.880	0.918	0.948	0.898	0.924	0.942	0.894	0.921	0.947
SVM	0.932	0.953	0.970	0.870	0.908	0.941	0.932	0.953	0.970	0.900	0.930	0.955
Logistic regression	0.932	0.953	0.970	0.870	0.908	0.941	0.932	0.953	0.970	0.900	0.930	0.955

Acknowledgments

- **Funding:** No funding.
- **Ethics Approval:** The author confirms that the work described has not been published before, and it is not under consideration for publication elsewhere.
- **Consent to Participate:** The authors agree to participate.
- **Consent for Publication:** The authors agree to publication in the Journal of Applied Intelligence.
- **Conflict of Interest:** We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.
- **Authors' Contributions:** All authors have equal contribution in the chapter.

References

1. Tan, S., Duan, Z., Zhao, S., Chen, J., Zhang, Y.: Improved reviewer assignment based on both word and semantic features. *Inf. Retr. J.* **24**(3), 175–204 (2021)
2. Adebisi, A., Ogunleye, O.M., Adebisi, M., Okesola, J.: A comparative analysis of tf-idf, lsi and lda in semantic information retrieval approach for paper-reviewer assignment. *JEAS* **14**(10), 3378–3382 (2019)
3. Anjum, O., Gong, H., Bhat, S., Hwu, W.-M., Xiong, J.: Pare: A paper-reviewer matching approach using a common topic space. *arXiv preprint arXiv:1909.11258* (2019)
4. Peng, H., Hu, H., Wang, K., Wang, X.: Time-aware and topic-based reviewer assignment. In: *DASFAA*, pp. 145–157 (2017). Springer
5. Bhattacharya, S., Banerjee, A., Goswami, A., Nandi, S., Pradhan, D.K.: Machine learning based approach for future prediction of authors in research academics. *SN Computer Science* **4**(3), 1–11 (2023)
6. Tomkins, A., Zhang, M., Heavlin, W.D.: Reviewer bias in single-versus double-blind peer review. *PNAS* **114**(48), 12708–12713 (2017)
7. Boyack, K.W., van Eck, N.J., Colavizza, G., Waltman, L.: Characterizing in-text citations in scientific articles: A large-scale analysis. *Journal of Informetrics* **12**(1), 59–73 (2018)
8. Stelmakh, I., Shah, N.B., Singh, A., Daumé III, H.: Prior and prejudice: The novice reviewers' bias against resubmissions in conference peer review. *Proc. ACM Hum. Comput.* **5**(CSCW1), 1–17 (2021)
9. Zhang, D., Zhao, S., Duan, Z., Chen, J., Zhang, Y., Tang, J.: A multi-label classification method using a hierarchical and transparent representation for paper-reviewer recommendation. *TOIS* **38**(1), 1–20 (2020)
10. Pradhan, T., Sahoo, S., Singh, U., Pal, S.: A proactive decision support system for reviewer recommendation in academia. *Expert Syst. Appl.* **169**, 114331 (2021)
11. Yan, S., Jin, J., Geng, Q., Zhao, Y., Huang, X.: Utilizing Academic-Network-Based Conflict of Interests: For Paper Reviewer Assignment. LAP LAMBERT Academic Publishing, Saarbruecken (2019)
12. Payan, J., Zick, Y.: I will have order! optimizing orders for fair reviewer assignment. *arXiv preprint arXiv:2108.02126* (2021)
13. Bhattacharya, S., Banerjee, A., Mazumder, A., Nandi, S.: Impact of author indexing from the co-authorship relation. In: *2022 MESIICON*, pp. 1–6 (2022). IEEE
14. Fiez, T., Shah, N., Ratliff, L.: A super* algorithm to optimize paper bidding in peer review. In: *UAI*, pp. 580–589 (2020). PMLR
15. Yong, Y., Yao, Z., Zhao, Y.: A framework for reviewer recommendation based on knowledge graph and rules matching. In: *2021 IEEE ICICSE*, pp. 199–203 (2021). IEEE

A Method for Detecting Retinal Microaneurysms in the Fundus Using CR-SF and RG-TF



S. Steffi 

Abstract The early detection of retinal microaneurysms significantly contributes to the prevention of visual impairment in individuals with diabetic retinopathy. During a fundus examination, this research describes a unique method for locating retinal microaneurysms. The strategy being discussed uses a circular frame of reference and radial gradient-based features. Additionally, the Boolean operators ‘AND’ and ‘OR’ may be used to interchange the set operations of intersection and union. Dilation is a kind of geometric transformation in which a figure’s size is changed, but the shape and proportions are kept the same. The first step in this approach involves preprocessing the fundus images and detecting potential microaneurysm candidates using morphological processing and adaptive thresholding algorithms. This work introduces two methods for extracting features, namely, circular reference-based shape highlights (CR-SF) and spiral gradient-based surface highlights (RG-TF), which aim to differentiate between microaneurysms and non-micro aneurysms. The robust backpropagation with PCA balanced data analysis machine learning technique is used to analyse the surface highlights that were retrieved from the candidates in terms of colour, shape, and appearance. During the testing phase, the extracted features from the test image are used to categorise the presence of microaneurysms and non-microaneurysms. The evaluation of the test was conducted using four unique datasets: MESSIDOR, DiaretDB1, e-optha-MA, and ROC datasets. The assessment included measuring many statistical parameters, including accuracy, sensitivity, specificity, AUC, and time complexity. The proposed technique demonstrates high levels of precision, sensitivity, specificity, and AUC, with values of 98.01%, 98.74%, 97.12%, and 0.9172, respectively. The evaluation result suggests that the proposed method for calculating the location of retinal microaneurysms surpasses traditional computation approaches.

S. Steffi (✉)

Department of Computer Science and Engineering, Nesamany Memorial Christian College affiliated to Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India

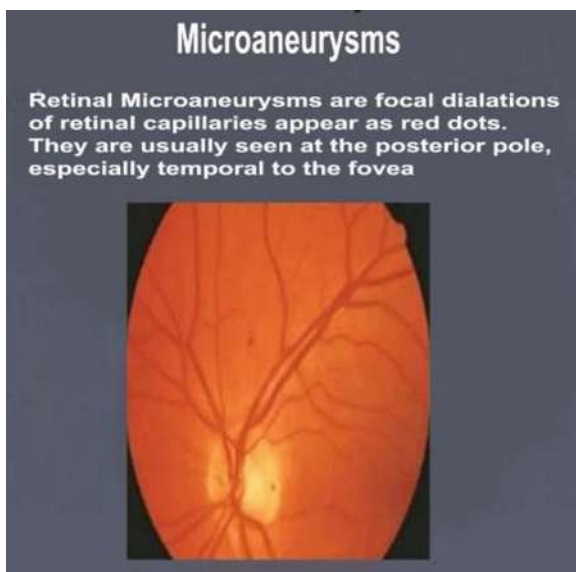
Keywords Neural network · Resilient back propagation (Rprop) · Morphological image processing · Diabetic · Adaptive thresholding · Microaneurysms

1 Introduction

A retinal microaneurysm refers to a small localised dilation of a blood vessel in the retinal layer of the eye. These protrusions have the potential to rupture, leading to the release of blood into the surrounding retinal tissue. Figure 1 depicts a comprehensive depiction of retinal microaneurysms, characterised by localised dilations of retinal capillaries appearing as red dots. These distinct features are often seen in the posterior pole of the eye, notably in the temporal area next to the fovea.

A widened retinal examination may reveal the presence of microaneurysms and aid in doing tests such as a fluorescein angiography. This imaging technique, which involves capturing images of the eye, is helpful in determining the onset and severity of the underlying condition responsible for these microaneurysms. A retinal microaneurysm may be caused by several forms of vascular infection or elevated blood pressure, however diabetes mellitus is the most prevalent underlying factor. Microaneurysms in isolation are unlikely to elicit any adverse consequences; nonetheless, it is important to exercise caution and be vigilant. An widened retinal examination may reveal the presence of microaneurysms and aid in doing tests such as a fluorescein angiography. This imaging technique, which involves capturing images of the eye, is helpful in determining the onset and severity of the underlying condition responsible for these microaneurysms.

Fig. 1 Overview of retinal microaneurysms be vigilant



A microaneurysm does not need independent therapy. The underlying systemic disorder responsible for these issues must be addressed. Treatment of underlying problems like diabetes, high blood pressure, or other contributing factors may usually successfully correct the majority of microaneurysms. Retinal microaneurysms are considerably decreased by adopting a healthy lifestyle and properly treating diabetes and high blood pressure. The primary impetus for this research is the identification of retinopathy at its nascent stage, enabling the implementation of preventative measures to safeguard the visual health of those afflicted with diabetes. Microaneurysms (MAs) are characterised as the first phase of retinopathy, manifesting as red spots seen in the superficial layer of the retinal walls. A novel model has been put out as an alternate approach to generate profiles of image intensity centred on the candidate MA in several directions. The primary obstacle in the detection of MA is in discerning the distinction between real positives and false positives, hence enabling the provision of reliable detection at the output. Over the course of recent years, a substantial multitude of algorithms have been created for the identification of potential microaneurysms. A diverse array of detection algorithms use both morphological and non-morphological techniques. However, a number of previously suggested approaches have shown suboptimal performance in terms of output quality. The use of the suggested technique has the potential to enhance efficacy.

The following portions of the paper are organised as follows: Sect. 2 provides an overview of retinal microaneurysms in individuals with diabetes. Section 3 presents an overview of the evaluations pertaining to fundus imaging. Section 4 illustrates the concept of morphological image processing. Sections 5 and 6 provide a comprehensive discussion on the topics of resilient backpropagation and principal component analysis (PCA). Section 7 provides an analysis and discussion of the obtained findings. In conclusion, Sect. 8 serves as the last segment of the paper.

2 Retinal Microaneurysms in Diabetics

Figure 2 depicts the morphological characteristics of diabetic retinopathy in contrast to a normal fundus picture: (a) illustrates the typical fundus picture, (b) highlights the location of the macula, (c) references the presence of macular edoema in the context of retinopathy, and (d) discusses the presence of cotton wool spots and haemorrhage in the fundus image.

3 Review on Fundus Imaging

Microaneurysms are an ocular disorder characterised by the presence of little red dots in the eye, often surrounded by yellow rings caused by leakage from blood vessels. Microaneurysms are characterised by the absence of further indications or

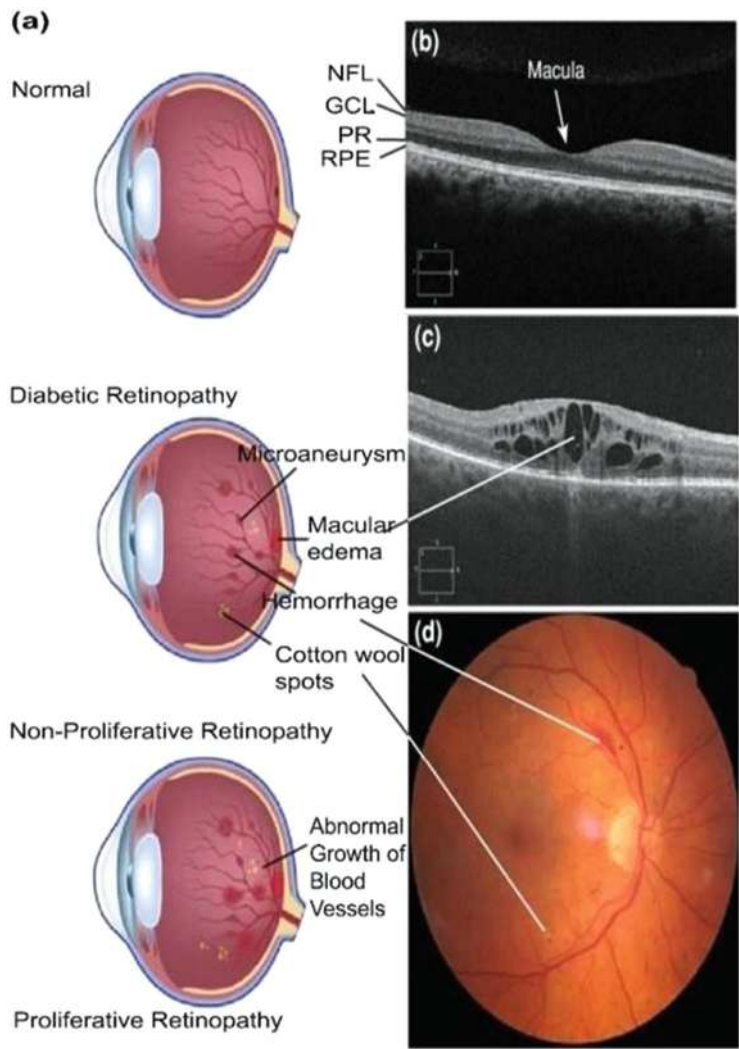


Fig. 2 Morphology of diabetic retinopathy vs normal

adverse consequences, and they do not have any discernible impact on visual acuity. Microaneurysms often function as the first indicators of diabetic retinopathy [1]. This suggests that the presence of microaneurysms in the eye is of great significance, since their detection may facilitate early intervention for diabetic retinopathy, hence reducing the risk of visual impairment. Diabetic retinopathy, a complication that arises in the latter stages of diabetes, affects a global population of 285 million persons, making it the predominant cause of worldwide visual impairment. Diabetic retinopathy occurs in individuals with diabetes as a result of elevated blood glucose

levels leading to damage to the microvasculature inside the retina. These damaged blood vessels cause leakage of fluid into the macula, resulting in impaired visual acuity. The newly formed blood vessels have the potential to leak into the posterior region of the eye, causing visual obstruction and subsequent loss of vision. At this stage, previously unutilised blood vessels develop on the surface of the macula in an effort to correct the irregular blood flow inside the retina. Although most people don't experience diabetic retinopathy until around 10 years after their diabetes first manifests, it's crucial to remember that diabetes' high blood sugar levels may cause diabetic retinopathy to manifest at any time. This observation suggests that it is increasingly important to initiate therapy as soon as possible at the onset of diabetic retinopathy. Furthermore, this implies that the presence of side effects such as microaneurysms, when combined with regular eye examinations, might be very crucial in the detection of diabetic retinopathy at its earliest stage. While there is currently no cure for diabetic retinopathy, early detection allows experts to use a broader range of treatments aimed at slowing down the progression of the problem. Interventions such as dietary modifications and physical exercise may provide support in regulating blood glucose levels and mitigating potential damage to ocular blood vessels. Moreover, in cases where blood vessels in the eye start to leak, laser surgery may be used to cauterise and effectively control or reduce the leakage.

The study of fundus pictures serves as a foundation for a comprehensive knowledge of retinal diseases associated with diabetes [2]. The morphology of the eye in the retinopathy state is shown in Fig. 3. Fundus imaging is the process of using reflected light to get a two-dimensional (2D) representation of the partly transparent, three-dimensional (3D), retinal tissues projected onto the imaging plane [3]. Therefore, the term 'fundus imaging' refers to the identification of any feature that arises in a two-dimensional image where the intensity of reflected light represents the quantity of light detected. Therefore, optical coherence tomography (OCT) imaging should not be confused with fundus imaging. However, the aforementioned modalities and methods all fall within the broad category of fundus imaging.

1. d-free photography and fundus photography both entail taking pictures that depict the intensity of reflected light within a certain wavelength range.
2. Colour fundus photography involves capturing an image that represents the combined reflection of red (R), green (G), and blue (B) wavelengths, as determined by the spectral sensitivity of the sensor.
3. Stereo fundus photography involves the capture of images that depict the combined intensity of reflected light from many distinct viewing angles, with the aim of achieving enhanced depth resolution.
4. The concept of SLO refers to the measurement of the cumulative amount of laser light reflected at a certain wavelength, obtained during a given time interval.
5. Adaptive optics SLO: The phenomenon of force exerted by light may be visualised as the cumulative effect of laser light that is reflected and optically corrected, taking into account the distortions in its wavefront.

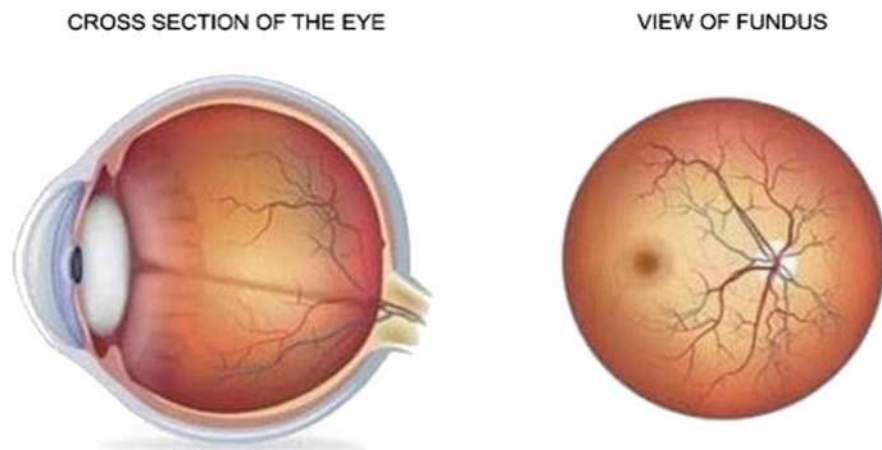


Fig. 3 Morphology in terms of cross-section of eye vs view on fundus pertaining to microaneurysm

The measurement of photons released by the fluorescein or indocyanine green fluorophore, which has been given to the person intravenously, is a component of the imaging procedures known as fluorescein angiography and indocyanine angiography.

Fundus imaging presents many distinct issues of a specialised nature. Given that the retina is often not illuminated, it is necessary for both external illumination entering the eye and the retinal image exiting the eye to pass through the pupillary plane. The primary technical problem in fundus imaging has always been the measurement of the student, often ranging from 2 to 8 mm in diameter. The complexity of fundus imaging arises from the fact that the presence of corneal and lenticular reflections reduces or eliminates image contrast, since the light and imaging bars are unable to adequately cover these reflections.

Therefore, the use of partitioned paths occurs inside the pupillary plane, resulting in optical apertures of just a few millimetres. It has been difficult to build imaging systems, and fundus imaging in particular requires costly equipment and expert ophthalmic photographers. Numerous important developments during the last decade or so have had a substantial influence on the subject of finance. This document presents a selection of fundamental image-processing algorithms. ArcNet, proposed in [13], restores cataractous fundus pictures without human annotation. This strategy makes restoration methods more feasible. ArcNet annotations are unnecessary. The high-frequency component from fundus images may replace segmentation to protect retinal components. The restoration model is trained on synthesised images and then adjusted for cataract photos. In Ref. [14], the classification of glaucomatous and normal fundus images is performed using a deep learning architecture based on CNN. It consists of 18 layers and is trained with the objective of extracting discriminative properties from fundus images. The architecture comprises one

fully connected layer, two max-pooling layers, and four convolutional layers. A two-stage tuning technique is suggested as a method for improving batch size and initial learning rate. Ref. [15] provides a novel framework for multi-task domain adaptation, which aims to automate the assessment of fundus picture quality. The proposed framework offers a method for evaluating the quality of images in a manner that is easily understandable, including both numerical scores and visual representations. This evaluation may be used to determine whether a image requires recapture, and if so, to make appropriate adjustments in real time. Ref. [16] introduced a novel unsupervised domain adaption framework, referred to as border and Entropy-driven Adversarial Learning (BEAL), with the aim of enhancing the performance of object detection (OD) and object classification (OC) segmentation. The primary focus of this framework is to address the challenges posed by ambiguous boundary regions.

4 Morphological Image Processing

A diverse array of image processing methods, referred to as morphology, are used to manipulate pictures by analysing their shapes [4]. Figure 4 depicts the morphological processes used for the processing of binary pictures. These procedures include (a) Original image, (b) Erosion operation, (c) Dilation operation, (d) Opening operation, and (e) Closing operation. Morphological image processing primarily involves the analysis of a pair of images with respect to predetermined geometric shapes referred to as structuring elements.

The word morphology refers to the scientific discipline that pertains to the study of the form and structure of organisms and plants. In order to capture and describe spatial forms, it is essential to extract from pictures relevant elements that are vital to morphological (shape-based) processing. Morphological preparation refers to a set of actions in which an object is subjected to organising elements, resulting in its reduction to a more revealing form. The organising components are fundamental geometric shapes that have been designed to represent various aspects of the information or the noise. Through the application of these organising components to the information, using unique mathematical combinations, one is able to enact morphological alterations on the data. The image in question refers to a two-dimensional representation that utilises light to enhance its visual impact. This particular image is classified into three distinct categories depending on the arrangement of numbers inside the two-dimensional cluster. These categories include the Two-fold Picture, Dim Tone Picture, and Colour Picture. The image exhibits a dual nature. The visual content of the Double Picture exhibits a mostly dark and white colour palette. Every and each pixel inside the image is represented by a binary value, namely, either a '0' or a '1'. A computer-generated image is referred to as a Parallel Picture when the range of grayscale values spans from 0 to 1.

The modification or intended altering of an existing picture is called image processing. The next step is to collect a picture, which, while it may seem clear,

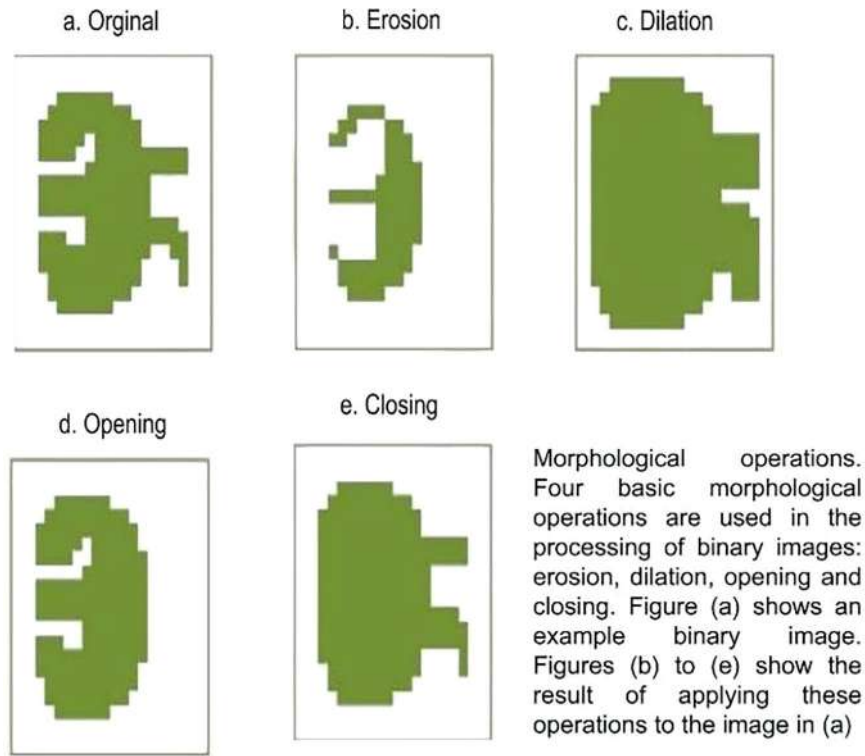


Fig. 4 Morphological operations on fundus image (a) Original (b) Erosion (c) Dilation (d) Opening (e) Closing

is not an easy operation since easily accessible image data is not always available. The software engineer need a fundamental approach for obtaining image data in a standardised and practical manner, sometimes referred to as a picture record organisation. The image repository system categorises and determines the storage capacity of images, while also providing further information pertaining to storage capacity via pixel values. The image record is composed of two main sections: the Header Section and the Data Segment. The image’s dimensions will at the very least be mentioned in the header. It is important to provide information on the size of a image when displaying or preparing it, since the absence of such information might be considered unorthodox. A signature or enchantment number appears in the header of the majority of record designs. a brief string of bytes intended to distinguish the data as an image with a particular format.

The organising component compares the matching neighbourhood of pixels with the organising component, which is present at all places or probable locations within the Double Picture. The concept of ‘Binary’ correction serves as the foundation for the morphological operation. When the procedure is characterised by coherence

rather than being focused on numerical calculations. Let us consider two 3×3 matrices for the purpose of organisation. In the provided image, there are three designated spots, denoted as A, B, and C, where the S1 and S2 organising components are intended to be placed.

The use of binary images was employed to assess the fitting and hitting of organising components S1 and S2. FIT: The organising component is considered to be in alignment with the image if, for every pixel in the organising component that is set to '1', the corresponding pixel in the image is also '1'. Regarding the above example, it can be seen that both S1 and S2 are suitable for the image at point 'A'. It is important to note that any organising component pixels that are set to '0' are disregarded when determining the suitability. S2 corresponds to the image labelled as 'B', whereas neither S1 nor S2 corresponds to the image labelled as 'C'. HIT: A structural component is said to have an HIT and Picture property if, for each pixel inside the component that is assigned a value of '1', the corresponding pixel within the Image is also assigned a value of '1'. In addition, we exclude picture pixels in which the corresponding organising component pixel is '0'. In the aforementioned instance, S1 and S2 made contact with the picture located in neighbourhood 'A'. The same principle applies at point 'B'. However, in neighbourhood 'C', S1 successfully impacts the image.,

Additionally, the set operations of Intersection and Union can be replaced with the Boolean operators 'AND' and 'OR'. Dilation is a geometric transformation that changes the scale of a figure while maintaining its geometry and proportional relationships. Dilation is a morphological process characterised by the enlargement of regions within a picture. Dilation is a phenomenon characterised by the process of expansion or enlargement of objects. The proliferation of these entities is dependent on the choice of the structural component and its way of propagation [3]. Dilation is a process that increases the size of an item by including additional pixels around its periphery. The mathematical representation of the dilation operation applied to an image 'A' using a structural element 'B' is denoted as $A \circ B$. In order to calculate the Dilation, the element 'B' is positioned so that its origin is located at pixel coordinates (x, y) , and the corresponding rule is then applied. If the variable 'B' makes contact with the variable 'A', the function $g(x, y) = \text{zero}$. Subsequently, iterate through all pixel coordinates. The process of dilation involves generating a new picture that displays the positions within the input image where the origin of a structuring element intersects. This process involves the addition of a pixel layer to an object, resulting in its enlargement. The process of dilation involves the addition of pixels to both the inner and outside bounds of regions. As a result, the holes encompassed by a single area will be reduced in size, and the gaps between various regions will be diminished. Dilation has the tendency to effectively fill in minor incursions inside the borders of a given area.

Both the size and the shape of the organising component have an impact on the results of Enlargement. A morphological technique called enlargement may be used with both parallel and dark-toned pictures. There is a discernible impact on the process of delineating the outside perimeters of the provided photos. In the context of digital image processing, a binary image refers to a kind of image that

consists of just two possible pixel values, often represented as black and white. The dilation operation may be defined as follows: $D(A, B) = A \hat{\cup} B$. In the given context, A represents the image, and B denotes the organising element inside a 3×3 arrangement. There is a need for various organising components in order to expand the overall perspective. Erosion: The phenomenon of erosion is responsible for the gradual degradation and contraction of geographical regions. The process of erosion induces the rebound of things. The magnitude of their recoil is contingent upon the selection of the organising component. Erosion is a process that reduces the size of an image by selectively removing pixels from its boundaries [3]. The process of disintegration of an image 'A' by an organising component 'B' may be denoted as $A \ominus B$. In order to calculate the erosion, the position of 'B' is adjusted so that its starting point aligns with the pixel coordinate (x, y) in the image. Subsequently, the algorithm is executed. The variable 'B' is compatible with 'A', where $g(x, y)$ represents an alternative value. Rewrite for all values of x and y representing pixel coordinates. Erosion is a geological process that results in the formation of a contemporary image, which delineates the various regions where organising components are situated in relation to the input picture. The erosion procedure is seen to remove a layer of pixels from an image, causing it to shrink in size. The erosion of pixels occurs at both the internal and exterior edges of regions. Erosion has the capacity to widen the gaps inside a certain locality, as well as increase the distance between different regions. Furthermore, disintegration is likely to result in the elimination of small outbursts near site borders.

The outcome of erosion is contingent upon the size of the organising component, with larger organising components exerting a more pronounced influence. Furthermore, the impact of erosion with a substantial organising component is similar to that achieved through repeated iterations of erosion using a smaller organising component of identical shape. Erosion is a morphological technique that may be applied to images with parallel and dim characteristics. The process of extracting the internal boundaries inside a given image yields discernible variations. In the context of Twofold Pictures: The erosion operation may be defined as follows: $E(A, B) = A \ominus B$. The image is denoted by the variable A . The organising component of the arrangement $3 * 3$ is denoted by B .

A multitude of organisational elements are necessary for the dissolution of the whole image. OPENING: The organised evacuation of border pixels inside a picture area. An successful administrative technique may be achieved by combining the processes of erosion and dilation. The act of opening separates the objects. As is well understood, the process of expansion results in the enlargement of an image, whereas erosion leads to its contraction. The act of opening generally results in the smoothing of a picture's shape, the breaking of boundary isthmuses, and the elimination of lean bulges. The act of opening an image 'A' via an organising component 'B' is denoted as $A \circ B$ and may be described as a Disintegration followed by an Expansion. This process is represented as $A \circ B = (A \ominus B) \hat{\cup} B$. The process of obtaining the opening operation involves performing an enlargement on a dissolved image. The purpose is to enhance the smoothness of the curves in the image. The act of creating larger distances between items that are in close proximity

serves to separate objects that are in contact but should not be, while also increasing the gaps inside objects.

The first phase consists of one or more erosions followed by a subsequent enlargement. **CLOSING**: The act of closing refers to the systematic process of filling in the border pixels of a picture's region. An successful administrative approach may be achieved via the combination of Erosion and Dilation techniques. In conclusion, it is important to establish a connection between the objects. The process of closing often involves the smoothing of regions on forms. In contrast to opening, closing typically involves the consolidation of boundaries and the elimination of elongated and narrow gaps. It also serves to fill small gaps and address any holes present inside the form. The process of closing an image 'A' using an organising element 'B' is denoted as $A \bullet B$ and defined as a combination of a dilation followed by an erosion. Mathematically, it may be expressed as $A \bullet B = (A \hat{A} B) \ominus B$. The process of obtaining a closing operation involves applying an erosion operation on an expanded image. The act of closing involves the joining of items that have been damaged or separated, as well as the filling of gaps that are considered undesirable in these objects. The process of closing involves the sequential application of one or more dilations followed by one erosion operation.

5 Resilient Backpropagation

The Rprop algorithm, which is a variant of the back-propagation algorithm, is a flexible computational method that may be used for training neural networks. It has similarities with the conventional standard back-propagation algorithm. However, it has two major focus points over the backpropagation algorithm. First, it is worth noting that training using Rprop often exhibits a higher speed compared to training with back propagation. In contrast to back propagation, Rprop does not need the specification of any free parameter values, such as the learning rate (and often an optional momentum term). One of the primary drawbacks of Rprop is its higher computational complexity compared to backpropagation.

A complicated scientific structure known as a neural network processes numerical inputs and produces numerical outputs. The input values, the quantity of hidden processing nodes, the activation functions of the hidden and output layers, as well as a set of weights and bias values, all affect the output values. A fully connected neural network has $(m * h) + h + (h * n) + n$ weights and biases, with m inputs, h hidden nodes, and n outputs. With four inputs, five hidden nodes, and three outputs, for instance, a neural network contains 43 weights and biases ($4 * 5 + 5 + (5 * 3) + 3$). The process of creating a neural network involves choosing the right weights and biases, which cause the network's calculated outputs to closely resemble the predicted outputs for a particular training dataset made up of well-known input and output values. The most popular method for training neural networks is the back-propagation algorithm. A factor called the learning rate must be taken into account when using the backpropagation technique. The value given to the learning

rate has a significant impact on the sensitivity of the efficacy of back propagation. In 1993, analysts created the Rprop algorithm with the intention of improving the back-propagation technique. Figure 1 is advised as a reference to have a thorough grasp of Rprop and to provide a preview of the material covered in this tutorial. The example programme starts by creating 10,000 fake data bits rather than utilising real data.

6 Principal Component Analysis (PCA)

Principal components analysis (PCA) is a mathematical procedure used to transform the columns of a dataset into a new collection of orthogonal features known as principal components. Through this process, a substantial portion of the data within the whole dataset is effectively condensed into a reduced number of feature columns. This enhances the reduction of dimensionality and the ability to envision the separation of classes or clusters, if present. In this tutorial, the implementation of principal component analysis (PCA) using the scikit-learn library will be shown. Subsequently, a detailed step-by-step explanation, accompanied by code, will be provided to elucidate the underlying principles of the PCA algorithm in a comprehensible manner.

PCA is a frequently employed method for reducing the dimensionality of a dataset. It involves transforming the columns of the dataset into a new collection of features called principal components (PCs). The data included inside a column represents the cumulative total of the variations it encompasses. The primary goal of principal components analysis is to represent the information contained within a dataset using the fewest possible number of variables.

The framework of the suggested technique is shown in Fig. 5. The framework of the approach suggested in this study is shown in the image above. The fundus picture undergoes first processing for analysis. Subsequently, two processes are performed on every individual picture. The picture is partitioned into a dataset specifically designed for the purpose of detecting tiny objects, while the other component generates the relevant vascular segmentation map. The output is transferred to the basic feature extraction network, and the multi-layer feature map that is produced is then combined using a multi-layer attention mechanism. The region proposal network (RPN) is then used to classify and modify the feature map in order to provide preliminary results for the detection of small objects. The vessel segmentation map and the starting prediction box coordinates are utilised to determine the confidence value for each prediction frame. In the last step, the confidence and probability values undergo a double screening process by majority voting, resulting in the ultimate outcome of tiny item detection.

The experimental dataset and the hospital data were analysed, and it was discovered that the diameters of MA had a distinct distribution between 10 and 40 pixels. The distribution's details are shown in Table 1 below. The aforementioned findings were validated by a cohort of ophthalmologists affiliated with Chengdu

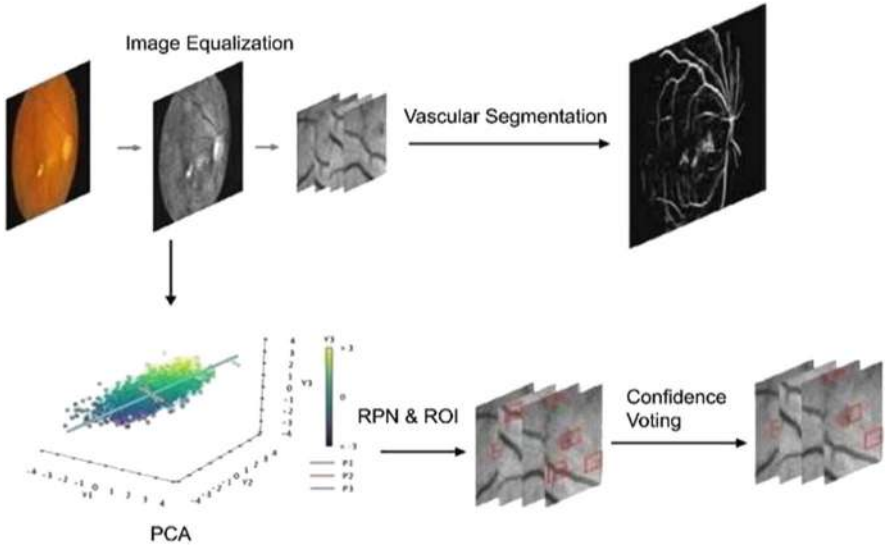


Fig. 5 Flow diagram of proposed framework

Table 1 Size distribution of Microaneurysms

Diameter (pixels)	<5	5–10	10–20	20–30	30–40	40–50	>50
%	0.02	1.58	47.40	39.70	9.80	1.2	0.3

Table 2 Distance distribution of microaneurysms

Distance (pixels)	0–10	10–20	20–30	30–40	40–50	50–100	10
%	80.70	7.60	4.20	1.30	2.20	2.20	1.

University at TCM. In order to optimise the functioning for the majority of medium-sized organisations, it is necessary to implement certain enhancements. The significance of flat layers conv2 and conv3 in the range of 10–40 pixels is crucial for the identification of MA. Consequently, the process of feature fusion and target identification was executed by using features extracted from layers conv2 to conv5. The experimental component of this study examines the significance of the conv2 and conv3 layers in the context of MA identification.

The detection of MA in secondary screening may be improved by the use of spatial reliability, as shown by previous studies [5, 6]. By including the target’s contextual context, the accuracy and effectiveness of MA detection can be enhanced. The term ‘MA’ refers to a roughly spherical item characterised by the presence of regular ridges of blood vessels, as well as little blood spots that result from the dilatation and leaking of blood vessels. Consequently, malignant angiogenesis often takes place in close proximity to vascular structures. The distance distribution matching the experimental statistics is shown in Table 2. The criteria for differentiating the MA is based on the distance between the MA and the blood vessels.

7 Result and Analysis

The identification of existing indicators, such as microaneurysms, through the examination of fundus images, in conjunction with appropriate therapy, offers significant benefits in terms of mitigating the risk of problems associated with diabetic retinopathy and reducing the likelihood of visual impairment.

The suggested computation consists of three parts.

- Enhancement of images through morphological processing.
- The process of removing and eliminating redstructures, namely, blood vessels, is preceded by identifying their position and then removing any bright artefacts.
- Finally, the authentic identification of microaneurysm candidates and other structures is achieved by the implementation of an extraction set (Fig. 6).

Two publicly accessible datasets with both normal and infected images were used to effectively assess the proposed technique. The images from the Diaretdb1 database showed a sensitivity of 89.22%, specificity of 91%, and accuracy of 92% for the identification of microaneurysms. For the e-optha database, the algorithm assessment for microaneurysm identification has a sensitivity of 83% and a specificity of 82%.

Using the detection sensitivity as an assessment measure, Table 3 compares the approach suggested in this work with other methods for identifying MA. This comparison is based on the data in this paper about finding small things. As shown, our method has a maximum sensitivity of 89.2.

The accurate identification of the quantity of microaneurysms in a computerised discovery system is associated with the progression of retinal infections and the timely diagnosis of such diseases. The use of genuine microaneurysm localisation proves to be a helpful tool in the screening of colour fundus images, resulting in time efficiency when counting microaneurysms in accordance with the Diabetic Retinopathy Grading Criteria.

As shown in Table 3, a number of MA detection algorithms built on various techniques were selected for experimental evaluation. Sensitivity was chosen as the main indicator in our investigation. To identify various types of lesions, including exudates, bleeding, and MA area, a single network was used [7]. The obtained findings were then compared with those of blood vessels [8].

The first steps include the removal and subsequent re-extraction of the relevant components. Furthermore, the identification of MA may be achieved by the use of methodologies such as pixel-by-pixel classification [9, 10]. This comparative analysis demonstrates that the suggested methodology exhibits a higher degree of efficacy in the detection of MA within fundus images.

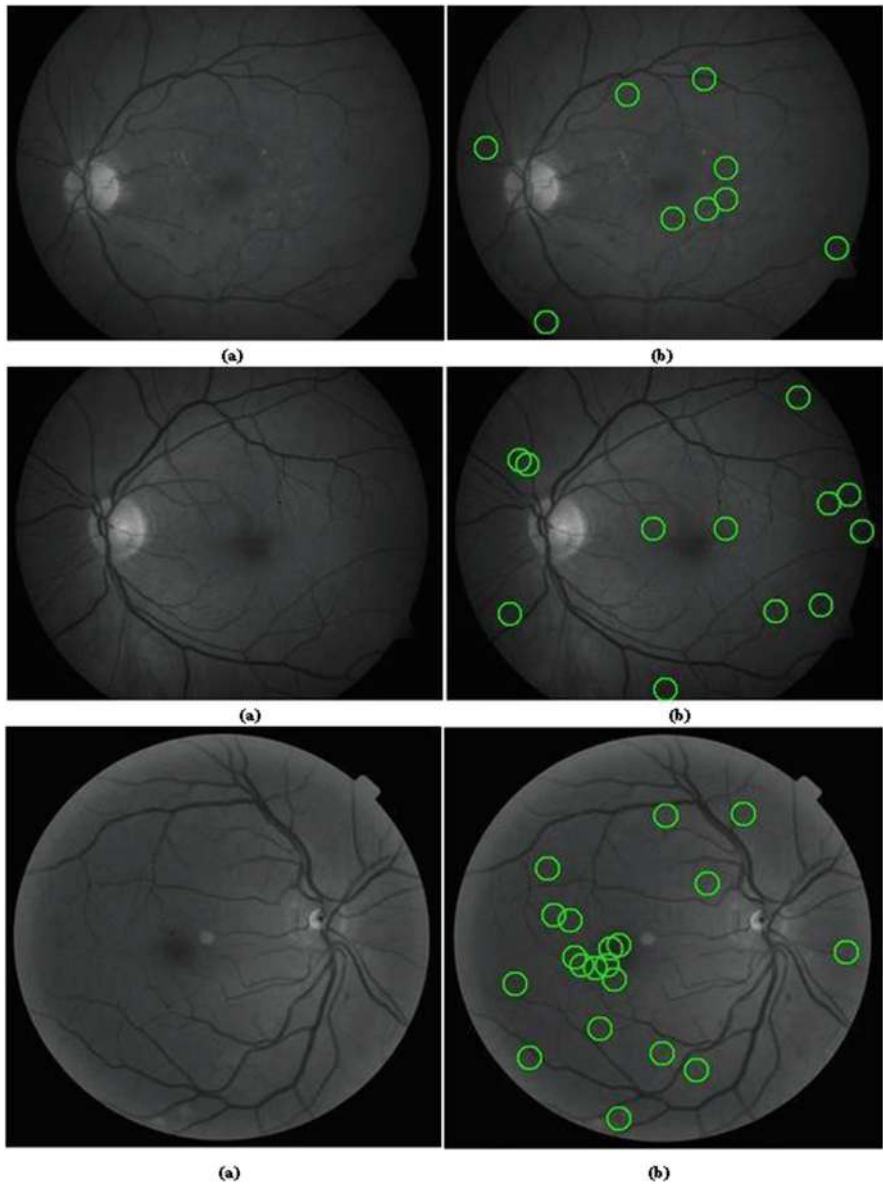


Fig. 6 Microaneurysms detection results (a) green channel images (b) detection results

Table 3 Comparison of performance of proposed system in terms of sensitivity with other literature results

Method	Dataset	Sensitivity
Tan [7]	CLEOPATRA	0.461
Dai [8]	ROC [11]	0.691
Khojasteh [9]	DIARETDB1 [12]	0.85
Adal [10]	DIARETDB1 [12]	0.646
Our proposed work	PCA-Rprop	0.892

8 Conclusion

Diabetic retinopathy is identified as a primary determinant in the escalation of visual impairment. The timely and accurate identification of microaneurysms is crucial for the diagnosis and classification of diabetic retinopathy. The paper introduces a unique approach for the automated identification of microaneurysms in retinal fundus images. The recommended technique consists of four fundamental steps: preprocessing, candidate extraction, feature extraction, and classification. Future research focuses on the application of deep neural network models for the analysis of microaneurysms in diabetic retinopathy using fundus photographs. The main objective is to extract features from two distinct deep neural networks. Subsequently, a crucial technique known as feature embedding is employed to combine the extracted features, aiming to enhance the performance of classification.

Acknowledgments

Additional Information and Declaration Funding information is not applicable/No funding was received.

Competing Interests All authors declare that they have no competing interests.

References

1. Henry E. Wiley, Frederick L. Ferris, in [Retina \(Fifth Edition\)](#), 2013

2. Renoh Johnson Chalakkal, Sheng Chiong Hong, in [Diabetes and Fundus OCT](#), 2020

3. Michael Abràmoff, Christine N. Kay, in [Retina \(Fifth Edition\)](#), 2013

4. P Soille. “Morphological Image Analysis, Principlesand Applications”, 1999

5. Wu Y., Ma W., Su Q., Liu S., Ge Y. Remote sensing image registration based on local structural information and global constraint. *J. Appl. Remote Sens.* 2019; 13:016518. <https://doi.org/10.1117/1.JRS.13.016518>.

6. Mughees A., Tao L. Multiple deep-belief-network-based spectral-spatial classification of hyperspectral images. *Tsinghua Sci. Technol.* 2018; 24:183–194. <https://doi.org/10.26599/TST.2018.9010043>.


7. Tan J.H., Fujita H., Siva Prasad S., Bhandary S.V., Rao A.K., Chua K.C., Acharya U.R. Automated segmentation of exudates, hemorrhages, microaneurysms using single convolutional neural network. *Inf.* 2017; 420:66–76. <https://doi.org/10.1016/j.ins.2017.08.050>.

8. Dai B., Wu X., Bu W. Retinal microaneurysms detection using gradient vector analysis and class imbalance classification. *PLoS ONE.* 2016; 11: e0161556. <https://doi.org/10.1371/journal.pone.0161556>

9. Khojasteh P., Aliahmad B., Kumar D.K. Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms. *BMC Ophthalmol.* 2018; 18:288. <https://doi.org/10.1186/s12886-018-0954-4>
10. Adal K.M., Sidibé D., Ali S., Chaum E., Karnowski T.P., Mériaudeau F. Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning. *Comput. Methods Progr. Biomed.* 2014; 114:1–10. <https://doi.org/10.1016/j.cmpb.2013.12.009>.
11. Niemeijer M., Van Ginneken B., Cree M.J., Mizutani A., Quellec G., Sánchez C.I., Zhang B., Hornero R., Lamard M., Muramatsu C., et al. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. *IEEE Trans. Med. Imaging.* 2009;29:185–195. <https://doi.org/10.1109/TMI.2009.2033909>.
12. Kauppi T., Kämäräinen J.K., Lensu L., Kalesnykiene V., Sorri I., Uusitalo H., Kälviäinen H. Constructing benchmark databases and protocols for medical image analysis: Diabetic retinopathy. *Comput. Math. Methods Med.* 2013; 2013:368514. <https://doi.org/10.1155/2013/368514>.
13. Li, H., Liu, H., Hu, Y., Fu, H., Zhao, Y., Miao, H., & Liu, J. (2022). An annotation-free restoration network for cataractous fundus images. *IEEE Transactions on Medical Imaging*, 41(7), 1699–1710.
14. Elangovan, P., & Nath, M. K. (2021). Glaucoma assessment from color fundus images using convolutional neural network. *International Journal of Imaging Systems and Technology*, 31(2), 955–971.
15. Shen, Y., Sheng, B., Fang, R., Li, H., Dai, L., Stolte, S., ... & Shen, D. (2020). Domain-invariant interpretable fundus image quality assessment. *Medical image analysis*, 61, 101654.
16. Wang, S., Yu, L., Li, K., Yang, X., Fu, C. W., & Heng, P. A. (2019). Boundary and entropy-driven adversarial learning for fundus image segmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I* 22 (pp. 102–110). Springer International Publishing.

Pilot Super-Resolution Network (PSRN)-Based Mango Fruit Classification



P. V. Yeswanth , Sammeta Kushal, V. Tharun Kumar, N. R. Ackshay, Ravindra Gangudi, Molapally Tharun Kumar, S. Deivalakshmi, Y. Thanya, and K. M. Lokesh Kumar

Abstract Mango as a fruit has many varieties, and identification of the same with the naked eye can be tiresome. Pilot Super-Resolution Network (PSRN) model is proposed to produce super-resolution images from that of the low-resolution mango fruit images and identify the varieties in this paper. The novel PSRN model is executed on the Mango Variety and Grading dataset for super-resolution factors 2, 4, and 6. The metrics recorded include PSNR values of 30.37, 32.936, and 35.428, SSIM values of 0.8275, 0.9146, and 0.9279, and classification accuracy values of 99.45, 98.86, and 97.64 on factors 2, 4, and 6, respectively.

Keywords Super resolution · Pilot Super-Resolution Network (PSRN) · Mango fruit classification · Low resolution · Image classification

1 Introduction

Millions of people all over the world enjoy the delicious and well-liked fruit mango. Its sweet, juicy flesh and distinct flavor make it a popular addition to a variety of cuisines and beverages, and it is also high in nutrients and health benefits. It's no surprise that this fruit is so popular, with over 400 different types, each with its own distinct taste, texture, and look. Mangoes are divided into species depending on features such as color, shape, and size. The Chaunsa, Anwar Ratol, Fajri, and others. Each of these kinds has a distinct flavor profile and is best suited to specific applications.

P. V. Yeswanth (✉) · S. Kushal · V. T. Kumar · N. R. Ackshay · R. Gangudi · M. T. Kumar · S. Deivalakshmi

Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India

Y. Thanya · K. M. Lokesh Kumar

Department of Computer and Communication Engineering, Amrita Vishwa Vidyapeetham, Chennai, Tamil Nadu, India

Image Super Resolution (SR) [1] is a technique for improving an image's resolution and quality over its low-resolution predecessor. It is a significant issue in computer vision and image processing, with applications ranging from surveillance to agriculture [2] to satellite imaging [3] and beyond. The purpose of image SR is to restore high-frequency information that is lost during the image's down-sampling process.

Reference-based super resolution (RefSR) [4–6] is a type of image SR in which the high-resolution image is reconstructed utilizing an additional reference image. The reference image is often a higher-resolution version of the low-resolution image. To increase the quality of the reconstructed image, RefSR methods use the similarity between the two images. The fundamental idea behind RefSR is to supply more information to the SR algorithm by using a high-resolution reference image. This additional information could be high-frequency features, edges, or textures that are lacking from the low-resolution image. There are two types of RefSR methods: reconstruction-based and hallucination-based.

Reconstruction-based methods recover the HR image using the reference image by solving an optimization problem that involves the Low Resolution (LR) input image and the High Resolution (HR) reference image [6]. Hallucination-based approaches, on the other hand, use machine learning algorithms to build the HR image using the LR input image and the HR reference image.

These approaches, however, simply consider the similarities between the reference and input LR images. There are two components to resolve the problem: one way to do this is to eliminate inconsistencies, such as identical object-matching problems [7]. On the other hand, acceptable content for the combination of features must be chosen. As a result, we offer a unique pilot super-resolution method for a reference-based super resolution that makes use of complementary information fusion.

2 Related Work

2.1 Related Work on Single Image Super Resolution

Image Super Resolution (SR) [8] is a critical task in computer vision that entails improving an image's resolution beyond its original size or resolution. Various methods for image SR have been developed in recent years, and great progress has been made in this sector. Interpolation-based approaches are the most often utilized classical Single Image Super Resolution (SISR) methods. These methods upsample a low-resolution image to a higher resolution by using interpolation techniques such as bicubic, bilinear, or nearest-neighbor interpolation [9]. Although simple and quick, these approaches frequently result in hazy and smooth photos. Researchers have proposed numerous approaches to circumvent this constraint, such as edge-Pilot interpolation and super resolution via patch-based synthesis [10].

Machine learning techniques, such as regression and classification, were used in learning-based methods to discover the relationship between the LR and HR images. The Deep Convolutional Neural Network (DCNN)-based method [11], which employs a deep neural network to understand the mapping between the LR and HR images, is one of the most prominent methods in this area.

Deep learning-based technologies have recently revolutionized the domain of SISR [15–19]. The Generative Adversarial Network (GAN)-based method [12], which employs a GAN to produce high-resolution images from low-resolution photos, is one of the most successful deep learning-based methods. Super-Resolution Generative Adversarial Network (SRGAN) [13], one of the most successful GAN-based algorithms, employs a GAN to develop HR images with realistic textures and details. The Residual Dense Network (RDN) [14] is another successful deep learning-based system that generates high-quality images by combining a dense network with attention processes.

2.2 Reference-Based Super-Resolution

Reference-based super resolution (RBSR) is the technique used for improving output quality by using a high-resolution reference image or series of images. The reference image is utilized as a guide in this technique to help improve the resolution of a low-quality photograph. The reference image can be a different viewpoint of the same scene or a similar photograph recorded under different conditions. Baker and Kanade proposed one of the first RBSR algorithms in 2002, which used an HR reference image to estimate the motion of the LR image and then utilized this motion to develop HR images. Since then, several techniques for RBSR have been developed, including interpolation-based, learning-based, and optimization-based methods.

The LR image is interpolated using the reference image as a guide in interpolation-based RBSR approaches. Bicubic interpolation, a high-quality interpolation method that can be used to improve image resolution, is the most widely used interpolation method. This procedure, however, frequently leads to a loss of detail and sharpness in the final output. In contrast, learning-based RBSR approaches employ a machine learning model to learn the link between the low-resolution image and the reference image. To learn the association between image pairings, the model is trained on a set of HR and LR image pairs. Once trained, the model can be used to produce HR images from LR images by referencing the reference image. An optimization technique is used in optimization-based RBSR approaches to discover the high-resolution image that best fits the low-resolution image and the reference image. These approaches entail defining an objective function that assesses the similarity of the LR image to the HR image, as well as the similarity of the HR image to the reference image. Following that, the objective function is optimized to identify the high-resolution image that best fits both the LR image and the reference image.

Deep learning-based approaches for RBSR have gained popularity in recent years. Deep neural networks are used in these methods to familiarize themselves with the relationship between the LR image and the reference image. Super-Resolution Generative Adversarial Network (SRGAN), which employs a generative adversarial network (GAN) to produce HR images from LR images and a reference image, is a deep learning-based algorithm for RBSR that proves to be the most efficient.

3 Methodology

3.1 Complementary Information Fusion

Figure 1 depicts the architecture of the proposed Pilot Super-Resolution Network (PSRN). The major goal is to consciously add the reference features to that of the LR features using the reference channel and attention block (AB) and residual block (RB). We designed the AB to selectively extract useful Ref features for LR images to utilize channel-wise and spatial data as shown in Fig. 2.

First, the low-resolution extraction stage is deployed, where we employ the EDSR network structure to retrieve the HR details of the input LR image via the ResBlock and pixel shuffle up-sampling modules. Second, the HR and switched Ref

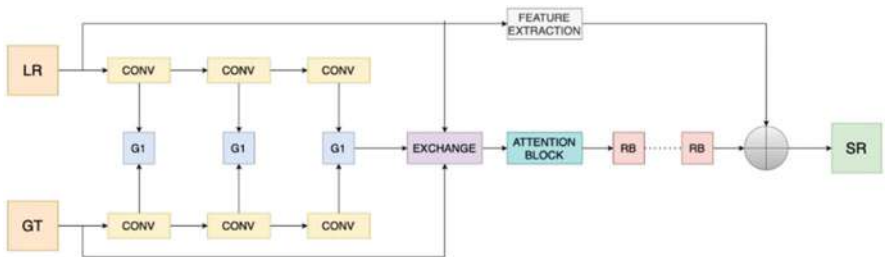


Fig. 1 Pilot Super-Resolution Network (PSRN)

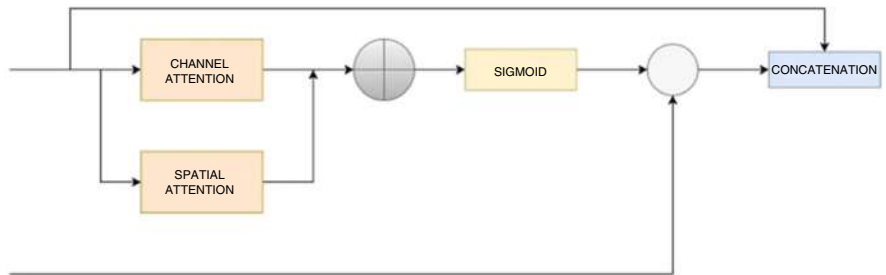


Fig. 2 Attention block

characteristics are entered into the AB module. Using the AB module, we adaptively pick Ref characteristics according to channel and spatial attention. The chosen Ref characteristics are then linked with the high-resolution characteristics to complete the input image. Lastly, we employ the RB block for feature modification to generate the HR picture result.

The two characteristics are usually connected directly in the conventional fusion mechanism:

$$F_l = \varphi l - 1 \parallel \psi l - 1 \quad (1)$$

where F_l represents the l th layer feature, and $\varphi l - 1$, $\psi l - 1$ refer to the input data and swapped Ref features, respectively. The connecting function is represented by \parallel . Even though fresh data is integrated with the linked characteristics, it is identical to the initial data, and it is simple to produce new details instead of recovering details from the initial image.

As a result, influenced by residual block, our approach integrates channel attention (CA) and spatial attention using the initial image features and joins the Ref features to generate the attention map following feature fusion:

$$F_{\text{Ref-AB}} = F_{\text{LR}} \parallel (\sigma(f_{\text{CA}}(F_{\text{LR}}) + f_{\text{SA}}(F_{\text{LR}})) \cdot F_{\text{Ref}}) \quad (2)$$

F_{LR} signifies the LR and F_{Ref} denotes the Ref feature maps. (\cdot) denotes the sigmoid function. The term $f_{\text{CA}(\cdot)}$ denotes channel attention and $f_{\text{SA}(\cdot)}$ refers to spatial attention mechanisms.

Channel attention (CA) A deep network's features contain many forms of information distributed over channels and spatial regions, each of which assists differently to high-frequency detail recovery. Average pooled and Max pooled characteristics are used by Channel Attention Module (CAM) at the same time to boost network representation power. Attention Block (AB) emphasizes informative feature maps for global feature recalibration, where per-channel summary data are created. For the pooling method, the variance is chosen instead of the average. Because SR eventually intends to restore high-frequency elements from images, high-frequency data about the channels can assist in creating attention maps. Excitation and scaling are carried out just as in RCAN.

$$f_{\text{CA}}(x) = W_{\text{CA}}^1 \left(\delta \left(W_{\text{CA}}^0 (f_{\text{var}}(x)) \right) \right) \quad (3)$$

where ReLU activation function is referred by δ , $W_{\text{CA}}^0 \in R^{C_r \times C \times 1 \times 1}$, and $W_{\text{CA}}^1 \in R^{C \times C_r \times 1 \times 1}$. $f_{\text{var}}(\cdot)$ refers to variance pooling.

Spatial Attention (SA) The data, which is given as input, is distributed across spatial positions along with that of feature maps. Edges and texture regions typically hold more high-frequency information. As a result, we devote greater attention to obtaining the data in the ground truth image to reconstruct the crucial textures of

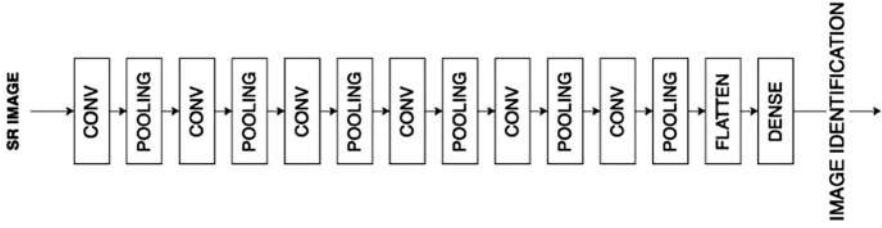


Fig. 3 Mango fruit classification model

the input image. It is beneficial to utilize the spatial attention mechanism to assist the network in having a discriminative capacity for distinct local regions and in turn weigh more attention to the regions that are more essential and difficult to recover.

To improve network representations, a number of spatial attention techniques were investigated. Average-pooling and max-pooling procedures were concatenated by the Channel Attention Module (CAM) along the channel axis for an efficient feature descriptor. To build a spatial attention mask, Attention Block (AB) utilized a two layer neural network accompanied by a sigmoid function that used the input's feature channel interdependence. RB employed only one 3×3 depth-wise convolution and did not hoard data per channel to save on channel-specific properties.

$$f_{S\ A}(x) = W_S^1 \left(\delta \left(W_S^0(x) \right) \right) \quad (4)$$

where Rectified Liner Unit (ReLU) activation function is referred to by δ , W_S^0 is the convolution function of filter size of 3×3 , and W_S^1 is the convolution function of filter size of 1×1 .

3.2 Classification of Mango Fruit

To produce good predictions, many previous classification systems relied solely on images with high resolution. Low-resolution images are then enhanced to give high-resolution images as a result that are then categorized. The model shown in Fig. 3 is used to differentiate Super-Resolution Mango fruits.

3.3 Data Set

The Pilot Super-Resolution Network (PSRN) is evaluated using the publicly available Mango Variety and Grading Dataset as mentioned in Table 1. This data

Table 1 Dataset information

Type of fruit	No. of images
anwar_ratool	200
chaunsa_black	200
chaunsa_white	200
chausna	200
dosehri	200
fajri	200
langra	200
sindhri	200

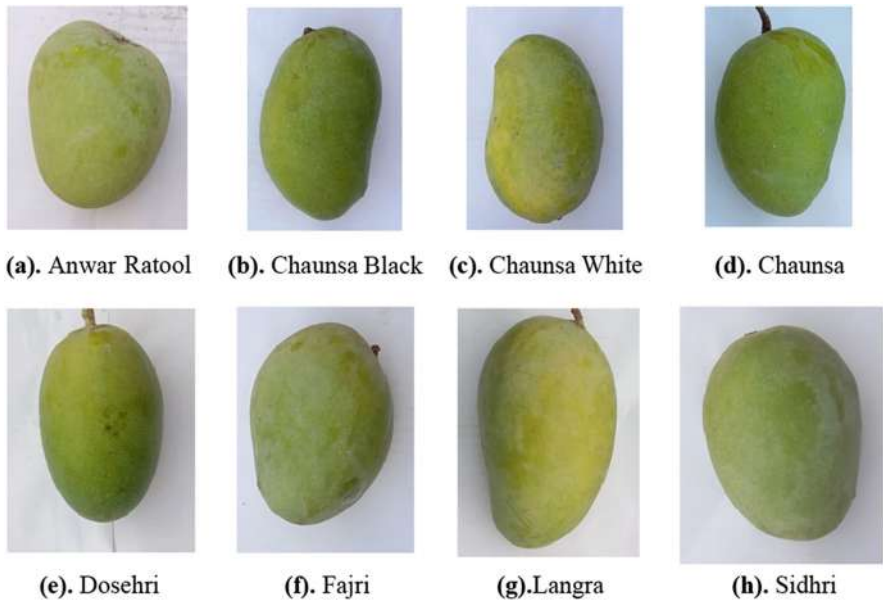


Fig. 4 (a) Anwar Ratool (b) Chaunsa Black (c) Chaunsa White (d) Chaunsa (e) Dosehri (f) Fajri (g) Langra (h) Sindhri

set is divided into eight classes. The backgrounds in these images are all uniform as shown in Fig. 4. There is also little variation in intensity.

In this case, 80% of all images are selected at random for training, 10% for testing, and 10% for validation.

3.4 Experimentation

Experimentation is carried out by capturing all 1600 images and then downsampling them by various variables to get low-resolution images, as illustrated in Fig. 5.

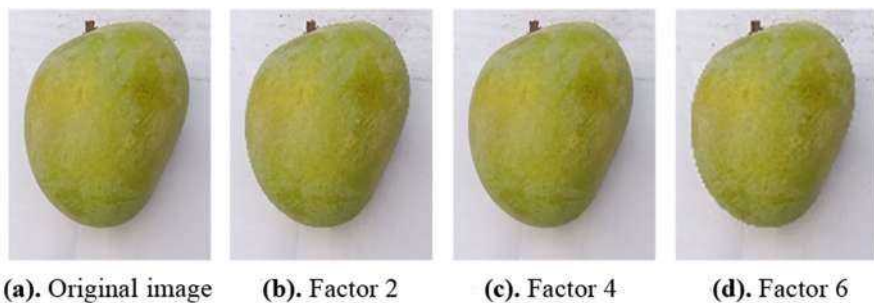


Fig. 5 (a) Original image (b) Factor 2 (c) Factor 4 (d) Factor 6

These low-resolution images will be sent into our classification model as input. To obtain high-resolution images, these are passed into the super-resolution block (PSRN). These images are then run via the classification block in Fig. 3 to identify the mango varieties. The super-resolution block PSRN uses the metrics PSNR and SSIM for training in this case. For training, the MSE loss function was utilized.

Figure 6 shows the super-resolution block training loss with respect to the number of epochs for different super-resolution factors 2, 4, and 6. Figure 7 shows the classification block training loss with respect to the number of epochs for different super-resolution factors 2, 4, and 6.

4 Results

4.1 Qualitative Analysis

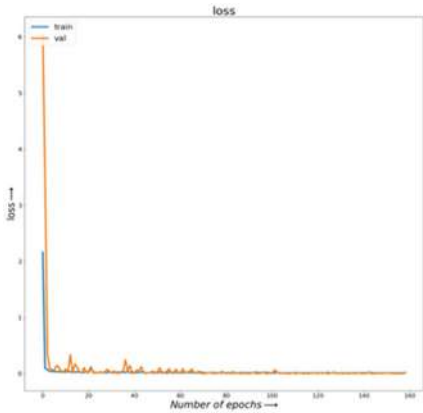
Figure 8 displays a qualitative study of SR images generated from LR photos using various scaling factors. Tables 2 and 3 show the performance analysis of our model for these various scaling variables. Tables 2 and 3 show that the proposed model predicts diseases with greater than 98.65% accuracy, with PSNR greater than 30 and SSIM greater than 0.8.

4.2 Quantitative Analysis

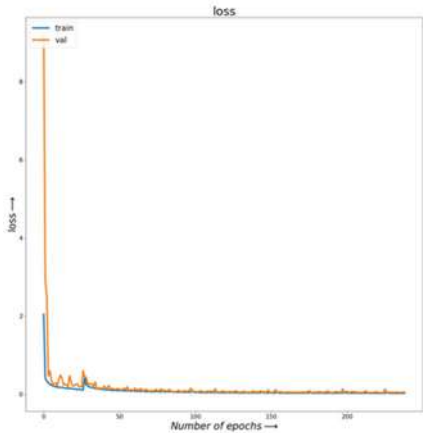
From the quantitative analysis it can be observed that, even though the PSNR increases from factors 2–6, the classification accuracy decreases as the classification model finds difficulty in detecting the type of mango fruit with more SR factor.

As the proposed PSNR model improves the resolution of the image its classification accuracy outperformed the existing models as shown in Table 4.

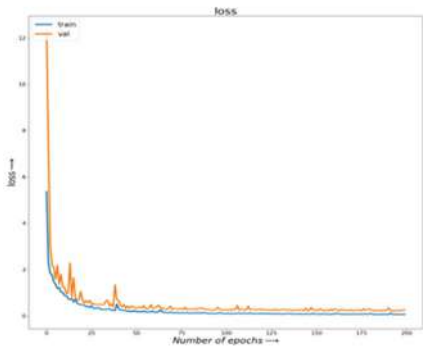
Fig. 6 (a) SR loss vs. No. of epochs (factor 2) (b) SR loss vs. No. of epochs (factor 4) (c) SR loss vs. No. of epochs (factor 6)



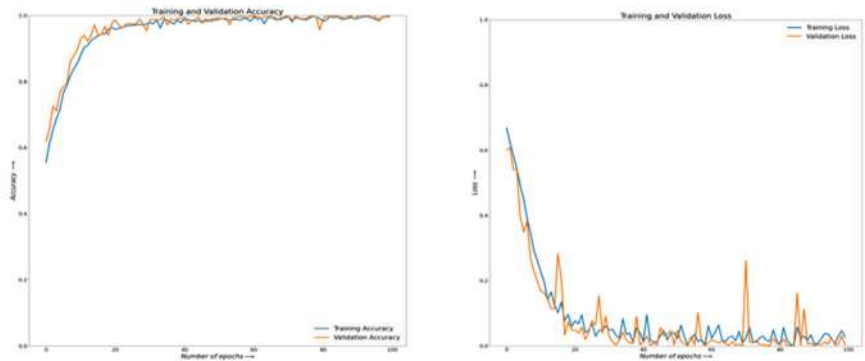
(a). SR loss vs No. of epochs (factor 2)



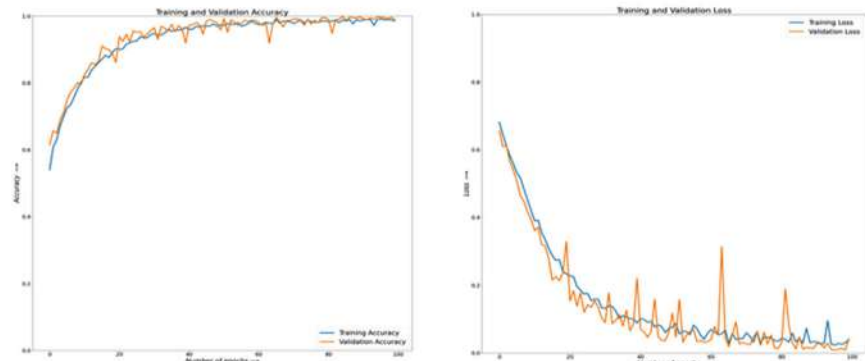
(b). SR loss vs No. of epochs (factor 4)



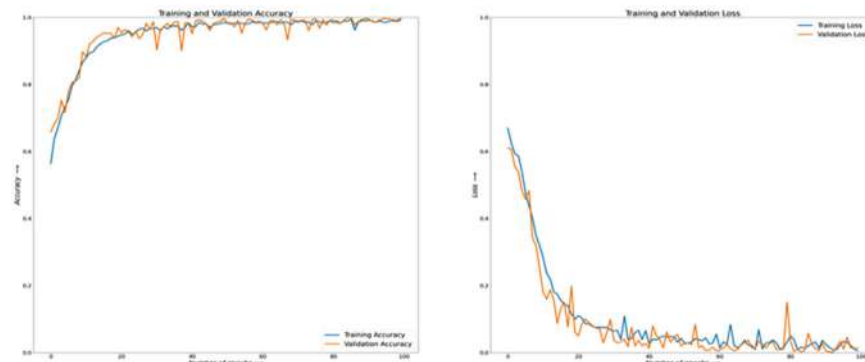
(c). SR loss vs No. of epochs (factor 6)



(a). Classification model Accuracies and losses of training vs. No. of epochs (factor 2)



(b). Classification model Accuracies and losses of training vs. No. of epochs (factor 4)



(c). Classification model Accuracies and losses of training vs. No. of epochs (factor 6)

Fig. 7 (a) Classification model accuracies and losses of training vs. No. of epochs (factor 2) (b) Classification model accuracies and losses of training vs. No. of epochs (factor 4) (c) Classification model accuracies and losses of training vs. No. of epochs (factor 6)

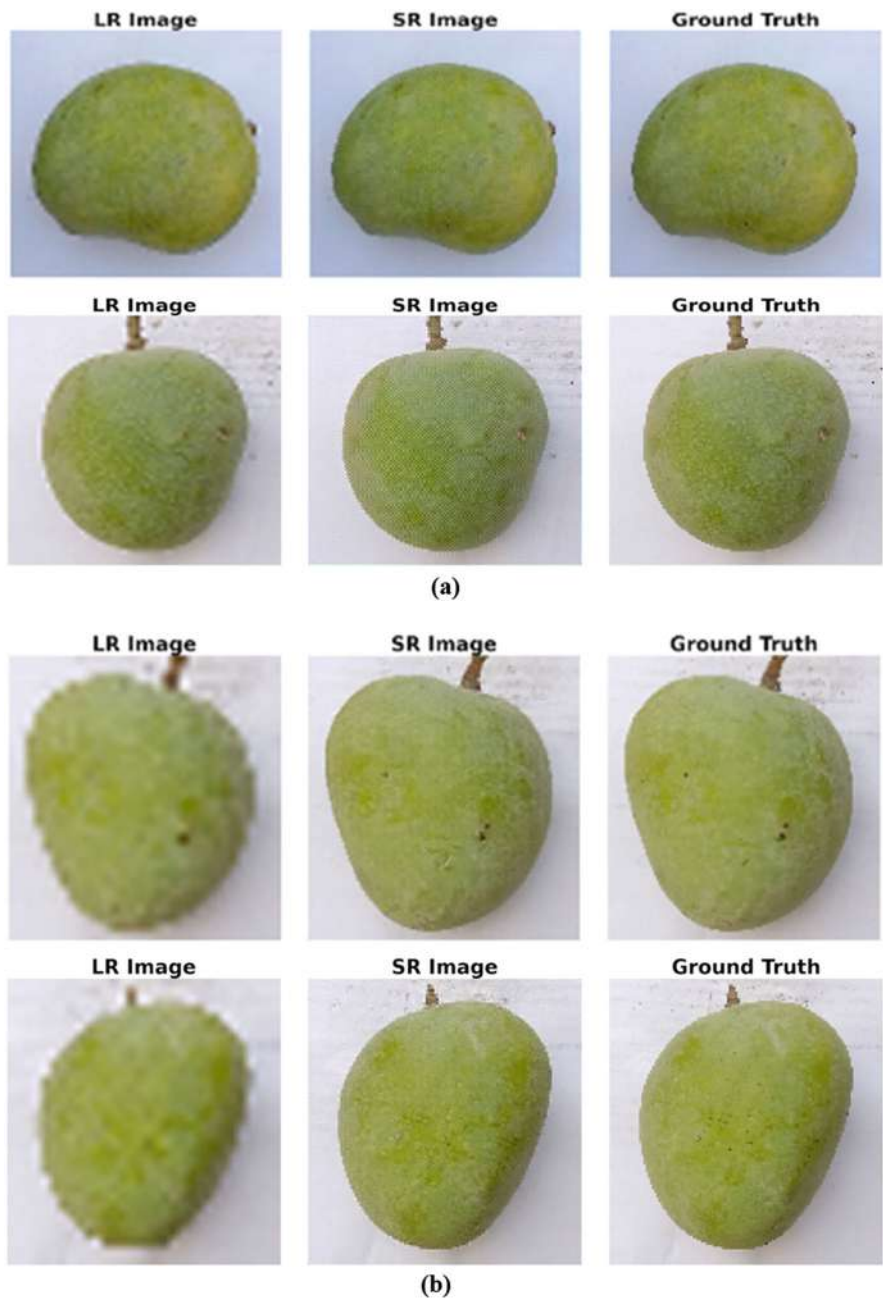


Fig. 8 (a) Super resolution with factor 2 (b) Super resolution with factor 4 (c) Super resolution with factor 6

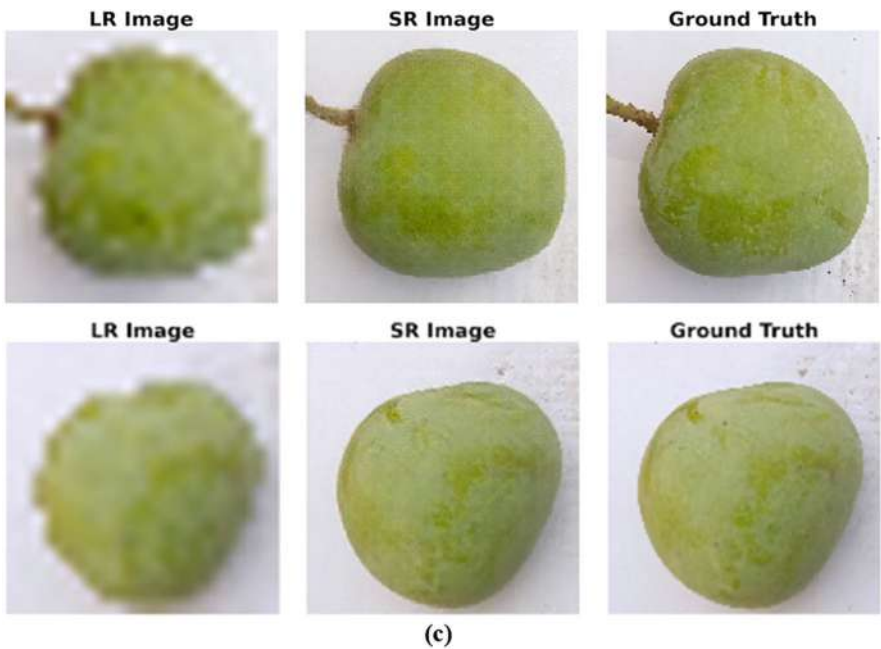


Fig. 8 (continued)

Table 2 Performance evaluation of PSRN

Metrics	SR image (factor 2)	SR image (factor 4)	SR image (factor 6)
PSNR	30.37	32.936	35.428
SSIM	0.8275	0.9146	0.9279

Table 3 Classification model performance evaluation

Metrics	SR image (factor 2)	SR image (factor 4)	SR image (factor 6)
Accuracy	99.45	98.86	97.64

5 Conclusion

This study proposes an accurate-time image super-resolution network based on Pilot Super Resolution (PSRN). Input images are super-resolved to obtain more exact categorization results. The model beat previous models such as ResNet50, GoogleNet, and EfficientNet. Following the retrieval of the shallow features, a super-resolution image is generated by first upsampling the features. Using super-resolution factors 2, 4, and 6, the proposed PSRN model achieves PSNR values of 30.37, 32.936, and 35.428, SSIM values of 0.8275, 0.9146, and 0.9279, and classification accuracy values of 99.45, 98.86, and 97.64.

Table 4 Comparison with existing models

Metrics	ResNet50	GoogleNet	EfficientNet	PSRN (Proposed)
Accuracy	95.01	96.45	97.73	99.76

Acknowledgments The authors are thankful to the director of the National Institute of Technology, Tiruchirappalli, for granting permission to use the GPU resources from the Center of Excellence—Artificial Intelligence (CoE-AI) lab.

References

1. Y. Li *et al.*, “NTIRE 2022 Challenge on Efficient Super-Resolution: Methods and Results.” pp. 1062–1102, 2022.

2. Yeswanth, P.V., Deivalakshmi, S., George, S. et al. Residual Skip Network-Based Super-Resolution for Leaf Disease Detection of Grape Plant. *Circuits Syst Signal Process* (2023). <https://doi.org/10.1007/s00034-023-02430-2>

3. K. W. Hung and W. C. Siu, “Single-image super-resolution using iterative Wiener filter based on nonlocal means,” *Signal Process Image Commun*, vol. 39, pp. 26–45, Nov. 2015, <https://doi.org/10.1016/J.IMAGE.2015.07.003>.

4. A. Muqeet, M. T. Bin Iqbal, and S. H. Bae, “HRAN: Hybrid residual attention network for single image super-resolution,” *IEEE Access*, vol. 7, pp. 137020–137029, 2019, <https://doi.org/10.1109/ACCESS.2019.2942346>.

5. B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, “Enhanced Deep Residual Networks for Single Image Super-Resolution.” pp. 136–144, 2017.

6. J. Jiang, Y. Yu, Z. Wang, S. Tang, R. Hu, and J. Ma, “Ensemble Super-Resolution with a Reference Dataset,” *IEEE Trans Cybern*, vol. 50, no. 11, pp. 4694–4708, Nov. 2020, <https://doi.org/10.1109/TCYB.2018.2890149>.

7. P. V. Arun, I. Herrmann, K. M. Budhiraju, and A. Karnieli, “Convolutional network architectures for super-resolution/sub-pixel mapping of drone-derived images,” *Pattern Recognit*, vol. 88, pp. 431–446, Apr. 2019, <https://doi.org/10.1016/J.PATCOG.2018.11.033>.

8. J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010, <https://doi.org/10.1109/TIP.2010.2050625>.

9. S. Vedula, S. Baker, T. K.-R. Techniques, and undefined 2002, “Spatio-temporal view interpolation,” *ri.cmu.edu*, 2001, Accessed: Apr. 03, 2023. [Online]. Available: http://www.ri.cmu.edu/pub_files/pub3/vedula_sundar_2001_3/vedula_sundar_2001_3.pdf

10. H. Chen, K. Sun, Z. Tian, C. Shen, Y. Huang, and Y. Yan, “BlendMask: Top-Down Meets Bottom-Up for Instance Segmentation.” pp. 8573–8581, 2020.

11. C. Dong, C. C. Loy, K. He, and X. Tang, “Image Super-Resolution Using Deep Convolutional Networks,” *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 2, pp. 295–307, Dec. 2014, <https://doi.org/10.1109/TPAMI.2015.2439281>.

12. A. Salmi, S. Benierbah, and M. Ghazi, “Low complexity image enhancement GAN-based algorithm for improving low-resolution image crop disease recognition and diagnosis,” *Multimed Tools Appl*, vol. 81, no. 6, pp. 8519–8538, Mar. 2022, <https://doi.org/10.1007/S11042-022-12256-W>.

13. D. Tiwari *et al.*, “Potato Leaf Diseases Detection Using Deep Learning,” *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2020, <https://doi.org/10.1109/ICICCS48265.2020.9121067>.

14. Y. Tai, J. Yang, and X. Liu, "Image Super-Resolution via Deep Recursive Residual Network." pp. 3147–3155, 2017.
15. Yeswanth, P. V., and S. Deivalakshmi. "Extended wavelet sparse convolutional neural network (EWSCNN) for super resolution." *Sādhana* 48, no. 2 (2023): 52.
16. Yeswanth, P. V., Khandelwal, R., & Deivalakshmi, S. (2022). Super Resolution-Based Leaf Disease Detection in Potato Plant Using Broad Deep Residual Network (BDRN). *SN Computer Science*, 4(2), 112.
17. P. V. Yeswanth, S. Kushal, G. Tyagi, M. T. Kumar, S. Deivalakshmi and S. P. Ramasubramanian, "Iterative Super Resolution Network (ISNR) for Potato Leaf Disease Detection," 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 2023, pp. 1–6, <https://doi.org/10.1109/IConSCEPT57958.2023.10170224>.
18. P. V. Yeswanth, R. Raviteja and S. Deivalakshmi, "Sovereign Critique Network (SCN) Based Super-Resolution for chest X-rays images," 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 2023, pp. 1–5, <https://doi.org/10.1109/IConSCEPT57958.2023.10170157>.
19. Yeswanth, P.V., Khandelwal, R., Deivalakshmi, S. (2023). Two Fold Extended Residual Network Based Super Resolution for Potato Plant Leaf Disease Detection. In: Misra, R., Rajarajan, M., Veeravalli, B., Kesswani, N., Patel, A. (eds) *Internet of Things (IoT): Key Digital Trends Shaping the Future. ICIoTCT 2022. Lecture Notes in Networks and Systems*, vol 616. Springer, Singapore. https://doi.org/10.1007/978-981-19-9719-8_16

Performance Evaluation of QoS in MAODV Routing Protocol in MANETS



Satish Dekka, Bosubabu Sambana , K. Narasimha Raju, D. Manendra Sai, M. Pallavi, and Kuppireddy Krishna Reddy

Abstract A Mobile Ad-hoc Network (MANET) is a wirelessly connected network of mobile devices that configures itself without the need for infrastructure. With the freedom to travel in any direction, every device within a MANET will regularly switch links with other devices. Each device operates as a router. The key difficulty in developing a MANET is ensuring that every device has the means to constantly update the data needed for accurate traffic routing. These networks can function alone or they can be linked to the wider Internet. Since MAODV is the reactive protocol linked to the Ad hoc On-Demand Distance Vector (AODV) routing protocol, it is comparable to AODV in terms of packet formats. The unicast Route Table and the packet types for the Route Request and Route Reply are derived from those used by AODV. In a similar vein, AODV defines a large number of the configuration settings that MAODV uses. A mobile node that wants to join or participate in a group inside an ad hoc network can do so dynamically, autonomously, and across several hops using the Modified Ad hoc On-Demand Distance Vector (MAODV) protocol. For managing routing in a mobile setting, there are various protocols. The MAODV (Modified Ad hoc On-Demand Distance Vector) routing protocol used

S. Dekka

Department of Computer Science and Engineering, Lendi Institute of Engineering and Technology (A), Vizianagaram, JNTU-GV, Vizianagaram, Andhra Pradesh, India

B. Sambana (✉)

Department of Computer Science and Engineering (Data Science), School of Computing, Mohan Babu University, Tirupathi, Andhra Pradesh, India

K. N. Raju · M. Pallavi

Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering (GVPCE), Andhra University, Visakhapatnam, Andhra Pradesh, India

D. M. Sai

Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, JNTU-GV, Vizianagaram, Andhra Pradesh, India

K. K. Reddy

Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering and Technology, Palamaner, JNTUA, Anantapur, Andhra Pradesh, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

133

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_10

in MANETs is therefore presented here. Packet delivery ratio, average end-to-end delay, and throughput are the measures utilized in simulation.

Keywords MANET · MAODV · AODV · NS2 simulator · Network transmission · Hybrid routing protocols

1 Introduction

A mobile ad hoc network (MANET) is a constantly self-configuring, infrastructure-less network of interconnected mobile devices. Every device in a MANET is free to move independently in any path and therefore transforms its links to other devices repeatedly. A MANET is an ad hoc network, which configures itself when mobile nodes alter their locations. It does not have any centralized support, and because of its dynamic nature, it faces some challenges. Among these challenges, routing is a prominent one. Each node communicates with all other nodes that are in its transmission range and with other nodes that are not in its transmission range via intermediate nodes. Communication is difficult because of the fact that the nodes recurrently change their locations. To solve this difficulty, a routing protocol is required. Routing protocols are categorized into three types. They are Proactive, Reactive, and Hybrid routing protocols.

Proactive: In proactive routing protocols, every node accumulates routing information in the form of tables and when any type of alteration accrues in network, topology needs to renew these tables accordingly. \rightarrow Reactive: In reactive routing protocols, route is determined when needed. A node initiates a path discovery throughout the network, only when it wants to transmit packets to its destination. \rightarrow Hybrid: Hybrid routing protocols merge the advantages of proactive and reactive routing protocols. The routing is mostly established with some proactively prospected routes and then provides the reactive routes required by additionally activated nodes.

MAODV implementation [6, 7] is the multicast extension of AODV [5]. Both AODV and MAODV are routing protocols for ad-hoc networks, with AODV for unicast traffic and MAODV for multicast traffic. NS2 [3], a widely used simulation tool, includes a standard implementation of the AODV protocol to analyze its performance [1], upon which we developed our initial MAODV implementations [2, 8]. But those MAODV implementations have two key limitations: (1) Only group members can send multicast traffic to the multicast group; (2) The multicast data packets are unicast, resulting in wasted bandwidth. This new version of MAODV allows each node in the network to send out multicast data packets, and the multicast data packets are broadcast when propagating along the multicast group tree.

Multicast Ad-Hoc On-Demand Distance Vector (MAODV) routing protocol is implemented. It is used to identify multicast routes on demand using a broadcast route-discovery mechanism. A source node originates a Route Request (RREQ) message when it desires to join a multicast group, or when it has data to launch to a multicast group but it does not have a route to that group. Only a member of

the desired multicast group may respond to a join RREQ. If the RREQ is not a join request, any node with a fresh enough route (based on group sequence number) to the multicast group may retort. If an intermediate node receives a join RREQ for a multicast group of which it is not a member, or if it receives a RREQ and it does not have a route to that group, it rebroadcasts the RREQ to its neighbors. As the RREQ is broadcast across the network, nodes set up pointers to establish the reverse route in their route tables.

2 Related Work

P. Mohapatra et al. [1] suggested in their paper “Multicasting in Ad Hoc Networks” that the widespread use of mobile and handheld devices is likely to popularize ad hoc networks, which do not require any wired infrastructure for intercommunication. The nodes of mobile ad hoc networks (MANETs) operate as end hosts as well as routers. They intercommunicate through single-hop and multi-hop paths in a peer-to-peer fashion. Most applications of MANETs require efficient support for multicast communications in which a node can communicate with multiple other nodes by exploiting the broadcast nature of wireless channels.

Elizabeth M. Royer et al. [2] represent their research work to discussion “Multicast AODV” may locate to connect multi-hop root paths, and maintain dynamic topology, provide loop freedom, and it used to quick route convergence using multi-channel allocation. Node connectivity with minimal memory and bandwidth usage, along with together with Unicast, broadcast, and multicast capabilities, node optimization used for improved overall routing expertise and reduced complexity analysis.

S.-J. Lee et al. [3] in their study “On-demand multicast routing protocol” discussed about ODMRP for multihop wireless mobile networks. ODMRP builds and maintains a mesh for each multicast group. Providing multiple paths by the formation of mesh configuration makes the protocol robust to mobility. ODMRP also applies demand-driven, as opposed to periodic, multicast route construction and takes a soft state approach in membership maintenance.

C. Chiang et al [4] presented “Adaptive shared tree multicast in mobile wireless networks.” Shared tree multicast is a well-established concept used in several multicast protocols for wireline networks (e.g., core base tree and PIM sparse mode). They extend the shared tree concept to wireless, mobile, multihop networks for applications ranging from ad hoc networking to disaster recovery and battlefield. They addressed various issues in design.

X. Hong et al. [5] surveyed “Scalable Routing Protocols for Mobile Ad Hoc Networks.” The growing interest in Mobile Ad Hoc Network techniques has resulted in many routing protocol proposals. This work mostly focus on the operational characteristics and scalability of routing protocols, and how they will connect together, if any difficulties in occurred channel allocation and deallocation and path designing, and new routing protocols in the future, and make clear how the

underlying network topology affects with the routing techniques. Additionally, they demonstrate how the routing approach varies with respect to different design factors like FSR, FSLs, OLSR, and TBRPF.

Thomas Kunz et al. [6] said “Multicasting in Ad-Hoc Networks: Comparing MAODV and ODMRP”. They connected to searching network pathway and resource channel allocation could be more efficient with the simulated experiments using the ODMRP outperforms and gives very poor PDR values, and it is the most important parameter to identifies data sharing conformation from source to destination through channel allocation, and in between communication internal node data transformation (speed = 1 m/s) through the MANET and MAODV protocols,. When the node speed is 10 m/sec, the directed movement achieves 250% to 500% and more, better values based on target node data transformation.

Cheng E et al. [7] said that represent their work on “On-Demand Multicast Routing protocol in Mobile Ad-hoc Networks” uses a flooding-based mesh and connect a forwarding group node channelized. A soft state node sends information to multicast node group members without explicit, and group membership frequent send messages. Multicast routes and group membership are established and updated on demand. When a multicast source has data packets to send to a group without any knowledge of the route, it broadcasts a Join-Query control message The Join-Query message is also periodically broadcast to update the membership and route information.

Kunz T et al. [8], “Reliable multicasting in MANETs,” provided an in-depth study of one-to-many and many-to-many communication in mobile ad-hoc networks. The original goal of this work was to design an efficient protocol that delivers packets from one or multiple senders to many receptions with high probability.to exploring the performance of a number of best-effort protocols and efficient broadcast protocols this mechanism indeed increases the packet delivery ratio.

Royer E. M. and Perkins et al. [9] discussed “Multicast Operation of the Ad-hoc On-Demand Distance Vector Routing Protocol” on Ad hoc On Demand Distance Vector Routing AODV an algorithm for the operation of such adhoc networks to over novel multicast capabilities which follow naturally from the way AODV establishes unicast routes AODV builds multicast trees as needs.

Royer E. M et al. [10] represent their work on “Multicast Ad-hoc On-Demand Distance Vector (MAODV) Routing” protocols and how the AODV and MAODV protocols are intended for used by mobile nodes in an ad-hoc networks. It offers quick adaptation to dynamic channel link conditions, low processing and memory overhead, and low network utilization, etc. and it creates bi-directional shared multicast trees and connecting multicast resources and receivers.

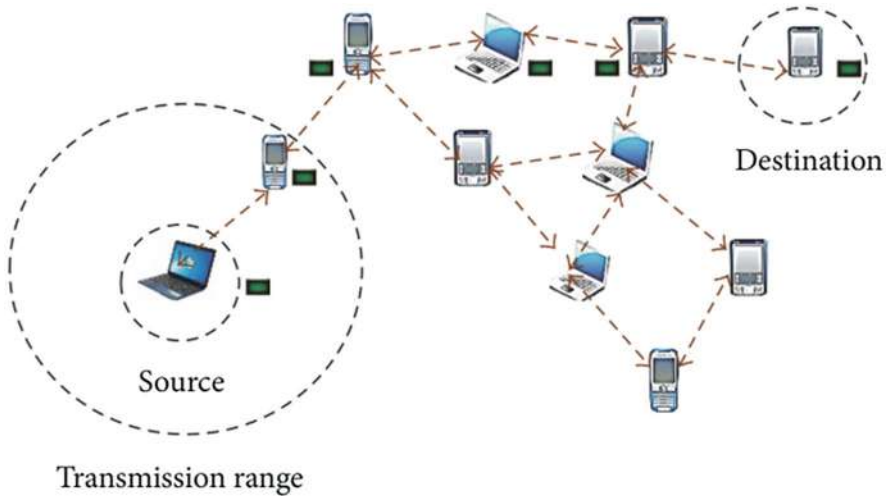


Fig. 1 Mobile ad-hoc network

3 Methodology

A mobile ad hoc network (MANET) is a constantly self-configuring, infrastructure-less network of interconnected mobile devices. Every device in a MANET is free to move independently in any path and therefore transforms its links to other devices repeatedly.

The flow of the Fig. 1 shown diagram represents the transmission of the data from the source to the destination through media and calculate the best results and high throughput. The mainly focused on this work “Performance analysis of MAODV” is that we can calculate different types of performance metrics such as Average Throughput, End-to-End Delay, and Packet delivery calculating and estimate performance analysis of MAODV.

- Average Throughput
- End-to-End Delay
- Packet Delivery Ratio

Average Throughput

A benchmark can be used to measure throughput. In data transmission, network throughput is the amount of data moved successfully from one place to another place in a given time period and is typically measured in bits per second as in megabits per second gigabits per second.

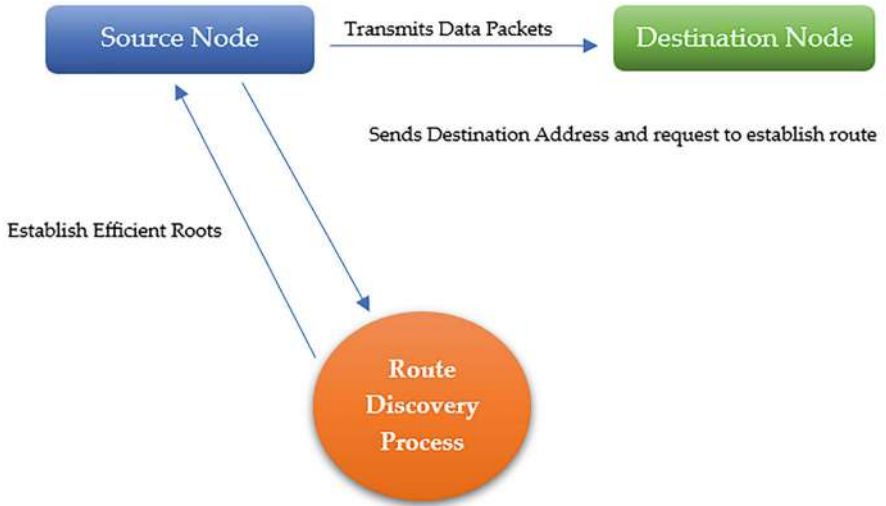


Fig. 2 Data flow diagram level-0

End-to-End Delay

The End-To-End delay calculates the duration of a packet. The notation for the End-To-End delay is: End-to-End Delay = End Time—Start Time.

Packet Delivery Ratio

Packet delivery ratio is the ratio of packets that are successfully delivered to a destination compared to the number of packets that have been sent by sender. In order to calculate packet delivery ratio, we need total number of packets sent and number of received packets.

Packet delivery ratio = Number of received packets/Number of sent packets* 100.

A data flow diagram (DFD) is a graphical representation of the “flow” as shown in Figs. 2, 3 and 4 of data through an information system, modeling its process aspects. A DFD is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. DFDs can also be used for the visualization of data processing (structured design).

3.1 System Architecture

Figure 5 shows the conceptual model that defines the structure, behavior, and more views of a system. An architecture is a formal representation of a system, and organized in a way that supports reasoning about the internal node structure connectivity and node path behaviors of the system. This system architecture can

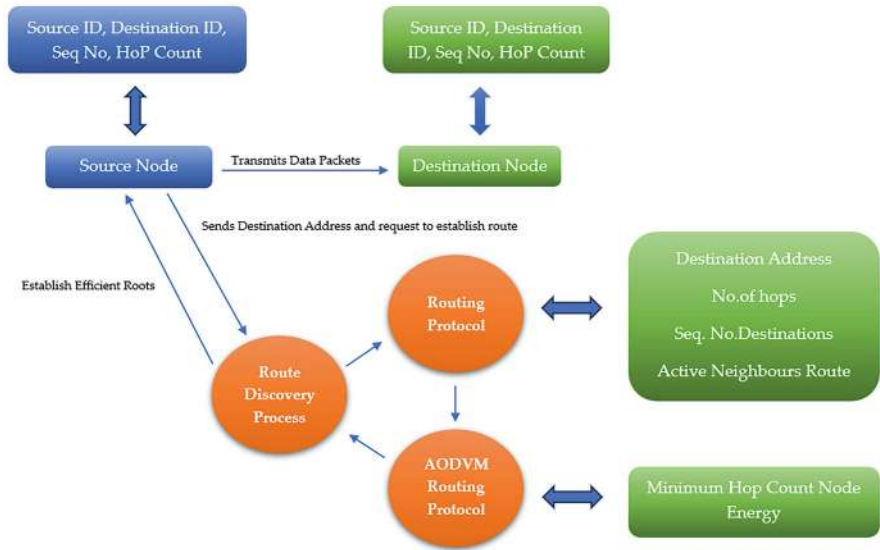


Fig. 3 Data flow diagram level-1

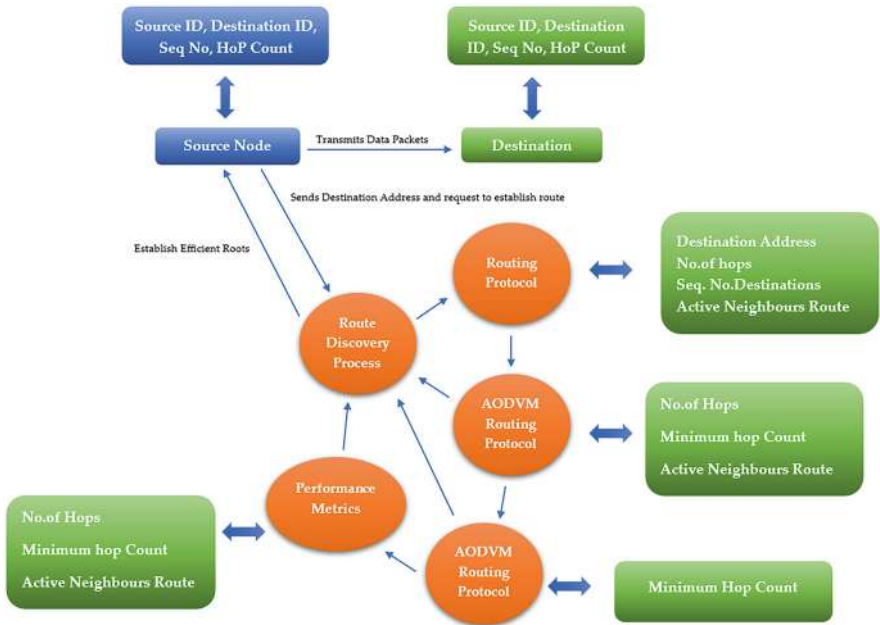


Fig. 4 Data flow diagram level-2

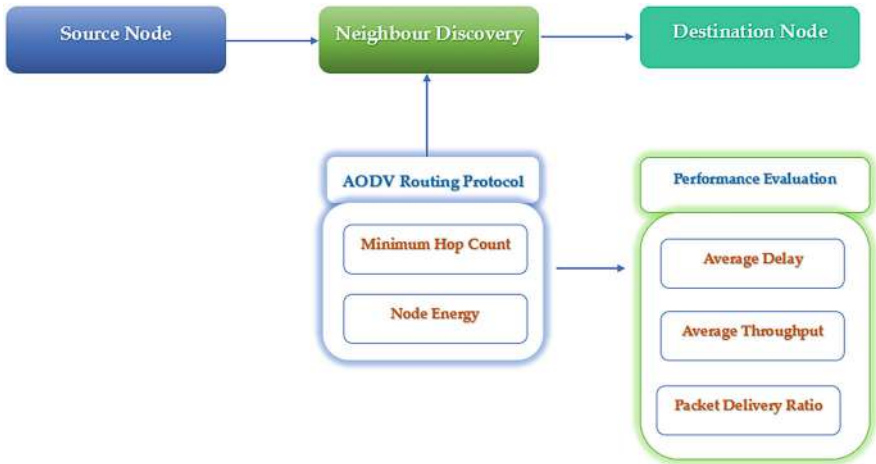


Fig. 5 System architecture

include system components and run through operational functionalities, through the proposed system.

3.2 Routing Protocols

Routing is the process of information Fig. 6 shown exchange node transaction from one host to another host in a network. Routing is the mechanism of forwarding packet toward its destination using most efficient path. Efficiency of the path is measured in various metrics like number of hops, traffic, security, etc.

4 Implementation

The implementation phase involves putting the work plan into action. It is the phase where visions and plans become reality. This is the logical conclusion, after evaluating, deciding, visioning, planning, applying for funds, and finding the financial resources of a work. Technical implementation is one part of executing a work.

The implementation phase consists of four sub-phases: Execution, Monitoring & Control, and Move to Production.

Monitoring and Controlling Progress of a Work’s Implementation When trying to send a message to a destination node without knowing an active route to it, the sending node will initiate a path discovery process. A route request message

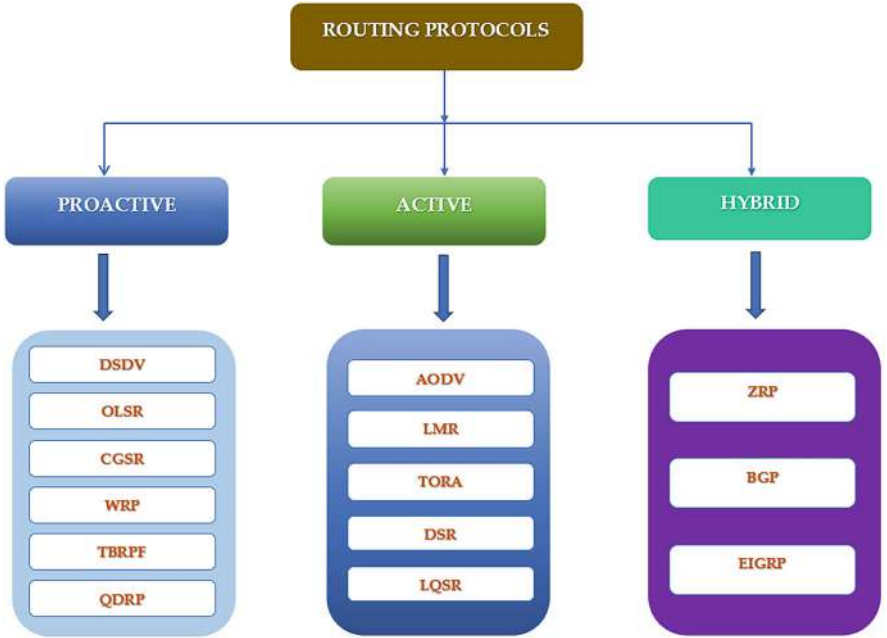


Fig. 6 Classification of routing protocols in mobile Ad-hoc networks

(RREQ) is broadcasted to all neighbors, who continue to broadcast the message to many connected nodes are communicated within the multicast channelized pathways.

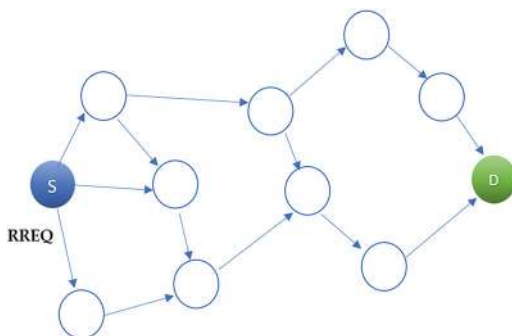
System Architecture

A system architecture Fig. 5 shown in the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. A system architecture shows how components are expand for systems developed, that will work together and to implement the overall system and identify the performance

Routing Protocol

Classification of Routing Protocols in Mobile Ad-hoc Networks and the process information Fig. 6 shown how to exchange from one host to the other host in a

Fig. 7 Source node initiates the path discovery process



network. Routing is the mechanism of forwarding packet towards its destination using most efficient path. Efficiency of the path is measured in various metrics like, Number of hops, traffic, security, etc.

The forwarding process is continued until the destination node is reached or until an intermediate node knows a route to the destination that is new enough. To ensure loop-free and most recent route information, every node maintains two counters: sequence number and broadcast_id. The broadcast_id and the address of the source node uniquely identify a RREQ message. Broadcast_id is incremented for every RREQ the source node initiates. An intermediate node can receive multiple copies of the same route request broadcast from various neighbors.

In this case if a node has already received a RREQ with the same source address and broadcast_id, it will discard the packet without broadcasting it further. When an intermediate node forwards the RREQ message, it records the address of the neighbor from which it received the first copy of the broadcast packet. This way, the reverse path from all nodes back to the source is being built automatically. Figure 7 shows RREQ packet that contains two sequence numbers: the source sequence number and the last destination sequence number known to the source. The source sequence number is used to maintain “freshness” information about the reverse route to the source while the destination sequence number specifies what actuality a route to the destination must have before it is accepted by the source.

When the route request broadcast reaches the destination or an intermediate node with a fresh enough route, the node responds by sending a unicast route reply packet (RREP) back to the node from which it received the RREQ as shown in Fig. 8. So actually the packet is sent back to reverse the path built during broadcast forwarding. A route is considered fresh enough, if the intermediate node’s route to the destination node has a destination sequence number that is equal or greater than the one contained in the RREQ packet.

As the RREP is sent back to the source, every intermediate node along this path adds a forward route entry to its routing table. The forward route is set active for some time indicated by a route timer entry. The default value is 3000 milliseconds, as referred to in the AODV RFC. If the route is no longer used, it will be deleted after the specified amount of time. Since the RREP packet is always sent back the reverse path established by the routing request, AODV only supports symmetric links.

Fig. 8 A RREP packet is sent back to the source

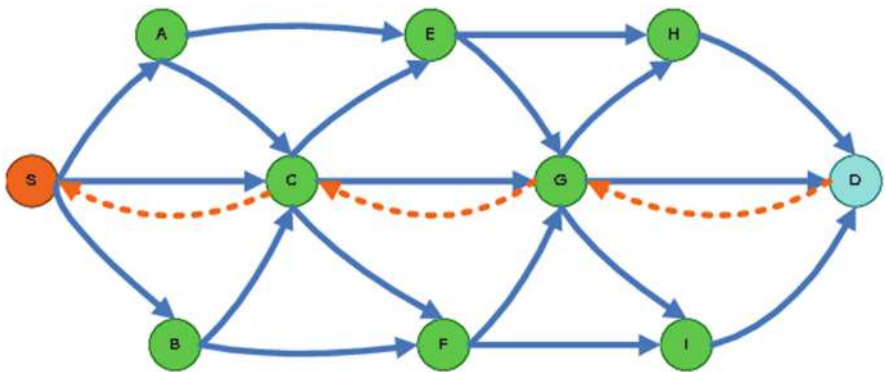
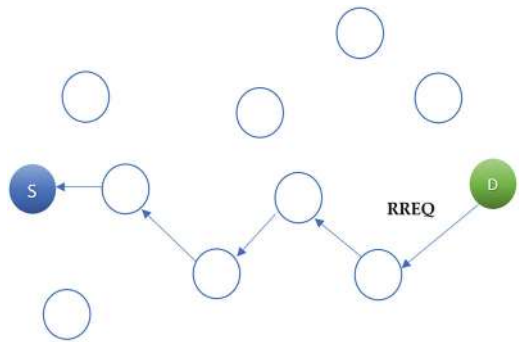


Fig. 9 Maintaining routes

4.1 Maintaining Routes

Mechanism of AODV routing protocol route maintenance is done by Means of Route Error (RERR) packets. When an intermediate node detects a link breakage, it generates a RERR packet. The RERR propagates toward all traffic sources having a route via the failed link and erases all broken routes on the way. A source upon receiving the RERR initiates a new route discovery if it still needs the route as shown in Fig. 9.

During multicast data packet forwarding, each node first checks if it is in the multicast tree. If it is not a tree member, it will check its Unicast Route Table to find the next hop for the multicast address. If it has the information, the data packets are forwarded toward the next hop; otherwise, it will send an unsolicited RREP back to the source node, in order to let the source node, initiate a new route discovery if it still needs a route to that multicast address. If the node itself is a tree member, it will follow its Multicast Route Table to forward the packets.

4.2 Algorithm Specification AODV

- Source initializes for communication
- Searches for neighbor nodes
- Then sends RREQ to neighbors until destination node receives the RREQ
- Route which has minimum hop count is selected

4.3 AODV Modified

The proposed protocol AODVM Figs. 10 and 11 as shown route discovery process is similar to the AODV protocol. The difference is AODVM and AODV protocols are determines to identifies internal node objective connectivity and an optimum route by considering the combination of residual energy of nodes on the path and hop count.

Each multicast group has a unique multicast group address. According to the MAODV specification [6, 7], each multicast group is organized by using a tree structure, composed of the group members and several routers, which are not group members but must exist in the tree to connect the group members [11].

Group-Hello (GRPH) messages in the whole network. The group leader also maintains the group sequence number, which is propagated in the network through the GRPH.

RREQ is the same as the RREQ used in AODV, without any flags and broadcasted in the network. But with the Group Leader Table, the source node may already know a route to reach the group leader. By using the information recorded in the Group Leader Table, the RREQ can be sent unicastly toward the group leader if this is the first time the node sends RREQ. During RREQ propagation, the reverse route toward the source node is constructed as described in the AODV protocol.

NS2 is a discrete event simulator used to simulate real-time network traffic and topology for analysis. For researchers, it is proven to be great tool. NS2 has been developed in C++ and TCL. Otcl, which is object-oriented TCL, is also used in NS2. For simulation purposes, TCL programming is used and for adding new module C++ is used.

4.4 Implementation of MAODV

Simulation Environment: Simulations will be done with NS 2.31. The simulation parameters used in NS 2.27 during the ad-hoc network simulation are configured as follows:

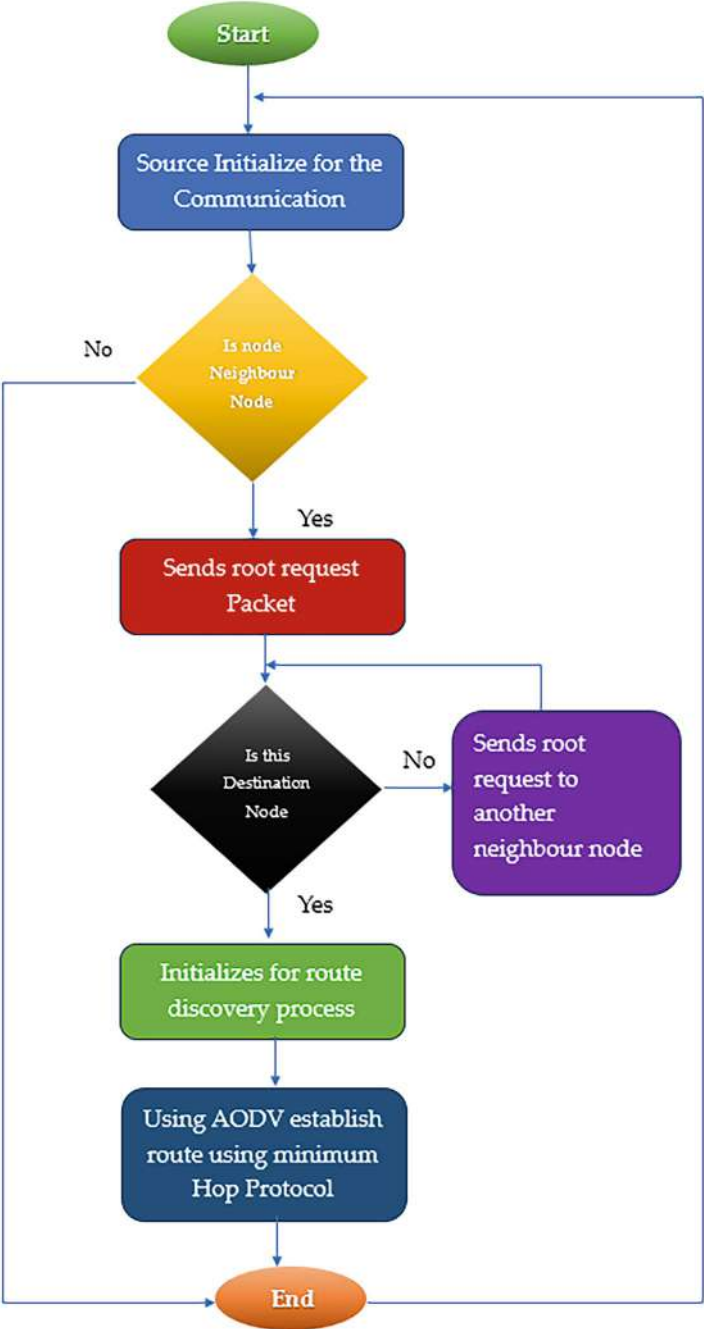


Fig. 10 Flowchart for Ad-hoc on-demand distance vector routing protocol

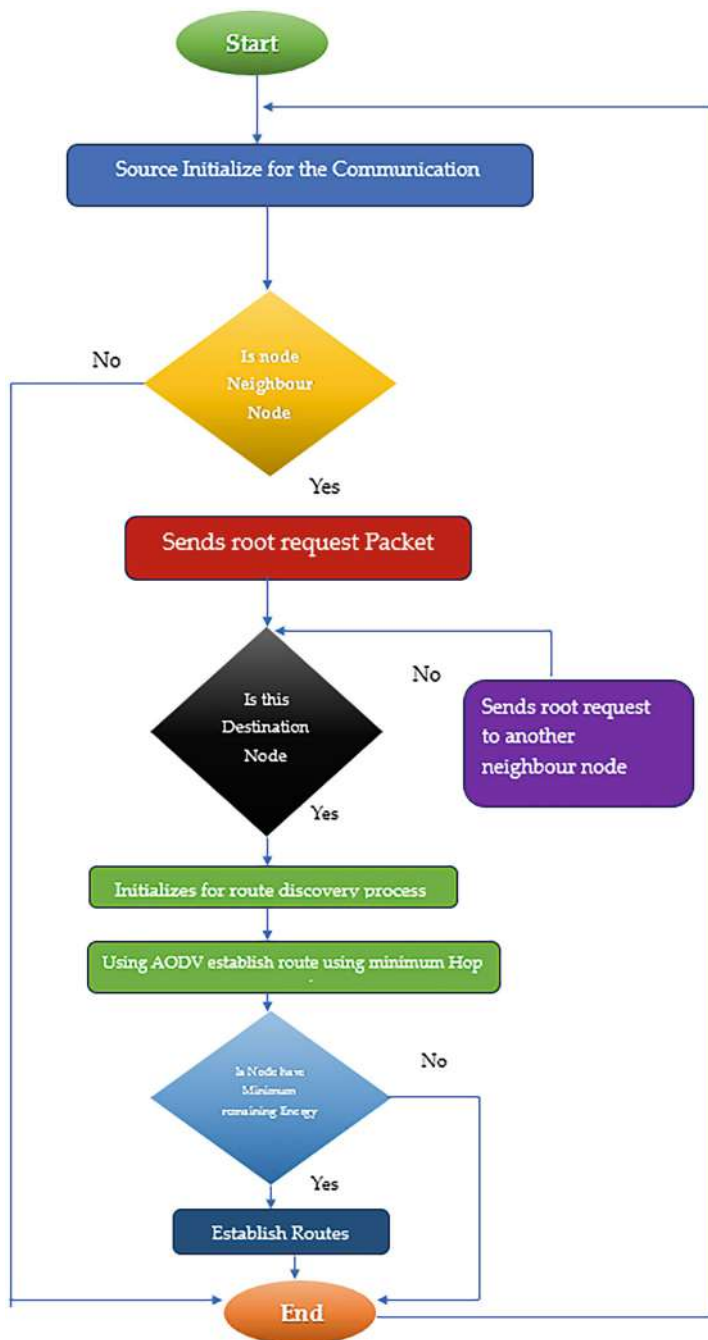


Fig. 11 Flowchart for Ad-hoc on-demand distance vector modified routing protocol

Parameter	Value
Channel type	Wireless channel
Radio propagation model	Two ray ground
Mac layer	CSMA/CA as in IEEE 802.11
Data rate at physical layer	11 Mbps
Queue type	Drop tail
Maximum queue length	50

5 Results

The complete result and comparative analysis are shown in Figs. 12, 13, 14, 15, 16 and 17.

6 Conclusion

The proposed MAODV routing protocol produces more number of multiple paths between source and destination node in time considering that they can be used if required, usually when the route link in use is broken and by running route discovery process in parallel with packet transfer phase and assure to have a number of adjacent route paths in advance. The proposed protocol has been tested and validated with the implementation under different scenarios. And also simulation-

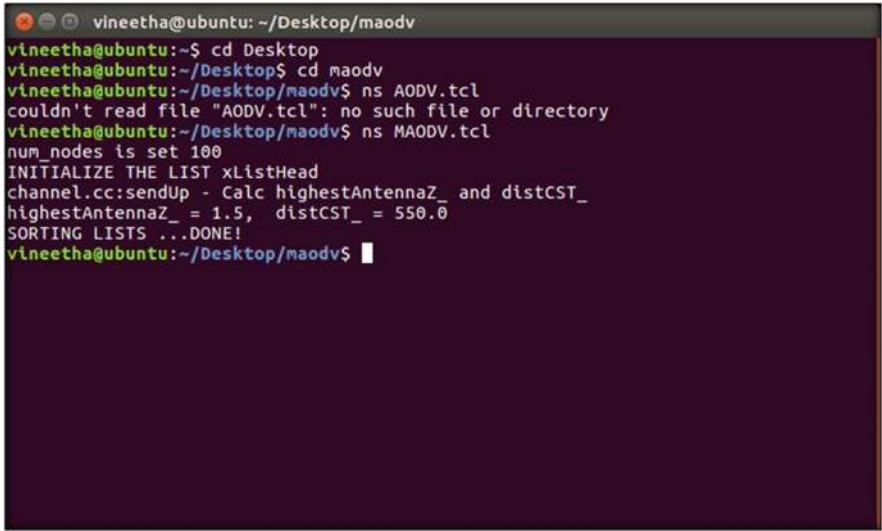
A screenshot of a terminal window with a dark background and light-colored text. The window title is 'vineetha@ubuntu: ~/Desktop/maodv'. The user 'vineetha' is at the 'ubuntu' machine in the directory '~/Desktop/maodv'. The terminal shows the following commands and output:
vineetha@ubuntu:~\$ cd Desktop
vineetha@ubuntu:~/Desktop\$ cd maodv
vineetha@ubuntu:~/Desktop/maodv\$ ns AODV.tcl
couldn't read file "AODV.tcl": no such file or directory
vineetha@ubuntu:~/Desktop/maodv\$ ns MAODV.tcl
num_nodes is set 100
INITIALIZE THE LIST xListHead
channel.cc:sendUp - Calc highestAntennaZ_ and distCST_
highestAntennaZ_ = 1.5, distCST_ = 550.0
SORTING LISTS ...DONE!
vineetha@ubuntu:~/Desktop/maodv\$

Fig. 12 Executing scripts

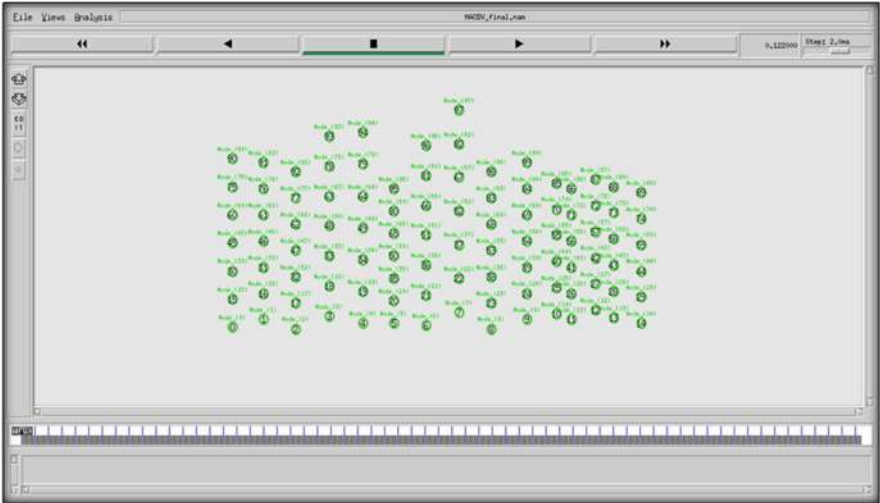


Fig. 13 Sample network creation

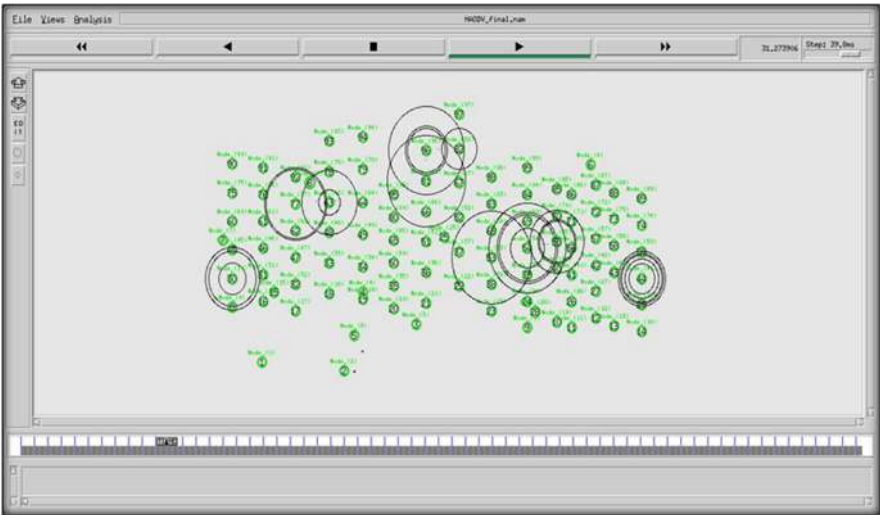


Fig. 14 Simulation using MAODV protocol in NAM

based experiments are performed to analyze the performance of MAODV routing protocol by evaluating packet delivery ratio, throughput, and packet loss. These results are compared with AODV, DSDV and DSR routing protocols by various number of nodes and mobility. The comparison shows that the proposed protocol MAODV for adhoc networks performs better as compared to the other protocols. Thus adding MAODV routing protocol is meaningful to optimize the performance



Fig. 15 Energy utilization of MAODV using X-Graph utility

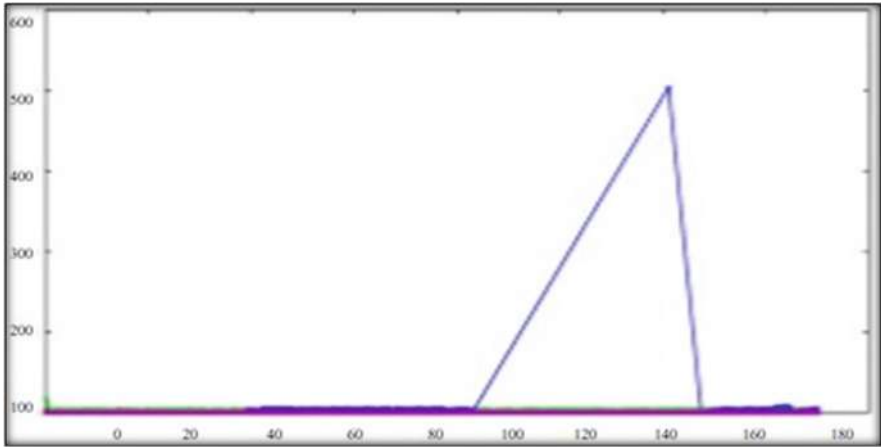


Fig. 16 ETE delay comparison for 5 nodes

of the MANET especially during the real-time traffic. In this case, the multicast tree may not be the most efficient. So a mechanism based on Group Hello messages can be improved in future. Currently, MAODV does not specify any special security measures. Route protocols are prime targets for impersonation attacks and must be protected by the use of authentication techniques involving generation of unforgeable and cryptographically strong message digests or digital signatures.

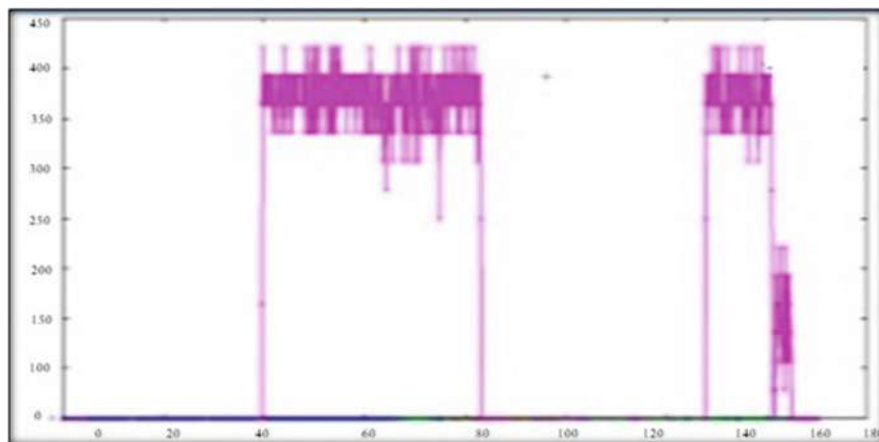


Fig. 17 Throughput comparison for 5 nodes

References

1. P. Mohapatra, J. Li, C. Gui "Multicasting in Ad Hoc Networks", book chapter, Kluwer Press, 2004.
2. Elizabeth M. Royer, Charles E. Perkins, "Multicast Operation of the Ad-hoc On-demand Distance Vector Routing Protocol", Proceedings of MobiCom '99, Seattle, WA, August 1999, pp. 207–218.
3. S.-J. Lee, M. Gerla, and C.-C. Chiang, "On-demand multicast routing protocol" Proceedings of IEEE WCNC, 1999, pp. 1298–1302.
4. C. Chiang, M. Gerla, and L. Zhang "Adaptive shared tree multicast in mobile wireless networks" Proceedings of IEEE Globecom '98, pp. 1817–1822.
5. X. Hong, K. Xu, and M. Gerla, "Scalable Routing Protocols for Mobile Ad Hoc Networks" IEEE Network, July-Aug, 2002, pp. 11–21.
6. Thomas Kunz and Ed Cheng "Multicasting in Ad-Hoc Networks: Comparing MAODV and ODMRP", Proceedings of the Workshop on Ad hoc Communications, Bonn, Germany, September 2001.
7. Cheng. E.; "On-Demand Multicast Routing in Mobile Ad Hoc Networks", M. Eng. Thesis, School of Computer Science, Carleton University, January 2001.
8. Kunz, T.; "Reliable multicasting in MANETs", Contractor Report, Communications Research Centre, Ottawa, Canada, July 2003.
9. Royer, E. M. and Perkins, C. E.; "Multicast Operation of the Ad-hoc On-Demand Distance Vector Routing Protocol", Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'99), Seattle, WA, USA, August 1999, pages 207–218.
10. Royer, E. M. and Perkins, C. E.; "Multicast Ad hoc On-Demand Distance Vector (MAODV) Routing", IETF, Internet Draft: draft-ietf-manet-maodv-00.txt, 2000.
11. Mohammed Rehman shareef et al. Data Migration Using Oracle In Government Agency Applications, NeuroQuantology, August 2022, Volume 20, Issue 8, 10613–10619. <https://doi.org/10.48047/nq.2022.20.8.nq221084>.

Minimize the Energy Consumption to Increase the Network Lifetime for Green IoT Environment



Satish Dekka, K. Narasimha Raju, D. Manendra Sai, M. Pallavi,
and Bosubabu Sambana 

Abstract One of the most important and complex issues in the IoT context is energy usage. Because the many nodes utilized in IoT contexts are energy-constrained and must operate for long periods of time without human interaction, their batteries quickly run out of power. As a result, nodes must manage their energy effectively. Sensing, communicating, and processing are regarded as the primary functions of IoT sensors. IoT connectivity uses more energy at the node level than at other times. The main motto of this study is to reduce energy consumption and lengthen network life. Although routing also has a major impact on energy usage, the first one deals with employing an energy model, and the second one deals with selecting an effective routing protocol. When picking a protocol, various other performance indicators such as throughput, delay, packet delivery ratio, and packet loss are also considered. The Network Simulator 3 is used as a tool for testing these strategies (NS3).

Keywords MANET · Green IoT · NS3 · Routing · Energy Model · Green networks

S. Dekka

Department of Computer Science and Engineering, Lendi Institute of Engineering and Technology (A), JNTU-GV, Vizianagaram, Andhra Pradesh, India

K. N. Raju · M. Pallavi

Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering (GVPCE), Andhra University, Visakhapatnam, Andhra Pradesh, India

D. M. Sai

Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, JNTU-GV, Vizianagaram, Andhra Pradesh, India

B. Sambana (✉)

Department of Data Science, School of Computing, Mohan Babu University, Tirupathi, Andhra Pradesh, India

1 Introduction

The Internet of Things (IoT) is the term used to describe the billions of physical objects that are connected to the Internet and used to exchange data globally. IoT has significant uses in every area of our life. IoT often consists of a heterogeneous environment consisting of MANETs and sensor networks with a variety of energy-constrained devices, such as sensors, RFID tags, mobile devices, etc. These devices' primary functions involve gathering, processing, and transferring data across nodes. These nodes must operate continuously over extended periods of time without any human assistance. Green IoT refers to IoT that conserves energy. It illustrates the problem of lowering IoT device energy usage to create a sustainable environment for IoT systems. Green networks in the IoT will aid in lowering emissions and pollution, utilizing environmental monitoring and conservation, and cutting back on operating expenses and power use.

In a Mobile Ad-hoc Network (MANET), each node has the freedom to move around the area and communicate with one another through wireless links without the need for a base station or centralized controller. These features make MANET realistic and effective in a variety of contexts, including military scenarios, sensor networks, rescue operations, on-campus students, etc. Power consumption is one of the most important design considerations in Mobile Ad-hoc networks as the nodes in MANET have battery limitations.

Sensing, communicating, and computing are regarded as the primary functions of IoT nodes and sensors. Out of which it is thought that communication uses up a lot of energy. There are various connections that link the energy issue with data in the IoT context. Since there are many different types of data processing, including data gathering, processing, and transmission, the data is one of the main factors that contribute to the energy consumption of the nodes on the Internet of Things.

These node batteries would quickly run out of power without effective energy management techniques, which would interfere with the operation of the application. It is quite difficult to relay data with these weakly powered sensors. As a result, managing energy consumption is one of the IoT's most important and complex concerns, and extending network lifetime is the goal.

2 Related Work

Omar Said et al. [1] proposed a system using three strategies for Energy management in IoT environments. The first strategy deals with data reduction techniques like data prioritization, fitting, and compression. The second strategy deals with a priority-based scheduling algorithm for processing the data at each node. The third strategy includes a fault tolerance system which is used as a backup or a complement to the about two strategies. Rakesh Kumar Lenka et al. [2] suggested using an

energy-efficient routing protocol for data transferring which will help reduce energy consumption and boost the network lifetime of WSN (Wireless Sensor Network).

Mohamed Er-Rouidi et.al [3] suggested an Energy Consumption Evaluation of Reactive and Proactive Routing Protocols in Mobile Ad-Hoc Network in an understanding and consumption shown in visualization. Amjad Anvari-Moghaddam et.al [4] proposed the Optimal Smart Home Energy Management Considering Energy Saving and a Comfortable Lifestyle towards humans. Dr. Shridhar A Joshi et.al [5] Examined IoT Based Smart Energy Meter wirelessly connected to users by the means of IoT and their energy efficiency techniques. Vinita Tahiliani et.al [6] presented the current state of the art research on energy optimization in IoT.

Saurabh Singh et al. [7] presented various technologies and issues regarding green IoT. They proposed a green IoT-Home Service (GIHS) model that provides efficient energy management in the home automation system. V. Preethi and G. Harish [8] presented a smart energy meter for automatic metering and the corresponding amount will be displayed on the LCD continuously. Communication between the user/household and substation is done using Zigbee. GSM network is used for sending SMS to the local authorities regarding theft cases. This meter can work as either a prepaid or postpaid meter.

Tri Kuntoro Priyambodo et al. [9] examined the effect of the route request parameters, on the Ad Hoc On-demand Distance Vector (AODV) protocol. The performance metrics used for measuring performance were Packet Delivery Ratio (PDR), throughput, delay, packet loss, energy consumption, and routing overhead. By reducing the combination value of RREQ_RETRIES, MAX_RREQ_TIMEOUT in AODV routing to (2, 10 s) and (3, 5 s), the protocol's performance can be improved. Ejaz et al. [10] proposed an optimized and schedule energy-efficient framework for smart cities in addition to energy harvesting which extended the lifetime of low-power devices. The performance analysis of this framework is weak because it depends on four appliances. It neglected the large quantity of data that may be transmitted in the smart city or through the IoT environment. Danielly B. [11] et al. provided an integrated and single system that allows monitoring of the energy consumption of the final consumers while providing useful information about the energy quality [12].

3 Methodology

The proposed energy management system uses two strategies to deal with the energy issues in the IoT network and increase the network lifetime, which is illustrated below. The strategies can be applied to any of the applications of IoT battery-based nodes like Wireless Sensor Networks (WSN), Mobile Ad-hoc Networks (MANET), and Radio Frequency Identification (RFID).

Strategy 1: Energy Model

An energy model is used to track the energy consumption rate at each node and to calculate the remaining energy at the node. It notifies the source in case of energy depletion and signals the harvester to recharge the source.

Strategy 2: Routing Protocol

The second strategy deals with choosing the best routing protocol and scheduling the events in the order of the oldest one executes first. Various performance metrics like throughput, packet loss, delay, jitter, packet delivery ratio, and average energy consumption are considered.

Algorithm

- Step-1: Create the Nodes required using the Node container.
- Step-2: Attach the channel and Net Device and add MAC.
- Step-3: Install the protocol stack onto the devices.
- Step-4: Assign an IP address to Net Devices.
- Step-5: Configure the Energy model.
- Step-6: Run the Simulation by setting the simulation time.
- Step-7: Calculate the performance metrics.
- Step-8: Create objects for trace metric, flow monitor, ascii tracer, and net aim to perform further analysis on output.
- Step-9: Destroy the Simulator object.

4 Experimental Setup

MANET (Mobile Ad-hoc Network) as a part of IoT considered for our simulation. The simulated environment of the MANET network is created using Network Simulator 3 [Table 1]. An energy model is implemented which will monitor the initial and remaining energy levels of each node in the network. Next, we are comparing all the manet routing protocols using an energy framework and filter them by average energy consumption and also comparing other performance metrics like throughput, delay, jitter, packet loss, PDR the best one is selected and implementing a queue-based scheduling algorithm to schedule these events.

4.1 Performance Metrics

The following performance metrics are considered for the evaluation of the environment

Average Throughput: Average throughput is the amount of data communicated from the source node to the destination node through the network in a unit of time stated in Kbps (Kilobits per second).

Table 1 Simulation parameters

Network parameters	Value assigned
Node density	20, 30, 40, 50
Simulation time	120 s
Pause or halt time	0 s, 10 s
Node velocity	20 m/s
Network region	300 X 1500
Source/sink connections	10
Transmit power set to	7.5 dBm
Mobility model used	Random Waypoint
Size of data packet	64 Bytes
Wi-fi rate	2 Mbps
Wifi operation mode	Ad-hoc
Data rate	2048 bytes per second
Propagation model	Constant speed propagation delay
Mac	Adhoc Wifi MAC
Routing protocols used	AODV, OLSR, DSDV, DSR

Packet Delivery Ratio (PDR): PDR is the ratio of total received packets to the total sent packets.

$$PDR = (\text{Received Packets Total}) / (\text{Sent Packets Total}) \times 100\%$$

Average End-to-End Delay (EED): EED is the average time interval between data packets produced at the source node and the effective delivery of these packets at the destination node.

$$EED = \text{Delay Sum Total} / \text{Packets Received}$$

Lesser values of end-to-end delay provide better and improved performance. It is derived in ms (mile seconds).

Packet Loss (PL): PL is the difference between the total packets sent and the total packets received.

$$PLR = ((\text{Total Packets Sent}) - (\text{Total Packets Received})) / \text{total packets sent}$$

Lesser values of packet loss ratio provide better and enhanced performance. PL is derived as the number of packets.

Average Energy Consumption: It is calculated for each node by subtracting the initial value (i) of energy at each node from the remaining energy (r) and then divided by total nodes (t).

$$AEC = (i - r) / N$$

Lesser values of AEC provide better and enhanced performance. Measured in Joules.

5 Results and Comparison

Throughput Packet data obtained from the simulation-based experiments have been used to calculate the throughput. Figure 1 illustrates the throughput of AODV, DSDV, and OLSR routing protocols with respect to different node densities. As compared to OLSR and DSDV, AODV has shown better performances for lesser and dense node densities.

Packet Loss Ratio (PLR) Figure 2 reveals the packet loss scenarios in AODV, DSDV and OLSR. Compared to three protocols the AODV and OLSR routing protocol has experienced a smaller number of packet losses. DSDV has more packet losses. It is observed that AODV has less packet loss when node density is increased and OLSR has more packet loss with increasing node density.

Packet Delivery Ratio Figure 3 shows the PDR directly measures the reliability of a network. Here, as compared to AODV and DSDV, the OLSR routing protocol has delivered the data packets better for every value of “N.” AODV with an increasing number of nodes has shown a better delivery ratio.

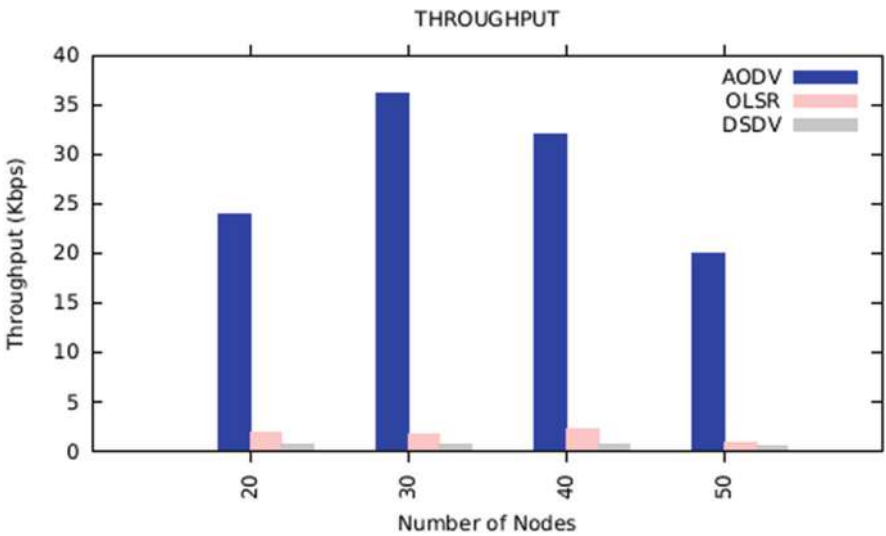


Fig. 1 Performance metrics of AODV

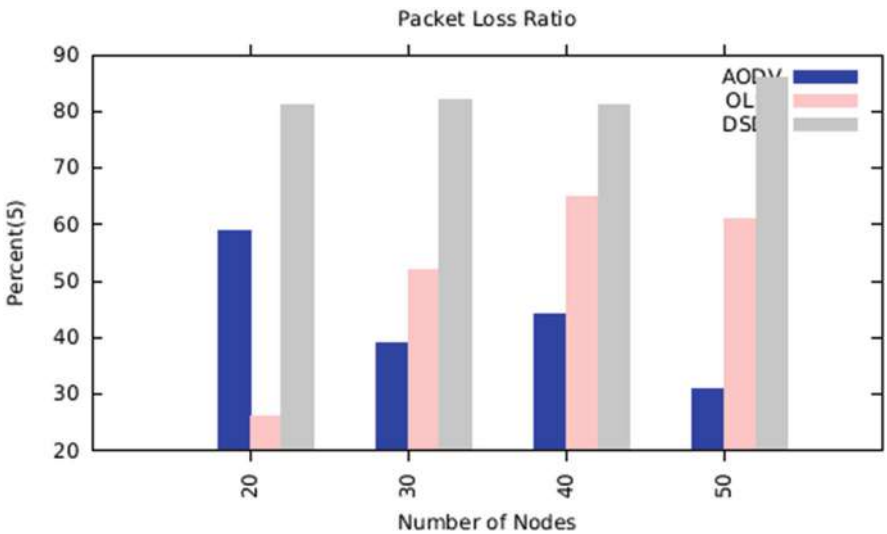


Fig. 2 Packet loss scenarios

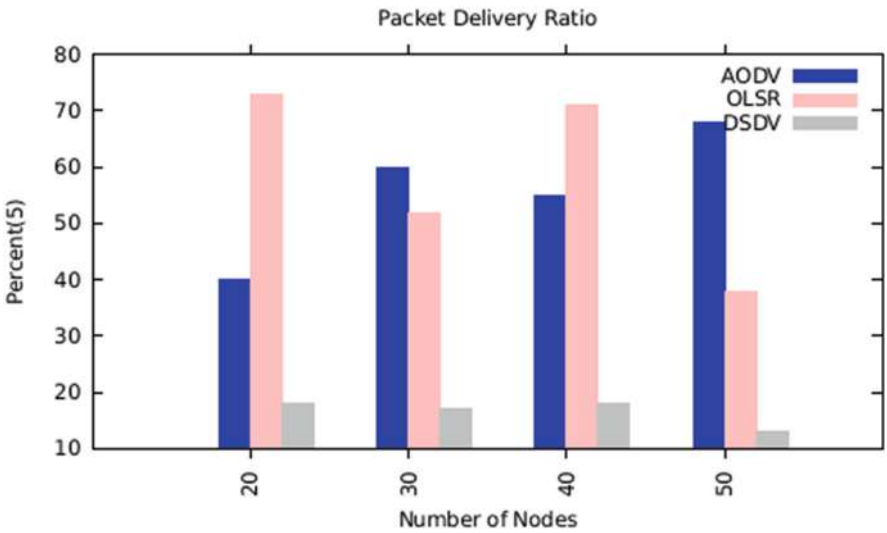


Fig. 3 Packet delivery ratio The figure considers a pause time of 10 s for 20 nodes

End-to-end Delay Delay scenarios reveal better performance of the OLSR as it encountered less delay during packet transmission. As compared to AODV, the DSDV protocol has a lesser delay for high node densities. Figure 4 shows the end-to-end delay during data transmission.

Jitter Delay Delay jitter is the variation of the delays with which packets traveling on a network connection reach their destination as shown in Fig. 5. For good quality

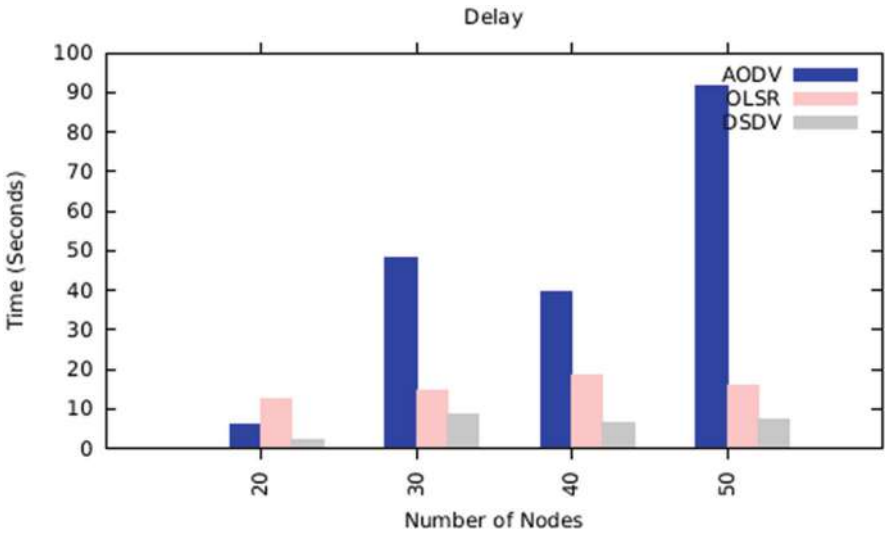


Fig. 4 End-to-end delay The figure considers a pause time of 10 s for 20 nodes

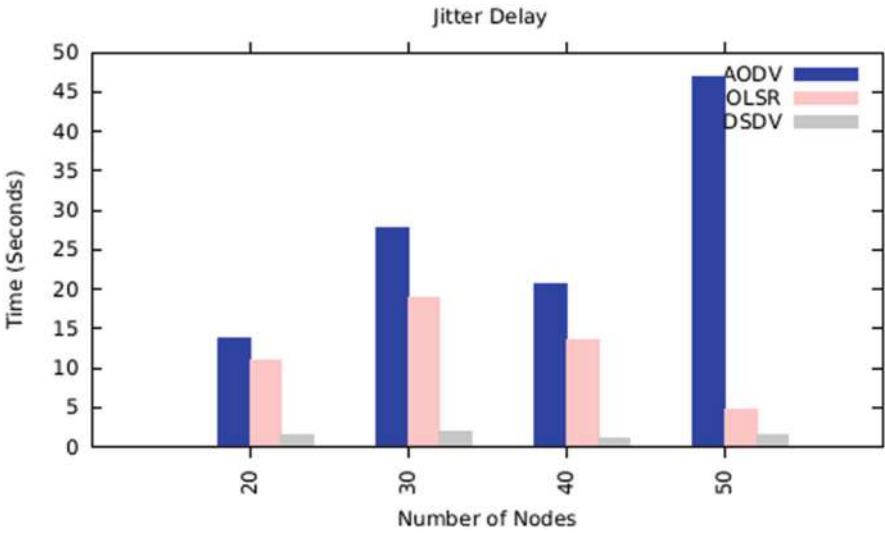


Fig. 5 Jitter delay

reception, continuous-media (video, audio, image) streams require that **jitter** be kept below a sufficiently small upper bound.

6 Conclusion

The proposed two strategies for conserving energy in IoT systems proved that using the best routing protocol and monitoring the energy levels of nodes will help in utilizing the node's energy efficiently, thus increasing the network lifetime. MANET environment is considered and tested using the proposed strategies. The proposed energy model as strategy one efficiently tracks, monitors, and manages the energy levels of the source, harvester, and nodes in the network. The results showed that among the protocols that we used for MANET, DSR is consuming a high level of energy whereas AODV showed optimal performance. The main weakness of DSDV is low throughput and high packet loss. Hence, AODV is considered as an optimal protocol for our energy model using MANET.

References

1. Omar Said, Zafer Al-MAKHADMEH AND AMR TO LBA Suggested Ems:" An Energy Management scheme for green IoT Environments". This work supported by the deanship of Scientific research with king Saud University through a Research Group under Grant RG-1438-027. February 27, 2020.
2. Rakesh Kumar Lenka, Amiya Kumar Rath and Suraj Sharma suggested "Building Reliable Routing Infrastructure for Green IoT Network". This work was supported by the IoT and Cloud Computing Lab, IIIT Bhubaneswar. September 6, 2019
3. Mohamed Er-Rouidi, Houda Moudni, Hicham Mouncif, [Abdelkrim Merbouha](#) Suggested "An Energy Consumption Evaluation of Reactive and Proactive Routing Protocols in Mobile Ad-Hoc Network". Published in [2016 13th International Conference on Computer Graphics, Imaging and Visualization](#). Added to IEEE Explore 12 May 2016.
4. Amjad Anvari-Moghaddam, Ashkan Rahimi-Kian "Optimal Smart Home Energy Management Considering Energy Saving and a Comfortable Lifestyle".
5. Dr. Shridhar A Joshi, Sri Jay Kolvekar, Y. Rahul Raj and Shashank Singh "IoT Based Smart Energy Meter "The smart meter is wirelessly connected to users by the means of IoT.
6. Vinita Tahiliani, Mayuri Digalwar "Green IoT Systems: An Energy Efficient Perspective" This paper presented the current state of the art research on energy optimization in IoT.
7. Saurabh Singh, Seo Yeon Moon, Gagman Yi, Jong Hyuk Park "Energy Consumption and Reliable Communications for Green IoT".
8. V. Preethi, G. Harish "Design and implementation of smart energy meter" This paper presents a smart energy meter for an automatic metering and billing system.
9. Tri Kuntoro Priyambodo, Danur Wijayanto and Made Santo Gitakarma "Performance Optimization of MANET Networks through Routing Protocol Analysis ".
10. Ejaz et al. proposed an optimized and schedule energy-efficient framework for smart cities in addition to energy harvesting which extended the lifetime of low power devices.
11. Danielly B. Avancini, Joel J.P.C Rodrigues "A new IoT-based smart energy meter for smart grids"
12. Mahammed Rehman shareef et al. Data Migration Using Oracle In Government Agency Applications, NeuroQuantology, August 2022, Volume 20, Issue 8, Page 10613–10619. <https://doi.org/10.48047/nq.2022.20.8.nq221084>.

Unveiling Hate Speech: Identifying Toxic Comments Targeting Women in Online Social Media Posts



M. Naveen Kumar, Kumari Gubbala, U. Mahender, Hari Keerthana Guna, and Chilaka Venkateswarlu

Abstract In today's digital world, where online platforms like Facebook, Twitter, and WhatsApp are ubiquitous, it is imperative to address the challenges faced by women in various roles, including as employees or team leaders within organizations. Unfortunately, women often encounter situations involving harassment, abusive language, toxic remarks, and hate speech on a daily basis. To tackle this issue effectively, it is crucial to employ efficient automatic detection tools capable of identifying such problematic comments across online forums, websites, blogs, YouTube, and social media networks. Manually searching for each toxic word within comments is time-consuming and inefficient. To address this, we conducted an analysis on a toxic comment dataset, utilizing natural language processing (NLP) techniques. NLP plays a vital role in processing text effectively and understanding its meaning. In our approach, we employed a transformer model for comment classification. Transformer models are powerful neural networks that grasp the semantics of sequential data by capturing relationships between different parts, such as words within a sentence. By leveraging the capabilities of transformer models, we aimed to identify and address toxic comments in a more efficient manner.

M. Naveen Kumar

Department of CSE-Data Science, CMR Engineering College, Medchal, Hyderabad, Telangana, India

K. Gubbala (✉) · H. K. Guna · C. Venkateswarlu

Department of CSE, CMR Engineering College, Medchal, Hyderabad, Telangana, India

e-mail: dr.kumarigubbala@cmrec.ac.in

U. Mahender

Department of CSE, CMR Engineering College, Medchal, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_12

Keywords Toxic comment · Hate speech · Transformer · DistilBERT · Word2Vec · NLP · ML

1 Introduction

Hate speech means remarks that target groups or individuals on the basis of perceived characteristics (such as race, religion, or gender) and which threaten social harmony [1]. Hate violence and intolerance are incited by hate speech. As sadly well known as it is, hatred can be devastating to a person's life. With the advent of new communications technologies, the scale and impact of this phenomenon have increased significantly [2]. It has become more common than ever to spread hate speech online, which threatens global peace as it spreads divisive rhetoric everywhere. Social media has been a major source of hate speech, so automated analysis of hate speech is an important tool in the fight against it [3]. Our goal is to create a model that is capable of detecting different types of toxicity such as threats, obscenity, insults, and identity-based hate. Classifying toxic comments has become one of our most important goals. Several approaches have been proposed recently in this field of research. The approaches, however, address some of the task's challenges, but others remain unresolved and need to be explored further. Natural language processing deals with common challenges of active research on this toxic comment classification [4].

2 Related Review

Samujjwal Goswami and coauthors [5] have trained a model for classifying fake news articles using the NLP technique known as TD-IDF (term frequency-inverse document frequency) vectorization. This technique provides the importance of each keyword within the news article or speech. They further employed logistic regression to classify the articles or speeches, aiding in the identification and handling of such deceptive content. Martins et al. [6] proposed a system that expands the original dataset with emotional information using natural language processing (NLP) techniques. This augmented dataset is then utilized for machine learning classification. The authors achieved an accuracy of 80.56% in identifying hate speech. They combined lexicon-based and machine learning approaches, incorporating sentiment analysis to predict hate speech within text based on emotional cues. In their work, Davidson et al. [7] introduced a crowd-sourced hate speech lexicon. They collected tweets containing hate speech keywords and classified them into three categories: hate speech, offensive language only, and neither. A multi-class classifier was trained to distinguish between these categories. The study revealed that racist and homophobic tweets were more likely to be classified as hate speech, while sexist tweets were typically categorized as offensive. Classifying tweets

without explicit hate keywords posed greater difficulty. Kovács et al. [8] proposed a deep natural language processing (NLP) model that combines convolutional and recurrent layers to automatically detect hate speech in social media data. Their model was applied to the HASOC2019 corpus, achieving a macro F1 score of 0.63 in hate speech detection on the test set. However, they noted that the high learning efficiency of deep neural networks also increased the risk of overfitting. These studies highlight various approaches to addressing hate speech detection, including TD-IDF vectorization, emotional information augmentation, lexicon-based methods, and deep NLP models. Each contributes to the ongoing efforts in combating hate speech and fake news, though challenges such as classification accuracy and handling nuanced cases still persist. Rahul et al. [9] used six types of machine learning classification algorithms to dataset to address the classification problem and to identify the best machine learning algorithm for toxic comment classification. Yuan et al. [10] introduced a two-phase deep learning model designed for discrimination analysis of tweets. Their model follows a sequential process, wherein it initially learns text representations using weakly labeled tweets that include specific hashtags. Subsequently, the model is trained on a small set of well-labeled training data to enhance its classification abilities. In this study, Ghinaa Zain Nabilah et al. [11] conducted an experiment to detect comments containing toxic sentences on social media in Indonesia. To accomplish this, a Pre-Trained Model specifically trained for Indonesian language was utilized. The study employed a multilabel classification approach and evaluated the classification performance of three models: Multilingual BERT (MBERT), IndoBERT, and IndoRoberta Small. The results of the study revealed that the IndoBERT model achieved the highest performance, with an F1 Score of 0.8897. This suggests that the IndoBERT model is particularly effective in accurately identifying and classifying toxic comments within social media discussions in the Indonesian language. Mehdad et al. [12] presented work on abusive language detection with simple character based and token based and also used two deep learning methods.

3 Research Methodology

3.1 Dataset Collection

- A dataset of 3,12,735 comments has been collected. Out of these 1,59,571 were used for training while the remaining 1,53,164 were used for testing. These comments have been classified into six toxic behaviors. The classes are “Toxic, Severe toxic, Obscene, Threat, Insult, Identity hate” provided with a large number of Wikipedia comments that have been labeled by human raters for toxic behavior.
- Toxic
- Severe_toxic

- Obscene
- Threat
- Insult
- Identity_hate

This is a multi-classification dataset to find different kinds of toxic comments.

3.2 *Preprocessing*

The preprocessing steps encompass several key tasks, including removing stop-words, tokenization, stemming, lemmatization, checking for missing values, and eliminating special symbols and hyperlink values.

3.3 *Word Embedding*

In Word2Vec, each word in the corpus is associated with a vector representation. To initialize these vectors, we typically begin with either a random vector or a one-hot vector encoding.

A one-hot vector is a binary representation where only one element in the vector is 1, and all other elements are 0. For instance, if there are 500 unique words in the corpus, each word will be represented by a vector of length 500. The position corresponding to the word's index in the vocabulary will have a value of 1, while all other positions will be 0. Once we have assigned vectors to each word, we proceed with the training process using a specified window size. The window size determines the number of neighboring words considered during the training process. We then iterate through the entire corpus, focusing on each word in turn. For each word, we aim to predict the probability of its neighboring words occurring within the window. The model is trained by adjusting the word vectors based on how well it predicts the context words in the given window. The ultimate goal is to learn word embeddings that capture semantic relationships and similarities between words in the corpus. This process enables Word2Vec to create dense, continuous vector representations for words, often leading to better performance in various natural language processing tasks.

3.4 *DistilBERT*

DistilBERT is a language model and a variant of the BERT (Bidirectional Encoder Representations from Transformers) model. The primary motivation behind DistilBERT is to reduce the computational complexity and memory requirements of

the original BERT model while maintaining a similar level of performance. It achieves this through a process called “knowledge distillation,” where a larger pre-trained model like BERT is used to teach a smaller model like DistilBERT, transferring its knowledge to the smaller model. After pre-training on a large corpus of text, DistilBERT can be fine-tuned on specific downstream NLP tasks with labeled data. Fine-tuning involves adapting the pre-trained model to the target task by adding a task-specific layer and training it on the task’s dataset. This transfer learning approach helps in achieving better performance on various NLP tasks, such as text classification, named entity recognition, sentiment analysis, and more. The efficiency of DistilBERT stems from its reduced model size compared to BERT, which allows it to run faster and require less memory. This advantage makes it particularly useful for applications where computational resources are limited or real-time inference is required. Overall, DistilBERT has proven to be a significant advancement in NLP by offering a compromise between model size and performance, making it more accessible for practical use in a wide range of NLP tasks and applications.

3.5 Model Evaluation

Model evaluation is a crucial step in assessing the performance and effectiveness of a machine learning or natural language processing model. It involves measuring how well the model performs on a specific task or set of tasks and comparing its predictions against ground truth or human-labeled data. Accuracy measures the proportion of correctly predicted instances out of the total number of instances. It is commonly used for classification tasks where the goal is to predict the correct class label either toxic or nontoxic comment. Figure 1 shows the several steps to classify the comments into toxic or nontoxic.

Transformers, a popular deep-learning framework, can be effectively used for toxic comment classification tasks.

Pre-trained Transformer Models: Transformers offer pre-trained models that have been trained on large-scale datasets, such as BERT, RoBERTa, or DistilBERT. These models have learned rich contextual representations from a vast amount of text data, which can be fine-tuned for specific tasks like toxic comment classification. In our model, DistilBERT model is used.

Fine-tuning: Fine-tuning involves taking a pre-trained transformer model and adapting it to the target task using labeled data. The pre-trained model’s weights are loaded, and a task-specific classification layer is added on top. The entire model is then trained on the toxic comment dataset using appropriate optimization techniques like gradient descent and backpropagation.

Training and Validation: Split the dataset into training and validation sets. During training, the model learns to predict the toxicity label based on the comment text.

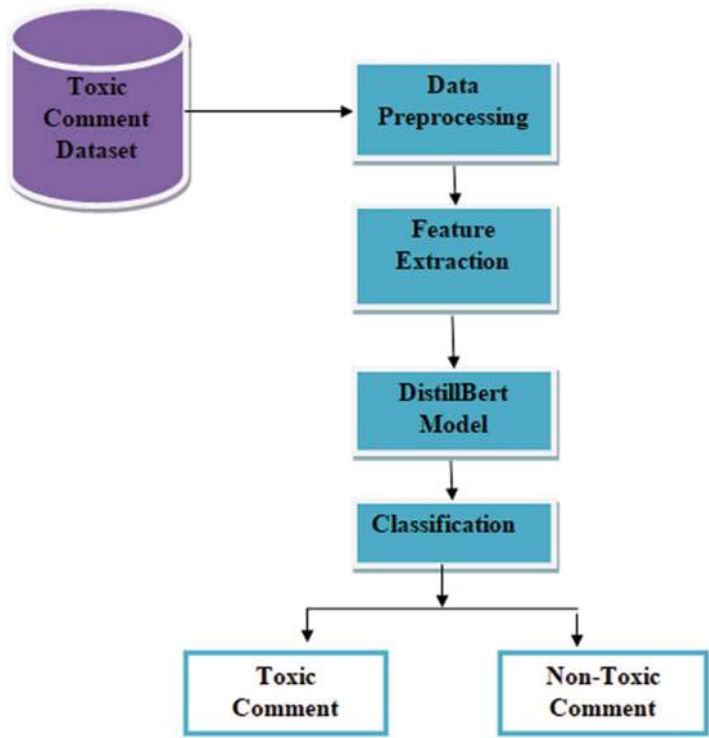


Fig. 1 Toxic comment classification

The model’s performance is evaluated on the validation set to monitor progress and prevent overfitting.

Hyperparameter Tuning: Experimented with hyperparameters such as learning rate, batch size, and number of training epochs to optimize the model’s performance.

Evaluation: Once the model is trained, evaluating its performance on an independent test set. Common evaluation metric for toxic comment classification is accuracy. These metrics provide insights into the model’s ability to classify comments as toxic or nontoxic.

4 Results

Multiple class labels are included in our dataset, such as toxic, severe_toxic, obscene, threat, insult, and identity_hate. According to dataset, toxic labels occur at the highest frequency, while obscene words occur at the second highest frequency. Here is a list of the comments in each category as shown in Fig. 2.

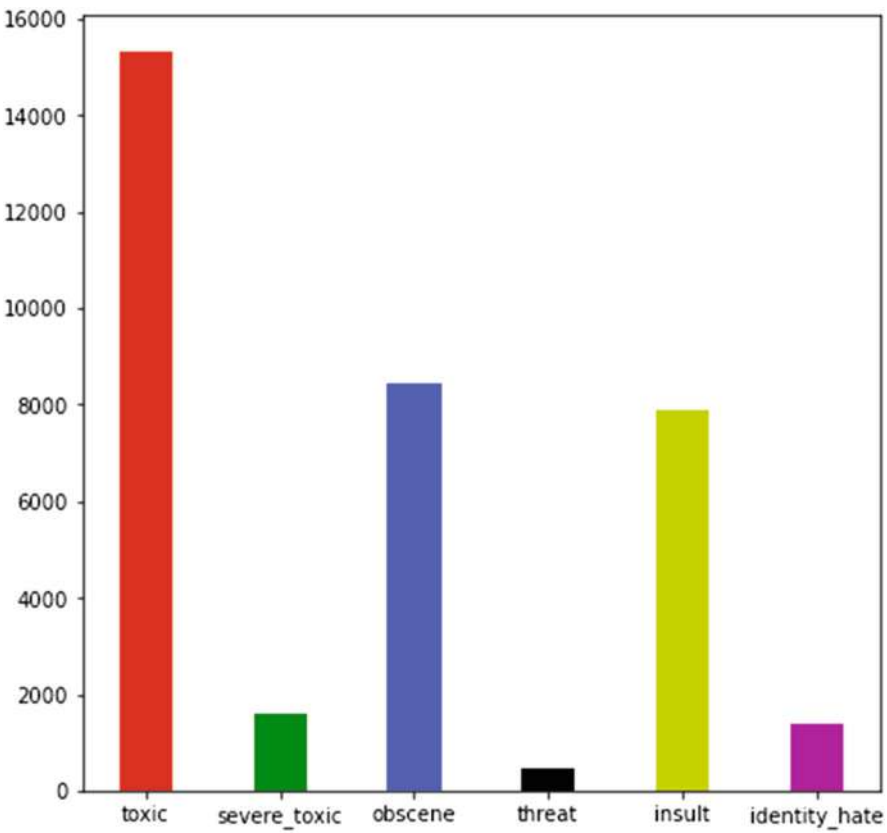


Fig. 2 Comments in each category

Figures 3 and 4 depict wordclouds showcasing the most frequent words in our dataset. Figure 3 specifically highlights the most commonly used words in non-toxic comments. These words, such as “article,” “Wikipedia,” “page,” “source,” etc., do not contain any toxic or hateful language. Their presence reinforces the absence of toxic content in the comments represented by these words.

Figure 4 illustrates the wordcloud representing the most frequently used words in toxic comments. This wordcloud includes words like “fuck,” “suck,” “fucking,” “shit,” etc., which indeed contain explicit toxic and offensive language. These words are indicative of hate speech and are particularly concerning as they target women. The presence of such words highlights the presence of toxic and derogatory language aimed at women within the dataset.

Figure 5 presents the classification of comments using the DistilBERT transformer model. The comments are assigned labels 0 and 1, where label 0 represents nontoxic comments and label 1 represents toxic comments. This classification process enables the identification and differentiation between toxic and nontoxic

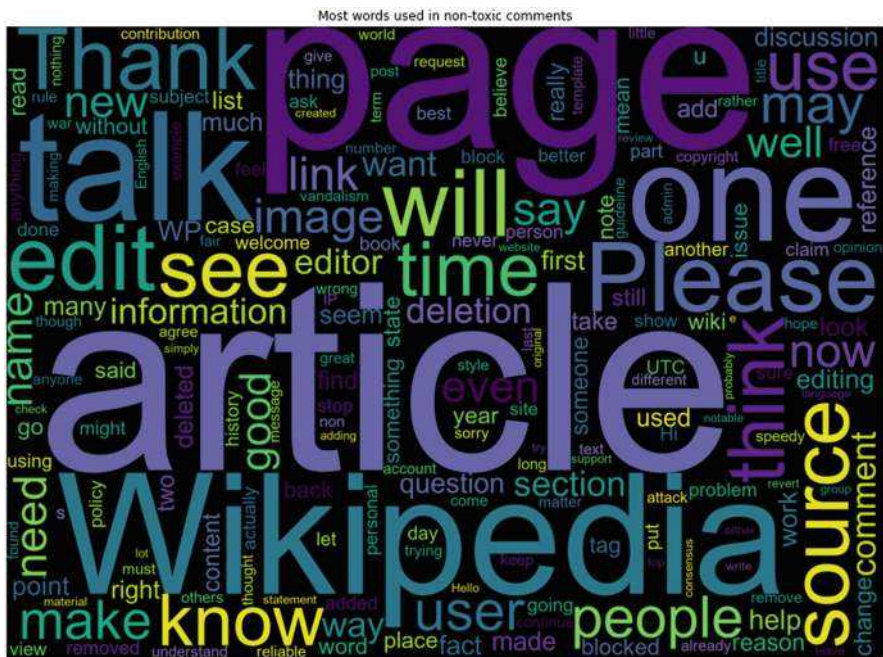


Fig. 3 Most words used in nontoxic comments

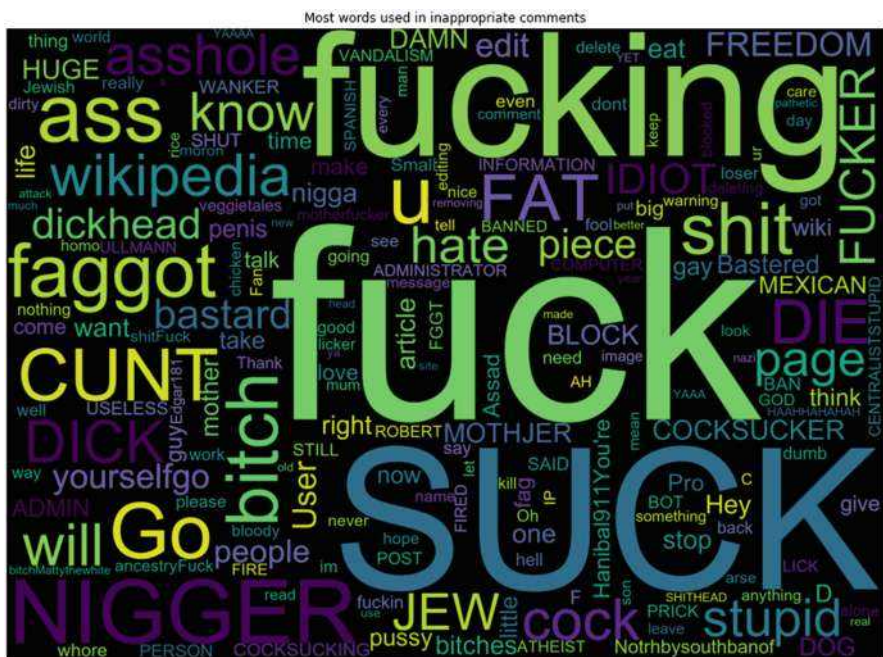


Fig. 4 Most words used in inappropriate comments

Out[6]:

	comment_text	labels
0	COCKSUCKER BEFORE YOU PISS AROUND ON MY WORK	[1, 1, 1, 0, 1, 0]
1	Bye! \n\nDon't look, come or think of comming ...	[1, 0, 0, 0, 0, 0]
2	FUCK YOUR FILTHY MOTHER IN THE ASS, DRY!	[1, 0, 1, 0, 1, 0]
3	GET FUCKED UP. GET FUCKEED UP. GOT A DRINK T...	[1, 0, 1, 0, 0, 0]
4	=Tony Sidaway is obviously a fistfuckee. He lo...	[1, 0, 1, 0, 1, 0]

Fig. 5 Classification of comments with 0—nontoxic and 1—toxic labels

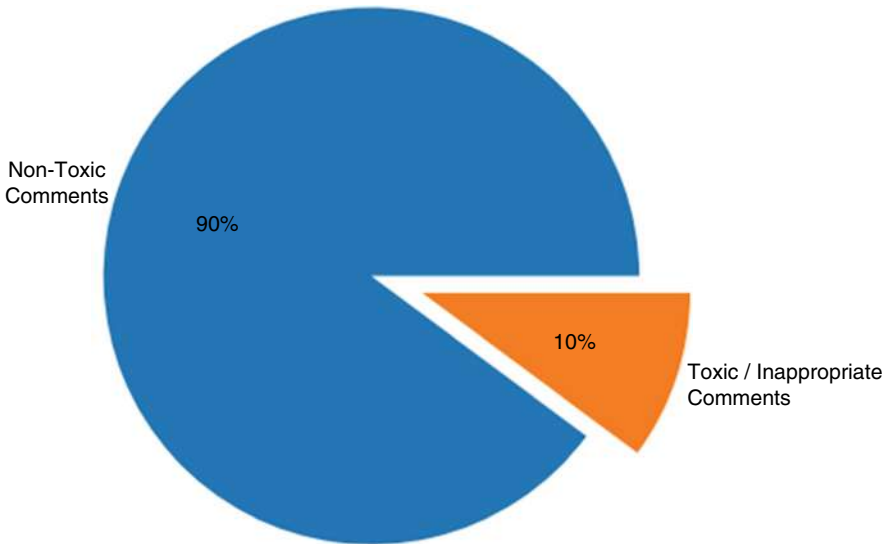


Fig. 6 Dataset division with toxic and nontoxic comments

content within the dataset. The dataset comprises 90% nontoxic comments and 10% toxic comments which is shown in Fig. 6.

The DistilBERT transformer model has exhibited remarkable performance, as evidenced by the learning curve depicted in Fig. 7. Over the course of multiple iterations, the model consistently reduced its loss values, indicating its robust learning capabilities. In addition, several experiments were conducted, incorporating established models like Support Vector Machines (SVM) with an accuracy of 88% and the BERT base model with 89% accuracy. Notably, our fine-tuned DistilBERT model achieved an impressive accuracy of 92%, as shown in Fig. 8. These results unequivocally demonstrate that fine-tuning the DistilBERT model yielded highly satisfactory outcomes.

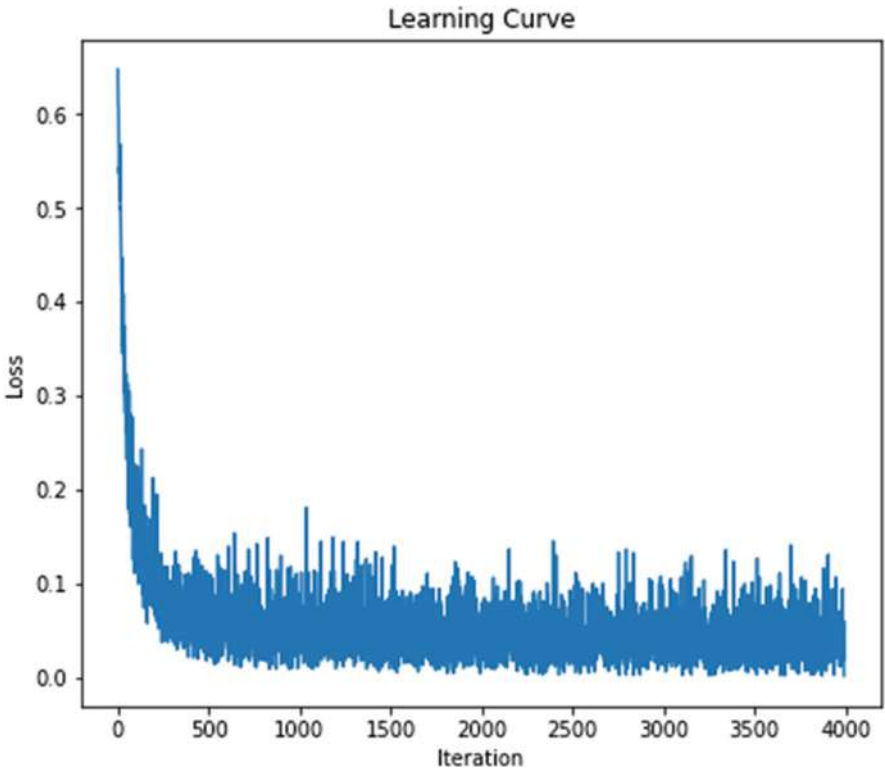


Fig. 7 Model performance for DistilBERT model

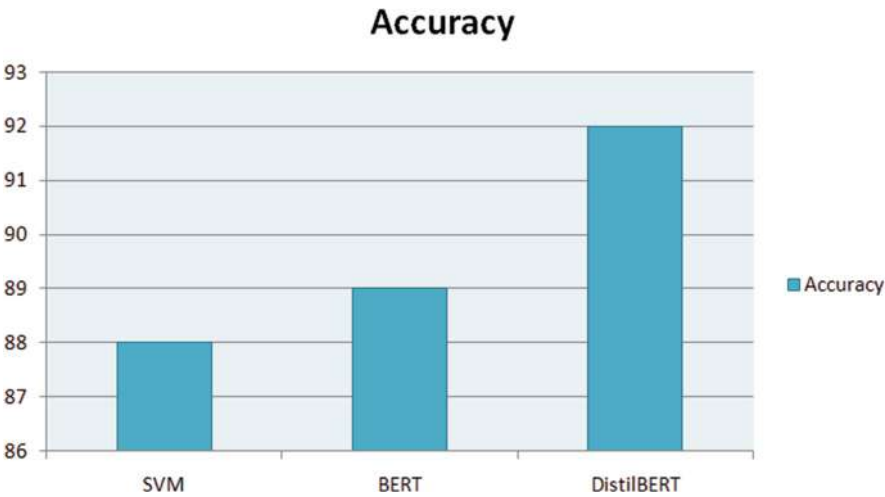


Fig. 8 Comparison of classification models SVM, BERT, and DistilBERT

5 Conclusions and Future Scope

Toxic comment classification presents a complex research challenge that has been addressed using various machine learning methods. This topic remains active and challenging, with ongoing efforts to improve classification accuracy. In this study, we employed a transformer model combined with natural language processing techniques to achieve precise results. The Word2Vec word embedding model was utilized to identify frequent words in distinguishing between toxic and nontoxic comments. Multilabel classification was performed on a substantial dataset, employing the DistilBERT transformer model for the classification task. Moving forward, we aim to explore different transformer models to further enhance the effectiveness of the results. Additionally, we plan to implement these models in emotion analysis to discern the moods of different individuals based on their comments. There is potential for further extension of this work to encompass diverse datasets and incorporate fine-tuning of different models in conjunction with natural language processing (NLP) tasks. For instance, word embeddings like Glove and Word2Vec could be integrated, considering the intensity levels of words. This approach would enhance the breadth and depth of the study, opening avenues for more comprehensive analysis and exploration within NLP.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Malik, Jitendra & Pang, Guansong & Hengel, Anton. (2022). Deep Learning for Hate Speech Detection: A Comparative Study. <https://doi.org/10.48550/arXiv.2202.09517>.
2. Alkiviadou N. The legal regulation of hate speech: the international and European frameworks. *Politička misao*. 2018;55:203–29. <https://doi.org/10.20901/pm.55.4.08>.
3. Alonso P, Saini R, Kovács G. Hate speech detection using transformer ensembles on the hasoc dataset. In: *Speech and Computer: 22nd International Conference, SPECOM 2020, St. Petersburg, Russia, October 7–9, 2020, Proceedings*, vol. 12335, p. 13. Springer Nature; 2020.
4. Arango A, Pérez J, Poblete B. Hate speech detection is not as easy as you may think: A closer look at model validation. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'19*, p. 45–54. Association for Computing Machinery, New York, NY, USA; 2019. <https://doi.org/10.1145/3331184.3331262.5>.
5. Samujjwal Goswami, Manoj Hudnurkar, Suhas Ambekar. “FAKE NEWS AND HATE SPEECH DETECTION WITH MACHINE LEARNING AND NLP”. *PalArch's Journal of Archaeology of Egypt / Egyptology*, vol. 17, no. 6, Dec. 2020, pp. 4309–22, <https://archives.palarch.nl/index.php/jae/article/view/1686>.
6. Martins, Ricardo, Gomes, Marco, Almeida, João, Novais, Paulo, Henriques, Pedro. (2018). Hate Speech Classification in Social Media Using Emotional Analysis. 61–66. <https://doi.org/10.1109/BRACIS.2018.00019>.

7. Davidson, T., D. Warmesley, M. Macy, and I. Weber. "Automated Hate Speech Detection and the Problem of Offensive Language". *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 11, no. 1, May 2017, pp. 512–5, doi:<https://doi.org/10.1609/icwsm.v11i1.14955>.
8. Kovács, G., Alonso, P. & Saini, R. Challenges of Hate Speech Detection in Social Media. *SN COMPUT. SCI.* 2, 95 (2021). <https://doi.org/10.1007/s42979-021-00457-3>.
9. Rahul, H. Kajla, J. Hooda and G. Saini, "Classification of Online Toxic Comments Using Machine Learning Algorithms," *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2020, pp. 1119–1123, <https://doi.org/10.1109/ICICCS48265.2020.9120939>.
10. Yuan S, Wu X, Xiang Y. A two phase deep learning model for identifying discrimination from tweets. In: Pitoura E, Maabout S, Koutrika G, Marian A, Tanca L, Manolescu I, Stefanidis K (eds.) *Proc. EDBT 2016*, Bordeaux, France, March 15–16, 2016, Bordeaux, France, March 15–16, 2016, pp. 696–697. [OpenProceedings.org](https://openproceedings.org/2016/edbt.2016.92); 2016. <https://doi.org/10.5441/002/edbt.2016.92>.
11. Ghinaa Zain Nabiilah, Simeon Yuda Prasetyo, Zahra Nabila Izdihar, Abba Suganda Girsang, BERT base model for toxic comment analysis on Indonesian social media, *Procedia Computer Science*, Volume 216, 2023, Pages 714–721, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2022.12.188>.
12. Mehdad Y, Tetreault J. Do characters abuse more than words? In: *Proceedings of the SIGDIAL2016 conference*, pp. 299–303; 2016. <https://doi.org/10.18653/v1/W16-3638>.

Identification of Retinal Fundus in Diabetic Patients Using Deep Learning Algorithms



Vidyullatha Sukhavasi  and Divya Yeluri 

Abstract Diabetic retinopathy (DR) is a complication of diabetes that causes retinal changes that affect vision. If not caught early, it can lead to blindness. Unfortunately, DR is not a reversible process where treatment only controls vision. Early detection and treatment of DR can reduce the risk of blindness. Compared with a computer-aided diagnosis system, the process of manually diagnosing retinal fundus DR images by ophthalmologists is time-consuming, laborious, expensive, and prone to misdiagnosis. Deep learning has become one of the most widely used methods in recent years and achieved great success in many fields, especially in the analysis and classification of medical imaging. The most popular and efficient deep learning technique for medical image processing used is a convolutional neural network. In this project, DR color fundus photos that were examined and analyzed were used to classify and recognize objects using deep learning techniques. Diabetic retinopathy was classified using a deep learning algorithm. The prediction of diabetic retinopathy is the main topic of this essay. The CNN model is trained using training datasets, and CNN will provide the likelihood of a diabetic eye infection.

Keywords Classification · Convolution neural network · Deep learning · Diabetic retinopathy · Feature extraction

1 Introduction

A collection of metabolic conditions known as diabetes mellitus (DM) are characterized by hyperglycemia as a phenotype, which is brought on by an absolute or relative shortfall in insulin synthesis or action. Damage, dysfunction, or failure of organs

V. Sukhavasi (✉) · D. Yeluri

Computer Science and Engineering, BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telangana, India

e-mail: divya.y@bvrithyderabad.edu.in

and tissues, such as the retina, kidney, nervous system, heart, and blood vessels, is linked to chronic hyperglycemia [1]. Both type 1 and type 2 diabetes patients can develop diabetic retinopathy, especially if they have a chronic condition and poor glycemic control [2]. The metabolic implications of persistent hyperglycemia are a part of the multifactorial etiology. The need for the early detection of DR has encouraged the search for more effective alternative screening techniques. The main factors involved in the genesis of diabetic retinopathy are the biochemical (earlier), hemodynamic, and endocrine factors, which interact with one another to cause nonenzymatic protein glycation, oxidative stress, accumulation of polyols, and activation of protein kinase C [3]. Automatic retinal image analysis models are choices that attempt to lessen the effort of doctors and offer a useful and affordable method [4].

Diabetic retinopathy is the most prevalent cause of blindness in adults in developed nations. The affected population is assumed to be 93 million. The World Health Organization and the US Centers for Disease Control and Prevention have estimated that there are 347 million cases of diabetes worldwide. Uncontrolled diabetes is linked to an eye disorder known as diabetic retinopathy (DR). The illness is at some stage in 40–45% of diabetic Americans. Today, DR is diagnosed by a skilled professional performing a thorough physical examination using digital color fundus images of the retina. By the time human experts submit their reviews, which is frequently a day or two later, the delayed results cause a lack of follow-up, inaccurate information, and delayed treatment.

This study focuses on the automated computer-aided detection of diabetic retinopathy (DR) using features derived from the output of various algorithms for processing retinal images, such as the diameter of the optic disc, lesion-specific features (micro-aneurysms, exudates), and image level features (prescreening, AM/FM, quality assessment). To detect the disease automatically machine learning algorithms like Naïve Bayes, SVM and KNN are used and compared with a deep learning model. Deep neural networks can now be used for the recognition and categorization of diabetic retinopathy due to the recent advancements in computer vision and the accessibility of vast datasets. No DR, mid-nonproliferative DR (MNPDR), moderate nonproliferative DR, severe nonproliferative DR, and proliferative DR are the five stages of DR as shown in Fig. 1. MNPDR is highly difficult to diagnose because it has symptoms like decreased retinal pericytes, increased leukocyte adhesion, and slowed retinal blood flow.

These are the plans for the remaining chapters of the text. A list of researchers that have researched in this field is provided in Sect. 2. The background research on machine learning methodologies and functions is discussed in Sect. 3. Diabetes prediction methods and outcomes are described in next sections by concluding the paper with future scope.

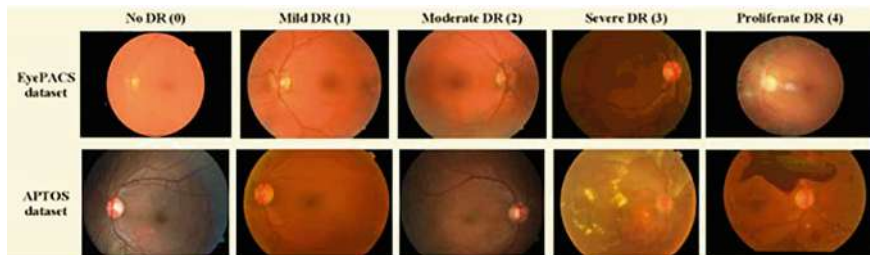


Fig. 1 Stages of diabetic retinopathy

2 Related Work

A CNN model for the referable eye disease due to diabetes was proposed by Antal et al. [5]. They employed two freely accessible datasets, Kaggle and DiaretDB1, using Kaggle for training and DiaretDB1 for testing. For classifying patients, they adopt a binary system in which the normal and mild phases are used as non-referable DR and other three stages are used as referable DR. Based on binary classification, the CNN model performs with sensitivity of 93.6% and specificity of 97.6% on DiaretDB1. A method based on SVM was suggested by Enrique Carrera et al. [6] to assist in the early identification of diabetic retinopathy. In this method, the blood vessels, microaneurysms, and hard exudates are separated during the preprocessing stage in order to extract features for SVM. The model's sensitivity on DiaretDB1 was 94.6% when evaluated on the STARE dataset.

In their chapter, Satish Kumar et al. [7] proposed a detection methodology for diabetic retinopathy that relied on a linear support vector machine approach. They take the color fundus image and extract the number of microaneurysms. Localized capillary dilatations are called microaneurysms, which are typically saccular (round). They are formed by absorbing brittle blood artery parts, and they are visible as little red dots on the retina. Using the DIARETDB1 database, their method could accomplish 96% and 92% specificity as well as sensitivity, respectively.

Sarni Suhaila Rahim et al. [8] provided several methods for microaneurysm detection. The system I have utilized adaptive histogram equalization, wavelet transform filtering, and morphological preprocessing. Pixel area, mean, and standard error are valid DR properties. Decision trees, K-nearest neighbors, multidomain SVMs, and radial basis function (RBF) based core SVMs are used for classification. Many attempts have been made to classify OCT images. For example, in the 1990s researchers proposed and developed Local Binary Models (LBPs) that can be used to classify OCT fundus images.

As demonstrated by Silva et al. in 2015, this type of imaging failed to differentiate nonproliferative DR [9]. Multicolor lasers and infrared rays are used in retinal imaging to improve OCT output, leading to a more accurate classification of fundus

images. This approach can detect defects such as disks but is still insufficient for accurate DR classification. To improve the performance of the model, the functional image was proposed by Gharaibeh et al. [10].

Exudate and hemorrhagic features are easily seen in computer-assisted diagnosis (CAD). This causes the fundus image to be divided into proliferative and nonproliferative disease, thus distinguishing moderate and large vascular abnormalities from low-level, less severe ones. CAD is a simple diagnostic method that contributes to the digital structure [11]. The usage of the approach data augmentation to lessen overfitting during model training was also noted by the authors. After reviewing new DL pipelines and ML procedures, Stolte and Fang [12], Attia et al. [13], and Asiri et al. [14] research papers discussed various DR grading jobs (such as grading, optic disc, blood vessels, lesions, and grading). In their discussion of current state-of-the-art (SoTA), CNN variations to classify DR, Valarmathi, and Vijayabhanu [15] drew attention to the inconsistent use of evaluation criteria for model evaluation in the literature. A thorough explanation of the operation of transformers for a variety of medical imaging goals, such as segmentation, classification, detection, and reconstruction, is given by Shamshad et al. [16].

3 Proposed Method

3.1 Dataset Characteristics

Gaussian-filtered retina scan images are used in the photographs to identify diabetic retinopathy. APTOS 2019 Blindness Detection has the original dataset available. These photos have been downsized to 224×224 pixels so that many pre-trained deep learning models may easily utilize them. All images have already been saved into the correct folders based on the stage of diabetic retinopathy using train.csv file. The corresponding images are located in five directories: 0—No DR, 1—Mild, 2—Moderate, 3—Severe, and 4—Proliferate DR are displayed in Fig. 2, along with a visualization of the data.

3.2 Preprocessing and Image Segmentation

Preprocessing procedures were employed to get each image into a format suitable for training due to the significant degree of noise in the data. Preprocessing involves scaling, cropping, and eliminating black areas from photographs. Images typically have a resolution of 4000×3200 , and cropping is used to remove the surrounding black area to obtain the retinal image center. The photographs are reduced to

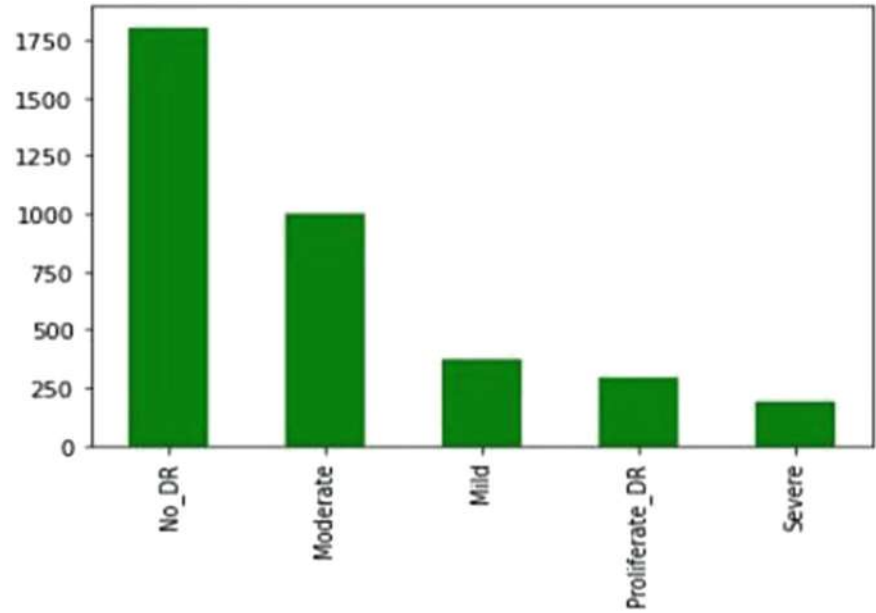


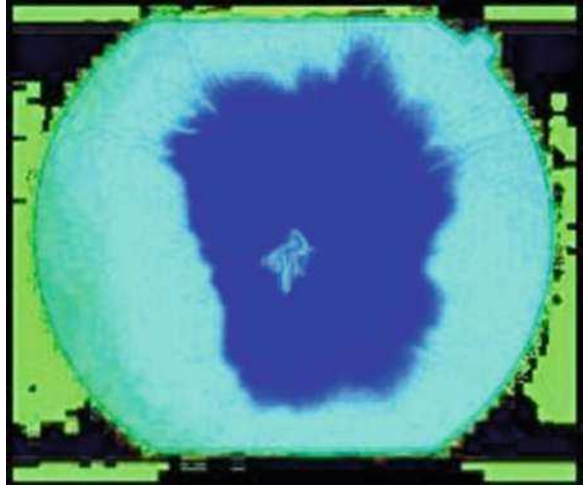
Fig. 2 Analysis of dataset characteristics

256×256 pixels in order to standardize the photos and reduce the number of data points. There are numerous completely blacked-out images in the dataset that must be removed. This is done by comparing each pixel’s arithmetic mean with zero. The image is dropped since it is black if the mean is 0. After preprocessing of retina images, segment smoothing, masking and bitwise AND was used. Smoothing is used to reduce noise from images. A smaller “image piece” is defined and used to alter a bigger image using the image processing technique known as masking. Here, we are using the blue ([0,0,255]) coloring to hide the optic disc and exudates that are yellow in color ([60,255,255]). Figure 3 gives the image after preprocessing.

3.3 Feature Extraction

Here, we are using two features, namely, the number of exudates as the first parameter and the number of hemorrhages and microaneurysms as the second parameter, for binary classification. In other words, we divide the total number of pixels in the image by the number of white pixels from the segmented images.

Fig. 3 Image after preprocessing



3.4 Classification

Naïve Bayes Naive Bayes is a probabilistic classifier based on the Bayes theorem that presupposes feature independence. To use Naive Bayes for retinal fundus identification, extract significant features from retinal fundus images. These characteristics could include statistical statistics, textural descriptors, or form attributes. Based on these features, Naive Bayes assesses the likelihood of a specific retinal fundus image belonging to the diabetes or nondiabetic class and provides the class label with the highest probability using Eq. (1).

$$P\left(\frac{y}{X}\right) = \frac{P\left(\frac{X}{y}\right) * P(y)}{P(X)} \quad (1)$$

$P\left(\frac{y}{X}\right)$ reflects the probability of the target variable (y) given the input features (X). Given the retinal fundus traits, it would be the possibility of retinal fundus abnormality in this situation.

$P\left(\frac{X}{y}\right)$ is the probability of the input features (X) given the target variable (y). It shows the possibility of observing retinal fundus features when a retinal fundus abnormality is present.

$P(y)$ is the prior probability of the target variable (y), which represents the overall probability of retinal fundus abnormality in the dataset.

$P(X)$ is the evidence probability, which represents the overall likelihood of observing the retinal fundus features in the dataset.

Support Vector Machine (SVM) SVM is a very effective supervised learning technique that is used for classification jobs. In the case of retinal fundus identification, SVM can be trained on a labeled dataset of retinal fundus images, with each image labeled to indicate whether or not the image belongs to a diabetic patient. SVM can be trained using image characteristics such as pixel values or more complex features such as texture or form descriptors. SVM determines the optimum hyperplane that separates the two classes by the greatest margin.

The SVM can be described using the following Eq. (2)

$$f(x) = \text{sign}(w^T * x + b) \quad (2)$$

where

$f(x)$ represents the forecast for a given input feature vector x .

w is the weight vector perpendicular to the hyperplane.

b is the bias term or hyperplane offset.

K-Nearest Neighbor(KNN) KNN is a nonparametric lazy learning method that classifies instances based on their similarity to neighboring examples. In the case of retinal fundus identification, KNN can be trained using a labeled dataset of retinal fundus images. The retrieved features from the photos are utilized to generate a feature space. When a new retinal fundus image needs to be classed, KNN searches the feature space for the k nearest neighbors and provides the class label depending on the majority class among those neighbors. The KNN algorithm assigns a class label to a new data point based on its K nearest neighbors' class labels. During the training phase, the algorithm does not learn a specific model but instead stores the labeled data points for later categorization.

Convolution Neural Network The proposed system workflow and architecture are shown in Figs. 4 and 5. Convolutional neural networks serve as its foundation, making it quick and less prone to lag. The suggested system is entirely cost-effective, simple to maintain, and energy-efficient. The retina may move away from the back of the eye as a result of the aberrant blood vessels linked to diabetic retinopathy stimulating the creation of scar tissue. Your vision may become severely impaired, experience flashes of light, or have floating spots appear in it. CNN methodology is quicker, less expensive, and more accurate. The proposed methodology was developed using more cutting-edge ML techniques, making it clear and easy to construct.

Convolutional neural networks are used in deep learning [17]. It is used in both image identification and natural language processing. By locating its features, a convolutional neural network (CNN) analyzes and forecasts a picture. The image captures the spatial and temporal relationships. Simple feature detection filters like corners and edges are picked up on by the early layers. The filtering for recognizing object sections is picked up by the middle layer. The final layers can recognize whole

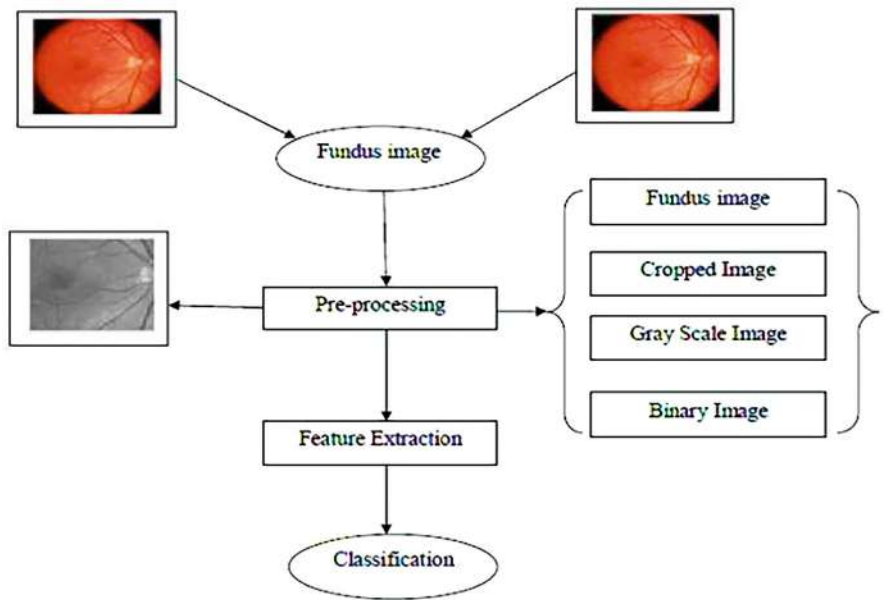


Fig. 4 Workflow diagram

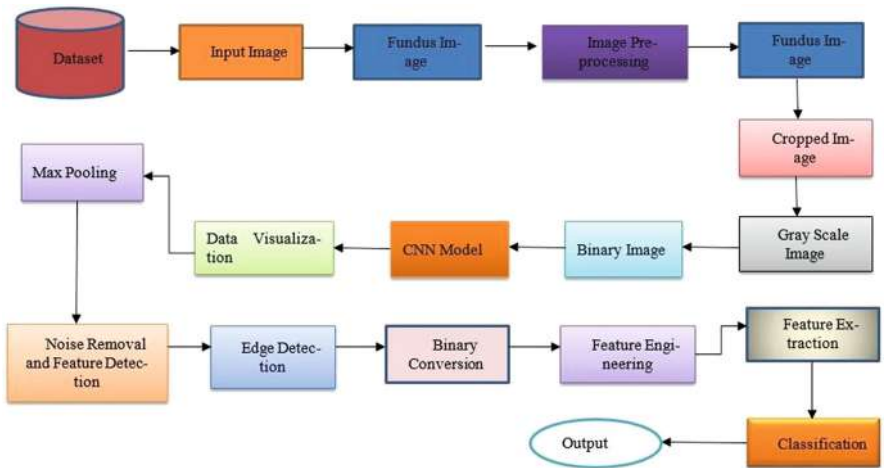


Fig. 5 Architecture of the proposed model

objects in a range of positions and shapes and have more complex representations. A convolutional neural network analyses the picture and recognizes it based on specific features. The basic steps of convolutional neural network are as follows:

1. Convolution
2. ReLU layer
3. Pooling
4. Flattening
5. Fully Connected Layer

In general, convolutional neural network architecture is generally versatile.

Convolution Layer. Convolution neural networks have many convolutional layers that can handle low to high levels depending on the layer we focus on. Early convolutional techniques also acquire higher-level features based on ideas of low-level features such as

RELU layer. Rectified Linear Units, often known as ReLU, are frequently employed in CNNs. Every negative value from the filtered image is taken out and replaced with zero in this layer. Only when the node input exceeds a specific threshold does this function become active. Hence, the output is zero when the input is less than zero.

Pooling layer. Convolutional neural networks are frequently used to categorize images. To lessen the chance of overfitting, we would prefer to lower the dimensionality of picture data because it is so high dimensional. Another type of layer commonly used in CNNs is the pooling layer. There are various types of pooling levels, with max-pooling being the most common. In essence, pooling reduces the spatial dimensions of the picture by applying mathematical techniques like average or maximum pooling. We frequently use pooling for picture classification because it usually functions as a noise suppressor, makes it translation-invariant, and helps capture important structural aspects of the represented images without losing the finer details.

Flattening. A single, extensive continuous linear vector is created by flattening the resulting 2-dimensional arrays from pooled feature maps. The fully linked layer receives the flattened matrix as input to categorize the image.

Fully connected layer. Convolution and pooling techniques are used to lower the dimensionality of the data before it is sent to the fully connected (Dense) layer. The fully connected layer just uses the image's "compressed" representation to conduct categorization and tries to fit a straightforward NN (multilayer perceptron). Convolutional layers, which are used to process images, consist of a learnable filter with a predetermined size (the kernel). ReLU, often referred to as rectified linear units, is frequently used in CNNs.

4 Result Analysis

True positive (TP) describes a situation when the test is positive and someone can recognize the illness. The situation known as a true negative (TN) occurs when a test is negative and a person is not given a disease diagnosis. False positive (FP) is the circumstance in which a test result is positive but the subject is unable to communicate it. False negative (FN) results occur when a person can actually experience a negative outcome. Naïve Bayes achieves 70%, SVM 68%, While KNN classifier yields accuracy of 76%. The deep learning model CNN attains accuracy of 93.09% by training on a dataset of size 2562 for 30 epochs with 81 observations in each epoch. The model accuracy and model loss can be viewed in Figs. 6 and 7.

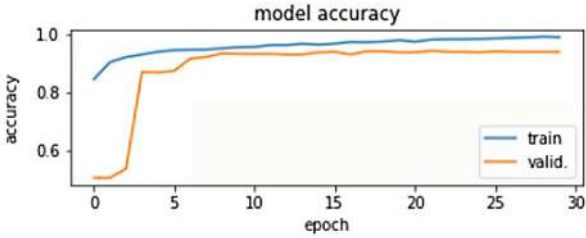


Fig. 6 Model accuracy

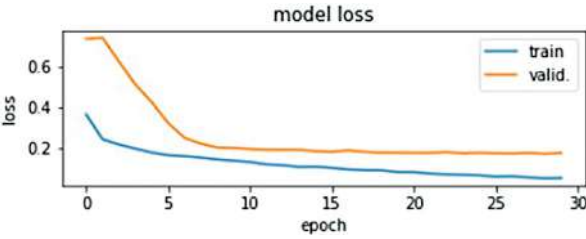


Fig. 7 Model loss

5 Conclusion and Future Scope

Using the current approach, humans can alleviate many of the challenges presented by traditional detection systems. This system has a significant influence and may decrease the likelihood of lags and other inefficiencies. Automated screening methods dramatically reduce the time needed for diagnosis, saving opticians time and money while allowing patients to receive treatment faster. Automated DR detection systems are essential for the early detection of DR. The stage of DR is determined by the type of lesions that occur on the retina. This chapter reviews the latest deep learning-based methods for the automatic detection and classification of diabetic retinopathy. Com mon publicly available fundus DR datasets are described, and deep learning techniques are briefly explained. Due to its efficiency, most researchers have used CNNs for the classification and detection of DR images. This research aims to propose an optimal model for the detection of diabetic retinopathy. The processing of retinopathy images is very important to obtain the appropriate features. The statistic correctly predicts severity, but in the case of noisy images, the risk of getting bad data leads to lower accuracy. It is also important to select the correct image features in order to produce accurate results. Using all available channels instead of just one helps the network learn more features and reduces overfitting by making the data more complex using alternative methods of noiseless image preparation.

References

1. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting diabetes mellitus with machine learning techniques," *Frontiers in Genetics*, vol. 9, 2018.
2. T. J. MacGillivray, J. R. Cameron, Q. Zhang et al., "Suitability of UK biobank retinal images for automatic analysis of morphometric properties of the vasculature," *PLoS One*, vol. 10, no. 5, Article ID e0127914, 2015.
3. P. Costa, A. Galdran, M. I. Meyer et al., "End-to-End adversarial retinal image synthesis," *IEEE Transactions on Medical Imaging*, vol. 37, no. 3, pp. 781–791, 2018.
4. P. H. Scanlon, C. Foy, R. Malhotra, and S. J. Aldington, "The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening," *Diabetes Care*, vol. 28, no. 10, pp. 2448–2453, 2005.
5. Bálint Antal, András Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Elsevier*, Volume 60, April 2014, Pages 20–27.
6. Carrera, Enrique V., Andrés González, and Ricardo Carrera. "Automated detection of diabetic retinopathy using SVM." *2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON)*. IEEE, 2017.
7. Kumar, Shailesh, and Basant Kumar. "Diabetic retinopathy detection by extracting area and number of micro aneurysm from colour fundus image." *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2018.
8. Rahim, Sarni Suhaila, et al. "Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening." *Neural computing and applications* 27 (2016): 1149–1164.
9. Silva, Caroline, Thierry Bouwmans, and Carl Frélicot. "An extended center-symmetric local binary pattern for background modeling and subtraction in videos." *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISAPP 2015*. 2015.
10. Gharaibeh, Nasr, et al. "An effective image processing method for detection of diabetic retinopathy diseases from retinal fundus images." *International Journal of Signal and Imaging Systems Engineering* 11.4 (2018): 206–216.
11. Fujita, Hiroshi, et al. "Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs." *Computer methods and programs in biomedicine* 92.3 (2008): 238–248.
12. Stolte, Skylar, and Ruogu Fang. "A survey on medical image analysis in diabetic retinopathy." *Medical image analysis* 64 (2020): 101742.
13. Attia, Abdelouahab, Zahid Akhtar, and Sofiane Maza SamirAkhrouf. "A survey on machine and deep learning for detection of diabetic Retinopathy." *ICTACT Journal on Image and Video Processing* 11.2 (2020): 2337–2344.
14. Asiri, Norah, et al. "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey." *Artificial intelligence in medicine* 99 (2019): 101701.
15. Valarmathi, S., and R. Vijayabhanu. "A survey on diabetic retinopathy disease detection and classification using deep learning techniques." *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE, 2021.
16. Shamshad, Fahad, et al. "Transformers in medical imaging: A survey." *arXiv preprint arXiv:2201.09873* (2022)
17. Sundari, M. Shanmuga, Vijaya Chandra Jadala, and Sai Kiran Pasupuleti. "Prediction of Activity pattern mining for Neurological disease using Convolution Neural Network." *2022 7th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2022

Smart Health Prediction Using Random Forest



Kooragayala Sukeerthi, Kurma Pooja Reddy, Shaik Thasleema,
and Paindla Sowjanya

Abstract This chapter's main goal is to discuss the application of information mining to clinical medical services. These information mining strategies can likewise be applied to research and schooling in various areas. The Brilliant Wellbeing Forecast Framework is the most up-to-date part of clinical science to arise. Information mining is a branch of software engineering that makes use of data already available in the healthcare industry to predict illness occurrence. We can separate new examples from gigantic datasets and gain data by utilizing AI and dataset administration procedures. The accompanying review directs a study on how information mining procedures and AI are consolidated to gauge sicknesses in view of client-side effects. The building of a model that can predict illnesses based on user symptoms is very helpful in delivering patients with prompt and effective medical treatment. In the field of medicine, several machine learning algorithms are utilized to forecast various diseases and assist clinicians in making quick diagnoses. Many lives can be saved by brief information examination and precise sickness forecast in light of side effects. Early illness identification helps specialists in endorsing exact medication.

Keywords Smart health · Random forest · Root Tkinter · Machine learning · Data mining

1 Introduction

Data mining is the process of extracting knowledge from unknown or ineffective datasets. The data is processed and converted into meaningful information using a variety of data mining techniques. Data mining can be applied to a variety of sectors, including business analysis, healthcare, and stock management. Data

K. Sukeerthi (✉) · K. P. Reddy · S. Thasleema · P. Sowjanya
Vignan Institute of Technology and Science, Deshmuki Village, Yadadri Bhuvanagiri District,
Telangana, India

mining techniques can be used to process a large amount of data in the medical profession.

It may have occurred to you or someone close to you in the past that you required immediate medical assistance but were unable to locate one. Developing a model that can anticipate diseases based on user symptoms is extremely beneficial in providing patients with quick and appropriate medical care. Many lives can be saved by prompt data analysis and accurate disease prediction based on symptoms. Early disease detection aids doctors in prescribing accurate medicine.

In the field of medicine, several machine learning algorithms are utilized to forecast various diseases and assist clinicians in making quick diagnoses. The accuracy of the results may vary depending on the data input. Smart health prediction using random forest is a machine learning approach that leverages the power of random forest algorithm to predict the health outcomes of individuals. This approach is based on a decision tree model that is constructed using a set of predictor variables that are known to be associated with specific health outcomes.

The random forest algorithm generates multiple decision trees and combines them to make a more accurate and robust prediction. It is an ensemble learning technique that utilizes bootstrap aggregating (bagging) to train multiple decision trees on different subsets of the data. The predictions from these individual decision trees are then combined to produce a final prediction.

Smart health prediction using random forest has various applications in health-care. It can be used to predict the likelihood of a particular disease or condition, to identify patients who are at high risk of developing a particular disease, or to predict the effectiveness of a particular treatment.

The approach is becoming increasingly popular due to its ability to handle large and complex datasets, and its ability to handle missing data. It is a valuable tool in the field of precision medicine, where personalized treatments are tailored to individual patients based on their genetic, environmental, and lifestyle factors.

2 Literature Survey

Vembandasamy et al. [1] have proposed heart contaminations recognizable proof using simple Bayes computation. They used Naïve Bayes computation and predicted heart diseases.

Clinical benefits are an unavoidable task in human existence. In the broad field of clinical science, prosperity through business and integrating considerations into business has access to enormous amounts of data and confidential information. By using data mining technologies, wise judgments are produced with this constrained knowledge. A couple of tests are done in the acknowledgment of cardiovascular disorders in the patient; however, with data mining, these tests could be decreased. However, there is a shortfall in examining at instruments to give feasible exploratory results the restricted information, so a structure is made using data digging estimations for requesting the data and to recognize the heart infections. Data

mining is probably a response to far most clinical consideration issues. Naïve Bayes computation is one such data mining procedure that serves in the finish of heart disorder patients. The heart disorders figure system (HDPS) is suggested in this work, which also separates two or three limitations and forecasts heart problems.

Wilson Wibamanto, Debashish Das, Sivananthan A/L Chelliah et al. [2] have proposed a splendid prosperity assumption using data mining. They used Naïve Bayes computation and predicted heart diseases. It may have worked out frequently enough that you or a family member require professional assistance right now, but they are not available due to a clarification. The Prosperity Gauge framework is a project that offers online guidance and end-client assistance. Here, we provide a method that enables users to seek second opinions on their medical problems online using a clear clinical benefits structure. Different secondary effects and infections/diseases linked to different systems are dealt with by the structure. The system allows users to communicate their results and problems. The next step is processing the client's incidental effects to look for potential links to various illnesses. Here, we determine the most reliable illness that may be linked to the patient's side effects using a few clever data mining approaches. If the system can't give proper results, it teaches the client about the sort in regards to disorder or unrest it feels client's after effects are connected with. If client's after exactly match no ailment in our informational collection, it shows the disorders client could potentially have based on their secondary effects. It similarly contains expert location, contacts close by Analysis, and manager dashboard for system exercises.

Dr. R. Thirumalaiselvi, S. Dilli Arasu et al. [3] have investigated consistent kidney disease based on data mining strategies. Chronic Kidney infection is a significant clinical issue affecting people overall. Steady disorders lead to depressingness and augmentation of death rates in India and other low and focus pay countries. The consistent afflictions record to around 60% of all passings all over the planet. Around 80% of deaths due to disease often take place in developing nations with poor wages. In India, the number of deaths attributed to ongoing disorder was most likely 5.21 million in 2008 and is expected to rise to 7.63 million in 2020, or roughly 66.7%. The process of extracting covered information from enormous databases is known as data mining. To predict renal problems, important data mining techniques such as packing, plan, alliance evaluation, backslide, time series, and progression exams were applied. Minor obstacles to preprocessing or at various stages were present in the methods that have been implemented so far. The various data mining techniques are investigated in this research to predict renal illnesses and to quickly resolve challenging challenges.

3 Existing System

Kidney disease is a challenging problem in under developed/developing countries. Around the world, 60% of fatalities are caused by kidney disease. Kidney disease can frequently lead to other chronic conditions including hypertension, diabetes, sickness, weak bones, and nerve damage. Using CNN makes it impossible to detect

medical care frauds and abuses. It is unable to generate an examination report quickly, is declining in functional competency, and has large functional expenses.

4 Proposed System

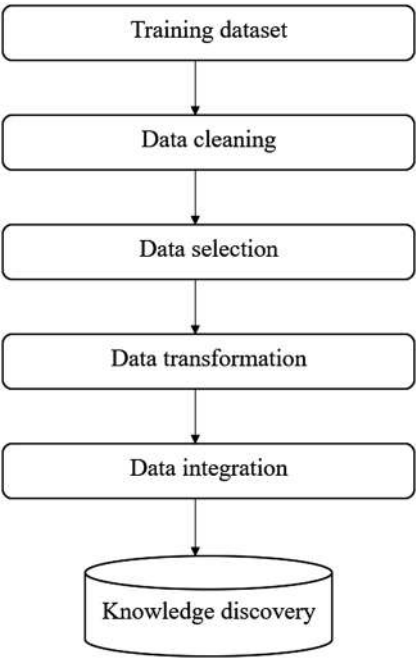
The most well-known machine learning classifiers employed in the proposed system include

- 1. KNN(K-Nearest Neighbor).
- 2. Decision Tree.
- 3. Random Forest.
- 4. GaussianNB.

The below Fig. 1 represents the preprocessing activity of the dataset. If the datasets offer suitable training, random forest has good accuracy.

5 System Architecture

Fig. 1 System architecture of data mining techniques to obtain clean data



6 Methodology

6.1 *Data Mining Techniques*

6.1.1 Data Gathering

This paper's information assortment comprises various records. The determination of the subset of all open information that you will be working with is the focal point of this stage. Preferably, ML challenges start with a lot of information (models or perceptions) for which you definitely know the ideal arrangement. Marked information will be data for which you are as of now mindful of the ideal result.

6.1.2 Preprocessing of the Data

Format, clean, and sample from your chosen data to organize it.

There are three typical steps in data preprocessing:

1. Designing
2. Information cleaning
3. Inspecting

Designing: It's conceivable that the information you've picked isn't in a structure that you can use to work with it. The information might be in an exclusive record configuration, and you would like it in a social data set or text document, or the information might be in a social data set and you would like it in a level document.

Information cleaning is the most common way of eliminating or supplanting missing information. There can be information examples that are inadequate and come up short on data you assume you really want to resolve the issue. These events could be eliminated. Moreover, a portion of the traits might contain delicate data, and it very well might be important or totally eliminate these properties from the information.

Inspecting: You might approach significantly more painstakingly picked information than you want. Calculations might take significantly longer to perform on greater measures of information, and their computational and memory prerequisites may likewise increment. Prior to considering the whole dataset, you can take a more modest delegate test of the picked information that might be fundamentally quicker for investigating and creating thoughts.

6.1.3 Feature Extraction

The next step is to include a quality reduction course in the extraction process. Highlighted extraction really modifies the traits instead of element choice, which positions the ongoing ascribes as indicated by their prescient pertinence. The first

ascribes are straightly joined to create the changed traits, or elements. Finally, the classifier calculation is utilized to prepare our models. We utilize the Python Normal Language Tool stash's classify module. We utilize the gained marked dataset. The models will be surveyed utilizing the excess marked information we have.

Pre-handled information was ordered utilizing a couple of AI strategies. Irregular woodland classifiers were chosen. These calculations are generally utilized in positions including text grouping.

6.1.4 Evaluation Model

The method involved in developing a model includes assessment. Finding the model that best portrays our information and predicts how well the model will act in what's to come is useful. In information science, it isn't adequate to assess model execution utilizing the preparation information since this can rapidly prompt excessively hopeful and overfitted models. Wait and Cross-Approval are two procedures utilized in information science to evaluate models.

The two methodologies utilize a test set (concealed by the model) to survey model execution to forestall overfitting. In light of its normal, every classification model's presentation is assessed. The result will take on the structure that was envisioned in the diagram portrayal of the ordered information.

6.1.5 User Interface

Tkinter is the inbuilt python module that is utilized to make GUI applications. It is one of the most regularly involved modules for making GUI applications in Python as it is basic and simple to work with. You don't have to stress over the establishment of the Tkinter module independently as it accompanies Python as of now. It gives an article-arranged connection point to the Tk GUI tool stash.

Some other Python Libraries accessible for making our own GUI applications are

1. Kivy
2. Python Qt
3. wxPython

Among all, Tkinter is generally broadly utilized.

Here are some normal use cases for Tkinter:

1. Making windows and exchange boxes: Tkinter can be utilized to make windows and discourse boxes that permit clients to cooperate with your program. These can be utilized to show data, accumulate information, or present choices to the client.

2. Building a GUI for a work area application: Tkinter can be utilized to make the connection point for a work area application, including buttons, menus, and other intuitive components.
3. Adding a GUI to an order line program: Tkinter can be utilized to add a GUI to an order line program, making it more straightforward for clients to communicate with the program and information contentions.
4. Making custom gadgets: Tkinter incorporates various underlying gadgets, like buttons, names, and text boxes, yet it additionally permits you to make your own custom gadgets.
5. Prototyping a GUI: Tkinter can be utilized to rapidly model a GUI, permitting you to test and emphasize on various plan thoughts prior to focusing on a last execution.

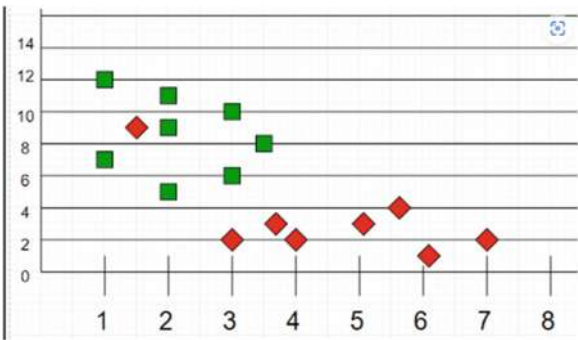
6.2 *Classifiers*

6.2.1 **Random Forest**

An AI technique called Irregular Woodland is outfit-based and operated. You can combine several computation types to create a more convincing forecast model, or use a similar learning technique at least a few times. The phrase “Irregular Timberland” refers to how the arbitrary woodland method combines a few calculations of the same type or various chosen trees into a forest of trees. The irregular timberland technique may be used for both relapse and characterization tasks. Coming up next are the essential stages expected to execute the irregular woods calculation.

1. Pick N records aimlessly from the datasets.
2. Utilize these N records to make a choice tree.
3. Select the number of trees you need to remember for your calculation, then, at that point, rehash stages 1 and 2.
4. Every tree in the forest foretells the categorization in which the new record will be included in the order of things. The classification that gets most of the votes is at last given the new record.
5. The advantages of Irregular Woodland.
6. The way that there are numerous trees and they are completely prepared utilizing various subsets of information guarantees that the irregular timberland strategy isn't one-sided.
7. The irregular woods strategy fundamentally relies upon the strength of “the group,” which reduces the framework's general predisposition. Since it is extremely challenging for new information to influence every one of the trees, regardless of whether another information point is added to the datasets, the general calculation isn't highly different.
8. In circumstances when there are both downright and mathematical highlights, the irregular woods approach performs well.

Fig. 2 KNN classification which is obtained during the application of KNN classifier



9. At the point when information needs esteems or has not been scaled, the irregular woodland method likewise performs well.

6.2.2 KNN

One of the most important yet fundamental representation estimations in artificial intelligence is K-Nearest Neighbors. Plan affirmation, data mining, and interference ID are three areas where it has significant applications and fits into the supervised learning field.

Since it is nonparametric as opposed to some calculations, like GMM, which assume a Gaussian movement of the provided data it makes no fundamental hypotheses about the dispersion of data, making it widely disposable in light of any conditions.

We are given some initial data (also called training data), which organizes the data points into groups identified by a specific attribute. The Fig. 2 showing KNN classification diagram. For example, consider the following table of data points containing two features.

6.2.3 Decision Tree

When calculating the class of a particular dataset in a decision tree, the process begins at the root hub and works its way up gradually. Considering the benefits of the root property and the record (genuine dataset) quality, this computation jumps to the companion hub and then follows the branch.

The calculation checks the quality worth with the other sub-hubs by and by for the accompanying hub prior to proceeding. It continues to do this until it arrives at the tree’s leaf hub. The whole calculation can be isolated into the accompanying strides for better association:

Stage 1: Begin the tree from the root hub, which contains all of datasets, prompts S.

Stage 2: Utilize the Trait Determination Measure to distinguish the dataset's best characteristic (ASM).

Stage 3: Subset the S into bunches that incorporate expected values for the best attributes.

Stage 4: Make the choice tree hub that has the best trait in sync four.

Utilize the subsets of the datasets made in sync 3 to iteratively make new choice trees in Sync 5

Stage 5: Proceed in this way until you reach a point where you can no longer classify the hubs, and at that moment, refer to the final hub as a leaf node.

6.2.4 Naive Bayes

A variation of depiction estimations taking Bayes' Theory into account is called Naive Bayes classifiers.

Each set of features being gathered is freed from the others, thus it's everything but a single estimation but a collection of computations where they all share a common standard. Take datasets into consideration in particular. Think of a hypothetical dataset that shows the weather conditions for a round of golf.

Each tuple groups the atmospheric conditions as suitable ("Yes") or unsuitable ("No") for playing golf.

Mathematically, Bayes' hypothesis is expressed as the going-with condition.

$$P\left(\frac{A}{B}\right) = \left\{ \frac{P\left(\frac{B}{A}\right) P(A)}{P(B)} \right\} \quad (1)$$

where A and B are occasions and $P(B) \neq 0$.

1. In essence, we are looking for the probability of occurrence A, assuming that occasion B is true. Additionally cited as proof is Occasion B.
2. $P(A)$ represents the priori of A (the earlier probability, such as the likelihood of an event occurring before evidence is observed). The proof is a quality that an obscure instance (in this case, event B) is valuable.
3. $P(A|B)$ denotes the probability of B, such as the probability of an event occurring once evidence is observed.

Currently, we may use Bayes' hypothesis in the manner described below in relation to our dataset:

$$p\left(\frac{y}{X}\right) = \text{frac} \left\{ p\left(\frac{X}{y}\right) p(y) \right\} \{p(X)\} \quad (2)$$

where X is a reliant component vector (of size n) and y is a class variable

$$X = (x_1, x_2, x_3, \dots, x_n)$$

Exactness: The level of precise expectations for the test information is implied by exactness. By separating the quantity of exact forecasts by the all out number of expectations, it very well might not be entirely set in stone.

7 Result Analysis

The outcomes are acquired by the examination of various calculations in the medical care expectation. The significant calculations incorporate Help vector machines, Brain organizations, strategic relapse, Irregular woodland, and so forth. Among this, the exactness is high in the brain organizations assuming appropriate preparation is given by the datasets. The calculations that are utilized in medical services in various fields of medication alongside their precision are talked about. This provides us with an overall thought of the level of calculations which are utilized in various sickness forecast. The Fig. 3 shows the comparative results of the proposed algorithms and accu-racies.

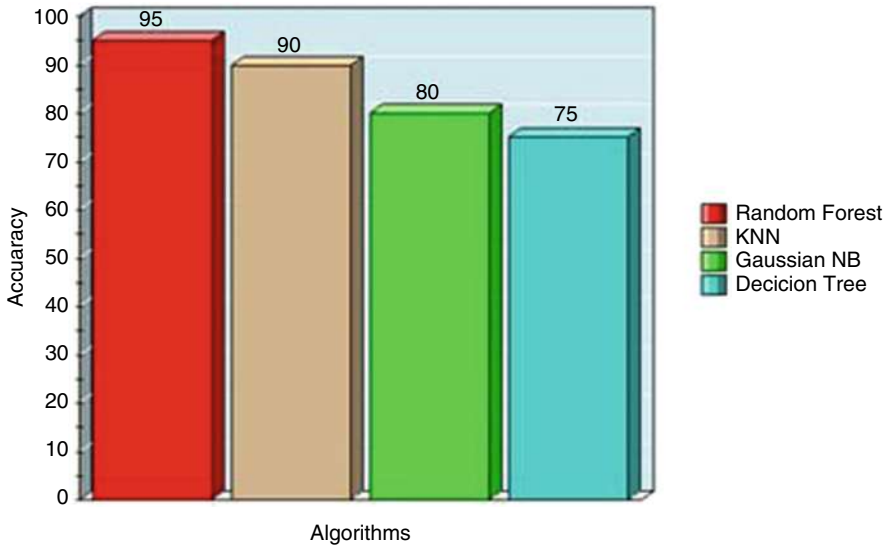


Fig. 3 Algorithm’s accuracy based on the results

8 Conclusion

Information mining has the most prominent significance in the space of clinical and specialized sciences. Clinical science can achieve some amazing feats with the aid of information mining and AI computation. Specialists can easily locate the illness and promptly administer medicine. Different infection stages can be carefully identified and treated in accordance with the needs of the patients. To make precise decisions, the data obtained through information mining might be valuable. Information mining will become significantly more advanced and capable of extracting distinct information hidden in healthcare data as IT advances.

References

1. Nikita Kamble, International Journal of Scientific Research in Computer Science Engineering and Information Technology, Vol. 2, Issue 5, 2017, "Smart Health Prediction System Using Data Mining".
2. Prof. Krishna Kumar Tripathi, International Research Journal of Engineering and Technology (IRJET), Vol.5 Issue:4, Apr-2018, "A Smart Health Prediction Using Data Mining".
3. G. Pooja Reddy, International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol-8 Issue-6, April 2019, "Smart E-Health Prediction System Using Data Mining".

Vulnerability Assessment and Penetration Testing Using Parrot Operating System



Avvaru R. V. Naga Suneetha, J. K. V. Prasanthi, S. Nimisha,
and C. H. Meghna

Abstract Ethical hackers attempt to compromise the system on their own while complying by the law in order to assist the business in preventing malicious attacks. Employing the defense-in-depth strategy can help businesses ensure that their information assets are properly safeguarded. Through vulnerability assessments and penetration testing (VAPT), these vulnerabilities are frequently addressed by security professionals from around the world. In order to identify vulnerabilities, the target system's security posture is assessed using a variety of automated tools as well as manual testing methods including Nmap, netdiscover, and OpenVAS Greenbone. Penetration testing simulates the actions of a malicious attacker trying to take advantage of the target system's vulnerabilities. Metasploit is one of the Pentest's tools. The input vector for this stage is the identified set of vulnerabilities from the vulnerability assessment. The efficiency of the security mechanisms present in the target system is evaluated through the VAPT procedure.

Keywords Nmap · Netdiscover · Metasploit · VAPT · Defense-in-depth

1 Introduction

In response to the rise in cyberattacks and the sophistication of malware and hacking tactics, organizations are using VAPT method of discovering and controlling security concerns. VAPT, also known as penetration testing, employs a “hands-on” approach to verify an IT system's overall security by acting out a hacking assault. VAPT testing can be used to evaluate a system's weaknesses and provide a thorough report on how a hacker can circumvent the system's present security safeguards. Finding all software defects that could be used by hackers is the goal of a VAPT audit. VAPT security audits are conducted in accordance with a systematic process

A. R. V. Naga Suneetha (✉) · J. K. V. Prasanthi · S. Nimisha · C. H. Meghna
Vignan Institute of Technology and Science, Deshmuki Village, Yadadri Bhuvanagiri District,
Telangana, India

that makes use of several tools, techniques, and methodologies. It is a method for figuring out whether the company is secure from outside threats, to put it succinctly. Today, there is a lot of talk about hacking issues and cyberattacks. Our computers and networks need to be secured by everyone. Through penetration testing and vulnerability assessments, you can get knowledge of attacks, security weaknesses, and how to fix them. The security posture of the target system is assessed to find vulnerabilities using a variety of automated tools and manual testing methods. A malicious attacker aiming to exploit the weaknesses of the target system is simulated during penetration testing. The determined collection of vulnerabilities from the vulnerability assessment serves as the input vector for this phase. The VAPT technique helps assess the effectiveness of the target system’s security controls. This uses the Parrot OS, a Linux-based operating system. The development of computing technologies has enabled greater success in both the professional and personal realms. Cybersecurity strategies are becoming more and more crucial as a means of securing people’s or organization’s online presence as people rely more on electronic devices like computers, smartphones, and other gadgets. Cybercriminals employ a variety of strategies, such as spreading malware, to gain unauthorized access and collect the data exchanged between clients and server, which is typically in plain text formats. Organizations must be informed about the different cyber-attacks in order to stop them.

Over the Kali Linux, Parrot OS is chosen because Parrot OS may be viewed as a handheld lab for a wide range of cyber security management tasks ranging from pen testing to reverse programming and digital forensics. It also provides everything you need to safeguard your data and build applications. ParrotOS clearly outperforms Kali Linux in terms of hardware needs. It not only requires less RAM to work effectively, but there are already numerous built-in utilities in the overview provided below [10] (Table 1).

Objective and Motivation

- The Parrot operating system, a Linux-based Debian system, plays a significant part in this study. In Parrot OS, there are several pre-built Open-source utilities.

Table 1 Comparison of tools between two operating systems [11]

Kali Linux	Parrot OS
Kali-Linux is a massive OS.	Parrot OS is a portable operating system.
It fails to include any built-in compilers.	It comprises a broad spectrum of preinstalled compilers and IDEs.
This Only facilitates penetration testing	A full development stack that includes the best editors, languages, and tools. There are additional multimedia bundles available.
It includes all tools used for hacking.	It includes all of the tools found in Kali. Anonsurf, Wifiphisher, and lots more

- There are three steps involved in the Vulnerability Assessment and Penetration Testing (VAPT) process: information gathering, vulnerability assessment, and penetration testing.
- The most efficient tools for information gathering include Nmap, Angry IP Scanner, and SNMP-Check, which let pen testers to learn about the target's operating system, IP and MAC addresses, ports, and services.
- To protect the network, use Nmap Enumeration Testing to pass and detect the firewall and create bespoke packets outside the firewall.
- OpenVAS Greenbone and Nmap-Vulnscan are utilized in vulnerability testing because they produce accurate vulnerability data.
- OpenVAS Greenbone generates a full vulnerability report for the selected host, as well as a recovery solution and Nmap-Vulnscan, on the other hand, analyses CVE scripts and reports on vulnerabilities that may be fixed.
- Finally, the pen tester joins the system via Metasploit, which lets the user to access the target console while maintaining organizational security.

2 Literature Survey

There is always a flip side to technology, as is the case with criminal hackers [1]. So, in this essay, Shanmugapriya concluded that hacking is currently a problem without a solution. The only way to stop a hacker is to become knowledgeable about hacking. To be safe, everyone should be aware of the hacking ethics and adhere to them. The method that can be utilized to find the undiscovered gap in the network is penetration testing. An overview of penetration testing and its tools is provided by Singh and Singh et al. in this work [2]. They have published a lot of literature and described the various stages of penetration testing as follows: Phase 1: Reconnaissance Second stage: scanning Phase 3: Getting Access Phase 4: Keeping Access Covering Tracks is the fifth phase. A tool that has proven to be particularly effective in network sniffing is Wireshark. The tool's ability to easily analyze the packets it pulls from a live network allows it to determine how vulnerable a system is to security vulnerabilities [3]. In this study, Sandhya S et al. constructed a website and used the pen-testing program Wireshark to sniff user data from the website and analyze it for security flaws. FilipHolik et al. have briefly described the fundamentals of penetration testing in this paper and demonstrated how to use the Metasploit framework and deploy it when conducting penetration tests [4]. The tools and aids that can be used to improve the workflow of penetration test methods are described in this study. Vulnerability Assessment and Penetration Testing: A Practical Approach by Mohit Arora offers an overview of VAPT and the procedures involved in the process [5]. The study emphasizes the necessity of VAPT for organizations and examines the advantages of performing frequent VAPT evaluations. A methodology for doing VAPT assessments is proposed in "A Framework for Conducting Penetration Testing and Vulnerability Assessment" by Vijayakumar Balasubramanian, Arun Kumar Srinivasan, and Shriram K.

Vasudevan [6]. The authors walk through the VAPT procedure step by step and explore the problems that organizations may experience while performing VAPT evaluations. Alzaabi and Alghathbar et al.'s study "A Review of Tools and Techniques for Vulnerability Assessment and Penetration Testing" gives an overview of the tools and techniques utilized in VAPT evaluations [7]. The study examines three common VAPT technologies and evaluates their benefits and drawbacks. Geer and Harthorne et al. characterized penetration testing as the art of locating an open door in their publication. It is not a science since science requires falsifiable hypotheses, as well as hackers as artists [8]. This article discusses characterization and specialization. Successful penetrations are defined as the unlawful acquisition of legitimate power. Shinde and Ardhapurkar et al. proposed a study that aids in the development of a flexible technique capable of detecting vulnerabilities in online applications employing VAPT [9]. This document illustrates the distinction between vulnerability assessment and penetration testing. Vulnerability Assessment and Penetration Testing (VAPT) are two methods of vulnerability testing that are frequently coupled to improve vulnerability analysis findings.

3 Existing System

A vulnerability examination thoroughly assesses the security flaws of a data system. It examines if the system is prone to any known flaws, assesses how destructive they are, and, if required, proposes patches or mitigations. There are several different types of vulnerability analyses. These include host assessment, network and wireless evaluation, database evaluation, and application scans. A pen test simulates an approved assault on a network in order to examine its security. Penetration testers utilize the same tools, methodologies, and procedures as attackers to identify and illustrate the financial repercussions of a system's weaknesses.

4 Proposed System

With the help of the agentless, Linux-based Parrot operating system, we conduct testing on the proposed system's Vulnerability Assessment and Penetration testing. Instead of reaching out to the assets from the server, agentless scanning does not need agents to be deployed on each device. Although the data is identical to that obtained using an agentless technique, no additional software needs to be installed or managed on any of the devices. This way of searching for security flaws is extremely simple and lightweight.

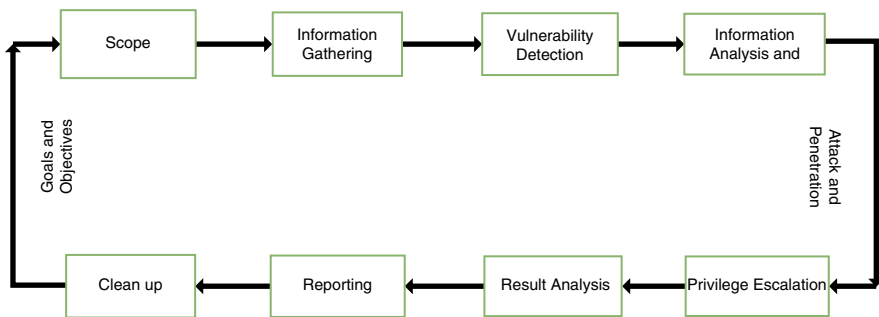


Fig. 1 System architecture

5 System Architecture (Fig. 1)

6 Methodology

- Gather the System Information
- Check for Vulnerabilities
- Penetrate the Vulnerable System
- Fix Vulnerabilities

6.1 *Gather the System Information*

It is the process of learning more about a potential target. This could be carried out as part of network security monitoring, penetration testing, or other cybersecurity tasks.

6.2 *Check for Vulnerabilities*

It assesses the significance of any known vulnerabilities, evaluates whether the system is vulnerable to them, and offers suggestions for mitigation or repair.

6.3 *Penetrate the Vulnerable System*

A pentest is a real-world simulated attack made against a computer system to assess its security.

6.4 Fix Vulnerabilities

The operating system can be updated, vulnerable ports can be blocked, etc., to address vulnerabilities that have been found.

Proposed Approach

Assessment of Vulnerabilities and Penetration Testing

Two types of vulnerability testing include penetration testing and vulnerability assessment. The tests are frequently combined because of their various strengths to produce extensive vulnerability analysis. Investigations of penetration and evaluations of vulnerabilities serve a pair of separate purposes. throughout a similar region of attention, frequently with variable results.

Risk assessment tools discover existing vulnerabilities but do not distinguish between usable faults. Vulnerability scanners tell businesses of the locations of active code issues. Penetration testing tries to take advantage of a system's vulnerabilities in order to determine whether it is possible to get unauthorized access or engage in other harmful conduct as well as to determine whether errors endanger the application. Exploitable weaknesses are discovered and their severity is evaluated during penetration tests. The goal of a test of vulnerability is not to identify every security fault in a system, but rather to demonstrate how severe a flaw might be in a real-world attack. Vulnerability assessment and penetration testing tools provide information on where and how many bugs are currently present in the code. Penetration testing tries to take advantage of a system's vulnerabilities in order to determine whether it is possible to get unauthorized access or engage in other harmful conduct as well as to determine whether errors endanger the application. Exploitable defects are detected and their severity is assessed during penetration tests. In order to provide a comprehensive picture of the risks and vulnerabilities existing in an application, VAPT tools collaborate.

7 Tools

7.1 Information Gathering Tools

Nmap

In order to scan hosts and services on a computer network, it sends packets and examines the response. With the use of Nmap, it can identify and get over the firewall as well as build custom packets to scan past the IDS/Firewall, which is also useful for enumeration testing. Some instructions are employed in order to obtain data. For instance;

To learn about the service/version, use the **-sV**- probe command on open ports.

Treats every host as being online.

-O-Enable OS detection. Scanning the OS of the systems in a network.

-A- It gives the complete information about the OS, MAC address and Versions.

Angry IP Scanner

Angry IP Scanner is a quick port and IP address scanner. It can scan any port as well as any range of IP addresses. It weighs little. It just performs a ping to see if each IP address is active, and then, if desired, resolves the hostname, finds the MAC address, and inspects ports. After downloading and installing the Angry IP scanner, launch the application and enter the IP range, including start and finish addresses, as well as the hostname and Netmask. Now, by pressing the start button, the IP address of the specified range is pinged. Some IP addresses are represented by the colors red, green, and blue. Whereas the Red color denotes that a port is closed, the Blue color denotes that a port is unavailable, and the Green color denotes that a port is open and active.

SNMP Check

In Internet Protocol networks, network-connected devices are managed and monitored using the Simple Network Management Protocol (SNMP). List the SNMP devices using SNMP-Check, and it outputs the results in a very readable and helpful way. It might be helpful for system monitoring or penetration testing. With the help of the command `snmp-check 192.168.1.1 -c public`, we may obtain all network data. Before configure SNMP using a switch before working on SNMP-Check.

Detection of Firewall

To identify filtering on ports, Nmap offers a streamlined method for identifying firewall filtering. To check the state of the filtering, this function can be used to test one port or several ports in succession. Calling Nmap with the IP address and destination port specified will do a Nmap firewall scan.

Bypass Firewall

While diagramming firewall rules can be useful, frequently the main objective is to get around restrictions. Nmap employs a variety of strategies for this, albeit the majority are only useful on networks with poor configuration. A firewall can't actually be gotten through. Every piece of incoming traffic is filtered in accordance with the preset rules. We are not permitted to access ports, services, or websites because the firewall might be blocking them. Therefore, the firewall must be got through in order to access them.

Create Custom Packets to Scan Beyond Firewall

The security tools designed to stop an unauthorized user from entering a network include an IDS (Intrusion Detection System) and firewall. Even IDS and firewalls have certain security flaws, though. IDS/firewalls can be bypassed while still transmitting intended packets to the target, though. Unwanted traffic entering a network is prevented by firewalls and IDS.

Enumeration Testing

The technique a penetration tester uses to find data about assets that are inside the scope is enumeration. An automated procedure is used by a pen tester to locate all

active IP addresses within the defined area as well as some basic details about these machines, such as their type and OS version. Enumeration is still a crucial tool for methodically scanning a target for data. By locating open ports, usernames, and passwords, enumeration can give hackers a step-by-step guide for breaking into a system.

7.2 Vulnerability Analysis

Nmap-Vulns

Nmap-vulns is among the most well-known vulnerability scanners currently in use. The three most well-known and often utilized CVE detection scripts in the Nmap search engine are vuln scan, vuln, and Nmap-vulners. We can learn vital information regarding system security holes thanks to these programs. Data security issues are listed in the Common Vulnerabilities and Exposures (CVE) database as being known to the public. It is a paradigm for locating threats to the security of information systems. It was discovered that data security had flaws. It serves as a reference model for locating risks and holes in the security of information systems.

OpenVAS Greenbone

OpenVAS is a strong vulnerability scanner. It supports a large range of high-level and low-level industrial and Internet protocols, optimizes speed for thorough scanning, and includes a sophisticated internal programming language. The stream, which has a lengthy history and receives daily updates, provides the scanner with tests for locating vulnerabilities.

7.3 Tools for Pentesting

Metasploit

It gives testers the ability to launch exploits, conduct reconnaissance, scan systems for vulnerabilities, and more. Ethical hackers can investigate systematically vulnerable servers and networks using the sophisticated Metasploit platform. It can be easily modified and used with the majority of operating systems because it is an open-source framework.

Operating System Used

Based on Debian, Parrot OS is a Linux distribution with an emphasis on security, privacy, and development. Debian, the most cutting-edge and well-known universal operating system that can operate everywhere, is the foundation upon which Parrot is built. The Parrot core may be used to power everything from your laptop to your phone, including servers, IoT boards, cloud containers (like docker), and more. It has a complete mobile arsenal for digital forensics and IT security operations. It

8.2 Vulnerability Analysis

OpenVAS Greenbone (Fig. 3)

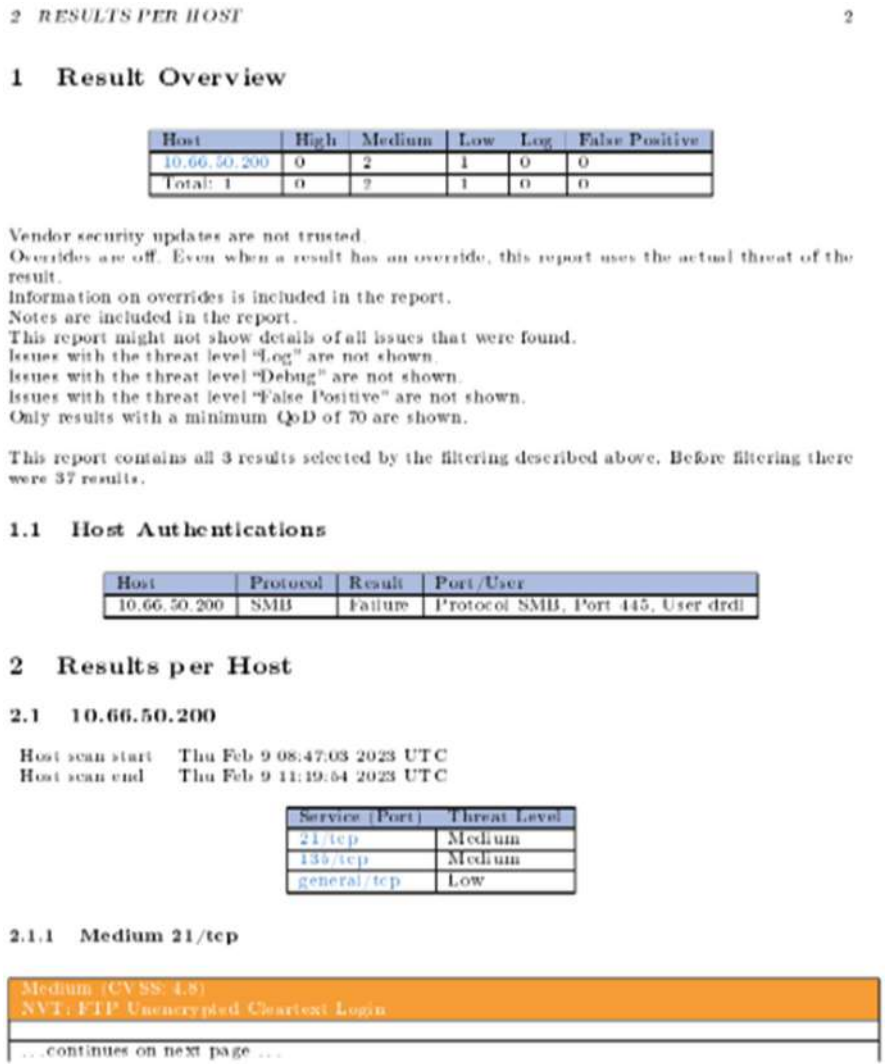


Fig. 3 Generated vulnerability report of the host

```
(Meterpreter 1)(C:\Windows\system32) > sysinfo
Computer      : USER1-PC
OS            : Windows 7 (6.1 Build 7601, Service Pack 1).
Architecture  : x64
System Language : en_US
Domain        : WORKGROUP
Logged On Users : 2
Meterpreter   : x64/windows
(Meterpreter 1)(C:\Windows\system32) > █
```

Fig. 4 Target system information

8.3 Penetration Testing

Metasploit (Fig. 4)

9 Conclusion

Based on Debian, Parrot OS is a Linux distribution with an emphasis on privacy, security, and development. Debian, the most cutting-edge and well-known world-wide operating system currently available, serves as the foundation for Parrot's architecture. The Parrot core can run servers, IoT boards, cloud containers (like docker), and more, in addition to your laptop and phone. It is equipped with a full mobile armament for IT security operations and digital forensics. Additionally, everything you need to build your own applications or protect your Internet privacy is included. India is becoming digitalized day by day by the government modernizing the system, but that doesn't imply hackers don't evolve and adapt to their surroundings. with high levels of protection and pen testers. On the other hand, finding the location and tracking out the attacker will be even more useful for cyber threat analysts to recover from the attacks.

References

1. Bella, G., Biondi, P., Bognanni, S., & Esposito, S. (2023). PETIoT: PEnetration Testing the Internet of Things. *Internet of Things*, 22, 100707.
2. Shanmugapriya, R. "A study of network security using penetration testing." In 2013 International Conference on Information Communication and Embedded Systems (ICICES), pp. 371–374. IEEE, 2013.
3. Singh, Harshdeep, and Jaswinder Singh. "Penetration Testing in Wireless Networks." *International Journal of Advanced Research in Computer Science* 8.5 (2017).
4. Naga Suneetha, A. R. V., & Narasimhareddy, K. V. (2019). Secure Energy Trade-offs in Wireless Sensor Networks. *Instrumentation, Mesures, Métrologies*, 18(1).

5. Sandhya, S., Purkayastha, S., Joshua, E., & Deep, A. (2017, January). Assessment of website security by penetration testing using Wireshark. In 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 1–4). IEEE.
6. Holik, Filip, Josef Horalek, Ondrej Marik, Sona Neradova, and Stanislav Zitta. “Effective penetration testing with Metasploit framework and methodologies.” In 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (*CINTI*), pp. 237–242. IEEE, 2014.
7. Arora, M., & Singh, A. (2014). Vulnerability Assessment and Penetration Testing: A Practical Approach. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(5), 764–770.
8. Balasubramaniyan, V., Srinivasan, A. K., & Vasudevan, S. K. (2014). A Framework for Conducting Penetration Testing and Vulnerability Assessment. *International Journal of Computer Applications*, 85(11), 7–12.
9. Alzaabi, M., & Alghathbar, K. (2018). A Review of Tools and Techniques for Vulnerability Assessment and Penetration Testing. *Journal of Information Processing Systems*, 14(6), 1446–1460.
10. ul Hassan, S. Z., Muzaffar, Z., & Ahmad, S. Z. (2021). Operating Systems for Ethical Hackers-A Platform Comparison of Kali Linux and Parrot OS. *International Journal*, 10(3).
11. ul Hassan, Syed Zain, Zainab Muzaffar, and Saleem Zubair Ahmad. “Operating Systems for Ethical Hackers-A Platform Comparison of Kali Linux and Parrot OS.” *International Journal* 10.3 (2021).

Assessing NSAID Threat Degree of Unfavorable Medical Reactions Using Machine Learning



Raja Vikram Gandham, P. Tharun Kumar, B. Mahesh Gopal, M. Karthik, and Krishan Dev Nidumolu

Abstract Non-steroidal anti-inflammatory drug (NSAID) use has been documented as being widespread among older people with persistent pain. Though NSAIDs are fundamental in maintaining their quality of life, the risk of the use of multiple medications drug interactions and adverse effects is of the utmost significance as the elderly usually require multiple medications for their co-morbidities. Prescriptions are likely to expose patients to hazardous interactions with medicines and potentially deadly adverse effects if they are not properly monitored and controlled. The purpose of this study was to evaluate the appropriateness of NSAID use and analyze the risk of potential interactions with NSAIDs. The purpose of this study was to evaluate the suitability of NSAID use and examine the risk of potential reactions to NSAIDs. The model is tested on a hold-out set after being trained on a portion of the data. The model can accurately forecast the danger level of medications such as weaknesses, side effects, etc. According to the results, which may help with drug development and regulatory decision making, the proposed work performs successfully as the RMSE of Grid Search-CV LR is 0.5 and the accuracy score is 80%.

Keywords Non-Steroidal anti-inflammatory drugs (NSAID) · Adverse drug reactions (ADR) · Machine learning algorithms · Drug reactions · Medical trails

R. V. Gandham (✉) · P. T. Kumar · B. M. Gopal · M. Karthik
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

K. D. Nidumolu
MS (Cyber Security) GHSCSE, FDU, Teaneck, NJ, USA

1 Introduction

A common class of drugs known as non-steroidal anti-inflammatory drugs (NSAIDs) are those that have analgesic, anti-inflammatory, and antipyretic characteristics. They have a vital role in controlling pain and inflammation caused by many kinds of illnesses, from sudden traumas to chronic ailments like arthritis. NSAIDs have a lengthy history of clinical usage and have undergone significant studies.

There are number of NSAIDs on the market, both as prescription medications and as OTC medicines. Acetylsalicylic acid (aspirin), ibuprofen, naproxen, diclofenac, and ketoprofen are a few examples that are commonly utilized. The strength, period of action, and significant side effects of each NSAID might vary. While some NSAIDs just inhibit the COX-2 enzyme, others also inhibit the COX-1 and COX-2 enzymes. This selectivity influences how they affect different biological processes and the risks that come with them. NSAIDs are popular and effective, but they are not without risks. The most important issues are gastrointestinal (GI) consequences, such as indigestion, stomach ulcers, and gastric hemorrhage. Use of NSAIDs for a prolonged period of time or at high doses increases the risk of developing these side effects. Furthermore, certain NSAIDs, particularly those that specifically inhibit COX-2 enzymes, have been linked to cardiovascular problems. A rise in strokes and cardiac events is one of these dangers. Additionally, NSAIDs, particularly in people with innate kidney illness, might alter renal function, resulting in fluid retention and possibly harming the kidneys. Additional factors to take into account when using NSAIDs are allergic reactions and medication interactions. Because all pharmaceuticals have the potential for adverse drug reactions, risk-benefit analysis is crucial for every prescription written. Even though NSAIDs are widely used and have known dangers, more study is still necessary to fully comprehend their effects, safety profiles, and recommended use. By examining the particular impacts and potential risks connected with the use of NSAIDs in a given environment, this study paper tries to close this gap. The study aims to add to the body of knowledge and aid in the process of making well-informed decisions by critically assessing the information that is currently available and undertaking a thorough analysis. A prior exposure is required for an allergic ADR, which is unconnected to dose. When a drug functions as an allergy or an antigen, allergies form. In some cases, clinical history and the proper skin testing can aid in the prediction of allergic ADRs. When a medicine's dose or plasma concentration exceeds the therapeutic range, whether purposefully or unintentionally, it can lead to undesirable effects. These unfavorable effects are referred to as drug toxicity (drug overdose). Abuse or dependency, major physiological harm (such include harm to the heart, liver, or kidneys), psychological impairment (hallucinations, unusual conduct patterns, and memory loss), or even can result from the misuse of recreational or therapeutic substances [3].

2 Necessity of ADR Prediction

Machine learning techniques can be quite helpful in interpreting different data kinds to forecast ADRs. These techniques utilize a range of input data, such as text mining, expression of genes, and chemical structures. There is a strain to shorten the duration of clinical trials by decreasing the count as it already requires a certain number of trial participants in the range 12–15 years to obtain a medicinal substance through a clinical trial before it is accepted. The reasons of adverse reactions to medicines that have been approved include drug–drug interactions, disease severity, previously mentioned lack of drug compatibility, physical characteristics of the person, which depend on race, ethnicity, and—most importantly—the patient’s individual characteristics like age, sex, weight, and overall health, which are though not the most important factors, are among the reasons behind adverse reactions to medicinal products [5]. It is advantageous to be able to collect prospective ADRs from a variety of sources since it improves the amount of data that can be tested and enhances the likelihood that lethal ADRs will be found early. Nonetheless, the largest flow in the current system continues to be under reporting of ADRs. The ADRs for a given medicinal medication product are identified using one of two different approaches (using frequentist and Bayesian principles). The Reporting Odds Ratio and the Proportional Reporting Ratio are the two most well-known frequentist techniques. All of these approaches compare the ADR reports are disproportionate to suspected drug abuse products and other medications in the database, despite the fact that the processes involved in each approach are different [6].

3 Literature Survey

Parolini (2020) proposed a framework for assessing toxicity of the non-steroidal anti-inflammatory drugs (NSAIDs) such as acetylsalicylic acid, paracetamol, diclofenac, ibuprofen, and naproxen toward freshwater invertebrates.

Modasi, A., Pace, et al. (2019) performed a systematic review and investigated the anastomotic leak rate following NSAID use vs control after colonic or rectal anastomosis. Meta-analysis was performed to assess for overall risk of anastomotic leak with NSAID use.

Meyer, P. F., Tremblay-Mercier et al. in 2019 did an Investigation of Naproxen Treatment Effects in Pre-symptomatic Alzheimer’s Disease (INTREPAD), a 2-year double-masked pharmaco-prevention trial, enrolled 195 families history with elderly (mean age 63 years) participants and screened carefully to exclude cognitive disorder (NCT-02702817).

Abu Esba, L. C., Alqahtani et al. in 2021 did a prospective cohort study between April 12 and June 1, 2020. Adults consecutively diagnosed with COVID-19 were included. Information on NSAID use was collected through a telephone questionnaire, and patients were followed up for COVID-19 infection outcomes, including death, admission, severity, time to clinical improvement, oxygen requirement, and length of stay.

Micallef, J., Soeiro et al. in 2020 identified the relation among non-steroidal anti-inflammatory drugs, pharmacology, and COVID-19 infection.

Olojede et al. proposed a nanomaterial-based medication delivery method as potential holder for COVID-19 patients [23]. This review presents new perspectives on the use of nanomaterial-based drug delivery systems in the therapy of COVID-19 to improve the bioavailability, lessen the toxicity, and improve the efficacy of current medications.

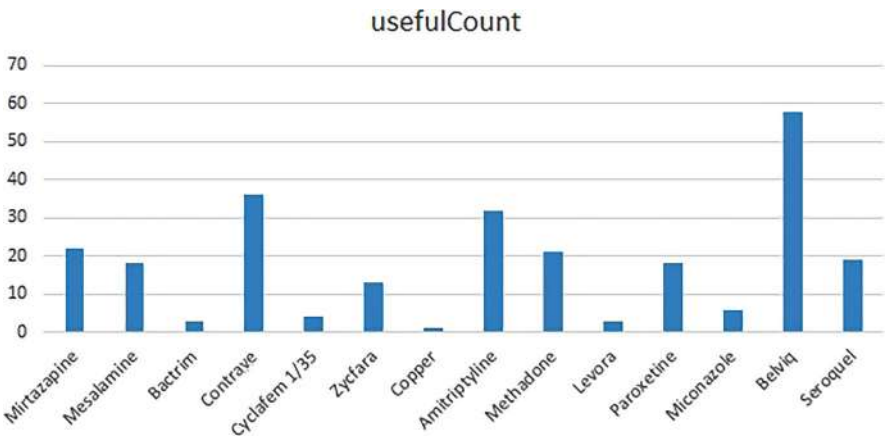
A risk assessment model for immunomodulatory drug-associated thrombosis in multiple myeloma patients was developed and validated by Sanfilippo et al. [9] published in *National Comprehensive Cancer Network Journal*. In the sensitivity analysis, from which patients being treated with previous anticoagulants ($n = 300$) were excluded, the resulting model had an HR of 1.34 ($P = .04$) for high- versus low-risk groups, with a c-index of 0.52.

Zeraatkar et al. [8] research includes collecting data From January 2004 to August 2019, the databases MEDLINE, EMBASE, and the International Pharmaceutical Abstract (IPA) were searched and considered studies that described warfarin-drug interactions with other medications.

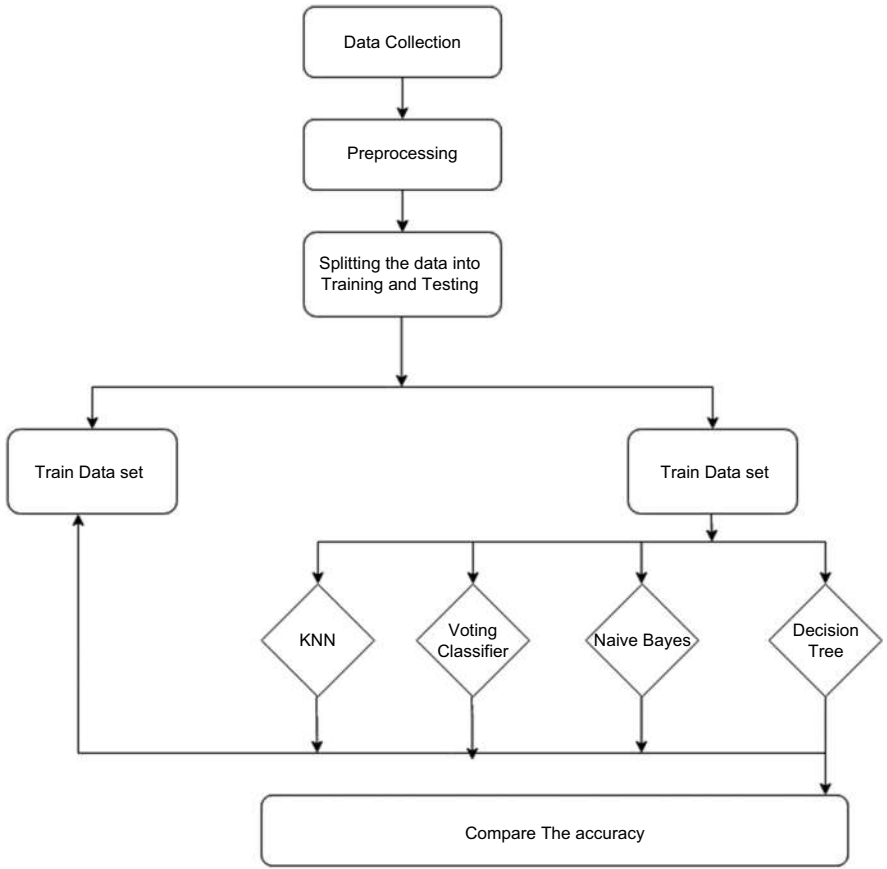
4 Dataset Description

The dataset is collected from [Kaggle.com](https://www.kaggle.com) and PubChem, which consists of 2,15,065 observations out of which 1,61,298 were used for training and 53,767 samples were used for testing purposes. The dataset contains following attributes:

- Unique id: Represents the id for each observation
- Drug Name: Name of the Drug
- Condition: Condition after people consume the Drug
- Review: Review Given by Users after the drug consumption.
- Rating: Rating Given By users.
- Date: Date of the given review.
- Useful count: Represents the count of number of people who found the Drug Useful.



Some of the collected data may contain missing fields and error or noisy data. So, the collected gone through pre-processing to get cleaned.



System architecture

Dataset Link

- (i) <https://www.kaggle.com/datasets/sahilfaizal/drug-prescription-based-on-consumer-reviews>
- (ii) <https://www.kaggle.com/datasets/deepalighodki/medicine>

5 Proposed Work

In this study, we constructed machine learning models that can predict ADRs and identify the molecular substructures connected to those ADRs without describing the substructures beforehand.

In this proposed work, we have applied KNN, Decision tree, Naive Bayes, Voting Classifier, and GridSearchCV-LR on the datasets and calculated their accuracies and RMSE values to find out the best model out of all.

5.1 Methodology

The work involves a series of steps, collection of data, pre-processing of the data to remove noisy data, and after that the data will be partitioned into training dataset and testing dataset, models will be trained using train dataset and tested with testing dataset.

Data Collection

The dataset is collected from [Kaggle.com](https://www.kaggle.com) and PubChem, which consists of 2,15,065 observations out of which 1,61,298 were used for training and 53,767 samples were used for testing purpose.

Dataset Link:

- (i) <https://www.kaggle.com/datasets/sahilfaizal/drug-prescription-based-on-consumer-reviews>
- (ii) <https://www.kaggle.com/datasets/deepalighodki/medicine>

Data Preprocessing

Data preprocessing is done to clean the data. We used data normalization techniques to bring all the data points to the same scale to make sure that different features are comparable.

Data Partitioning

The collected data is partitioned into training and testing set. A total of 1,61,298 samples were used for training and 53,767 samples for testing purposes.

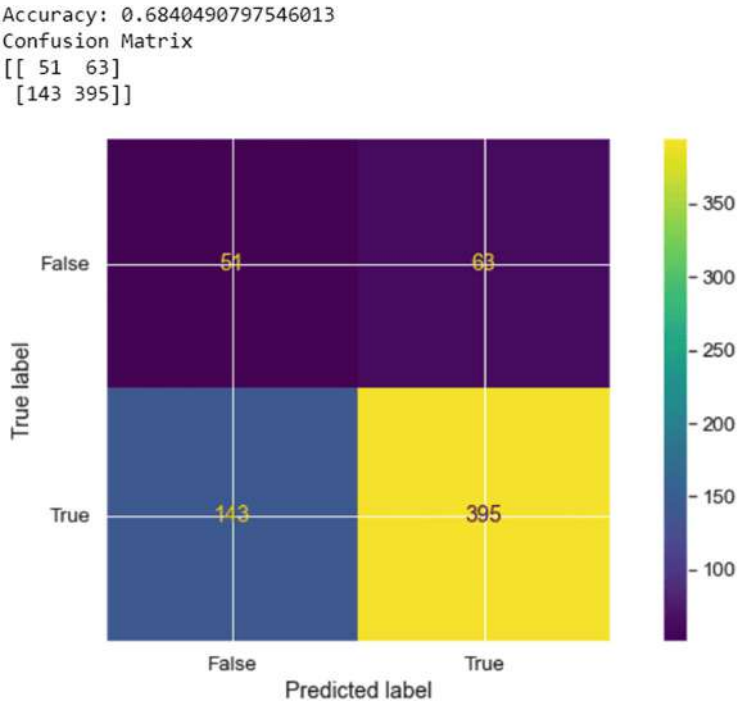


Fig. 1 Confusion matrix for KNN

Defining and Implementing ML Algorithms

In this step of the proposed work, we’ve used different ML algorithms and trained them on the training dataset that we partitioned in the previous stage. The ML algorithms used are as follows:

KNN

The k-nearest neighbors algorithm is a supervised machine learning classifier, popularly known as KNN or k-NN, that utilizes proximity to create predictions or classifications about how a single data point would cluster. It can be used to tackle concerns with segmentation or regression, but because it is typically utilized as a classification tool since it is based on the notion that nearby relevant patterns can be located. It should be stressed that the KNN method, instead of proceeding through a training stage, purely saves a training dataset. It is a member of the models from the “slow learning” family as shown in Fig. 1.

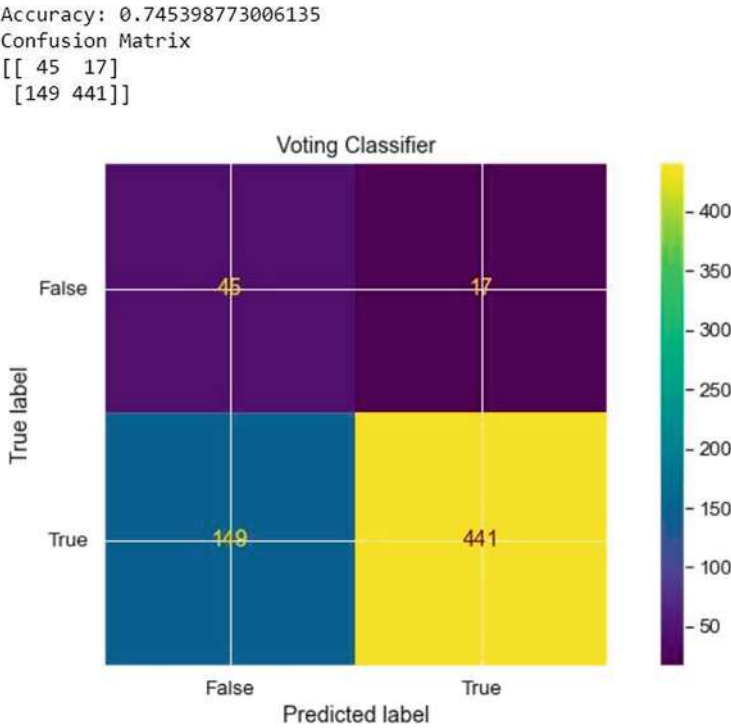


Fig. 2 Voting classifier confusion matrix

Voting Classifier

A computer-based estimator that establishes a number of base models or estimators and then delivers forecasts by averaging their result is an example of a voting classifier. The aggregating criteria can be paired with voting for every estimator output as shown in Fig. 2.

Naive Bayes

A family of basic “probabilistic classifiers” is referred as naive classifiers. Strict (naive) independence assumptions and the Bayes theorem between the features form the Bayesian basis of classifiers, which are used in statistics (see Bayes classifier). Although they are some of very simplest Bayesian network models, they can attain high degrees of precision when combined with estimating kernel density [9]. In most of the classifiers, it is conceivable to use maximum-likelihood training rather than a more intensive iterative approximation performed via assessing a confined proposition in constant time, which is faster as shown in Fig. 3 [9].

Bayes Theory

The Bayes theory relies on formulating a hypothesis (*H*) based on the available evidence (*E*). The following equation explains the Bayes Theory:

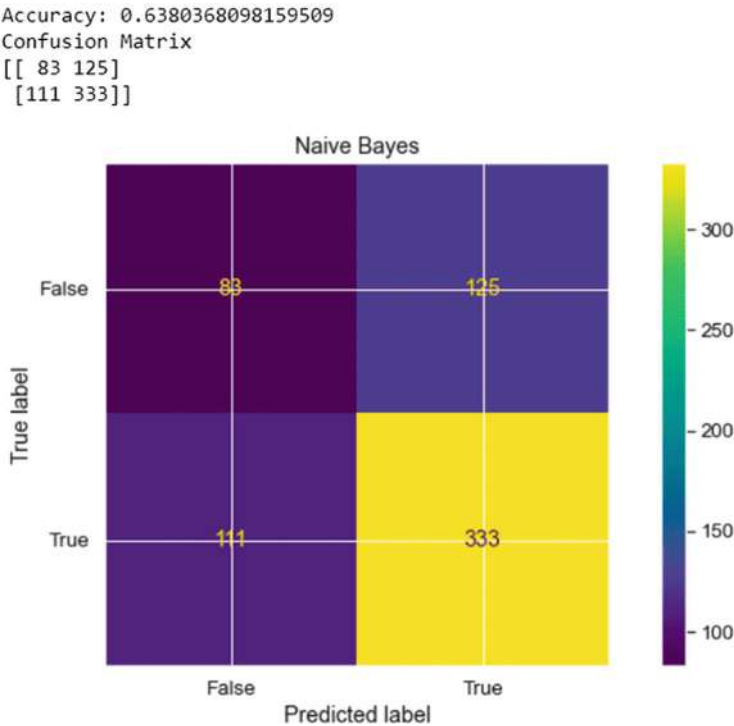


Fig. 3 Naive Bayes confusion matrix

$$P\left(\frac{H}{E}\right) = P\left(\frac{E}{H}\right) * \frac{P(H)}{P(E)}$$

Bayes Theory from a Machine Learning Standpoint

Training data are available to help you develop and refine your model. The data must then be verified before the model can be assessed and fresh predictions made. Finally, in the training data, you must refer to the input attributes as “evidence” and the outcomes as “outputs.”

Decision Tree

You define the input and the matching output in supervised machine learning approaches like decision trees while a unique criterion is used to repeatedly segregate the training data. The two aspects that can be employed to explain the tree are decision nodes and leaves. The leaves symbolize the choices or consequences. The data is separated. by the decision nodes as shown in Fig. 4.

Shannon Entropy is also another word for entropy, is a measure of the degree of uncertainty or unpredictable data’s nature, and is represented by the sign $H(S)$ for a finite collection S .

Accuracy: 0.6441717791411042
 Confusion Matrix
 [[88 126]
 [106 332]]

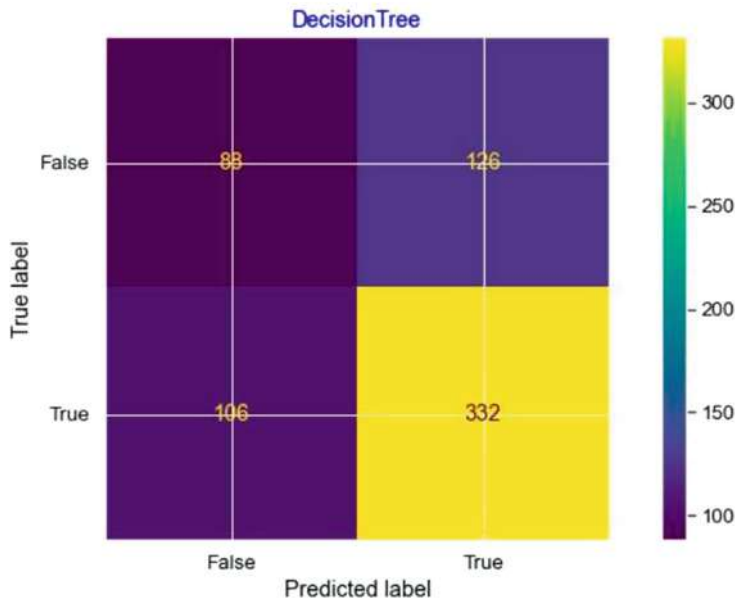


Fig. 4 Decision tree confusion matrix

$$H(S) = \sum_{x \in X} p(x) \log_2 \frac{1}{p(x)}$$

Information Gain

The Kullback-Leibler separation, denoted by the sign $IG(S, A)$ for a set, is another name for information gain. As a result of selecting a particular characteristic A , S indicates the actual variance in entropy. It measures the relative variance in entropy with respect to the independent constituents.

$$IG(S, A) = H(S) - \sum^n p(x) * H(x)$$

GridSearchCV-LR

Picking the correct model hyperparameters is one of the primary concerns in designing a machine learning model. This configuration known as hyperparameters specifies the intricacies of how the model method will be trained.

Although the preponderance of ML models has extensive documentation, preferring hyperparameters may remain tricky. Grid Searching is a methodology for

determining the ideal parameters that comprise frequently testing various options to evaluate how they impact the performance of the model. Fortunately, the GridSearchCV tool from Scikit-learn was designed to automate this operation. While it iterates over parameters, GridSearchCV also preserves the accuracy scores of the hyperparameter combinations, allowing a convenient examination of how those options affect the model.

The following equation can be used to illustrate the logistic regression model using hyperparameters:

$$P(y = 1/x) = 1/1 + e^{-z}$$

- where $p(y = 1 | x)$ is the likelihood that the event will occur given the input variables.
- A vector of input variables makes up x .
- The input variables' linearly summed weights (coefficients), z , are as follows:
The exponential function is

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

6 Experimental Results

For each technique that is being applied on the drug dataset, different kinds of accuracy values were obtained and with the help of confusion matrix we are able to find the exactness of each of these techniques. In Figs. 5 and 6, we have included the confusion matrix of each of the techniques Obtained, and with the help of confusion matrix we are able to find the exactness of each of these techniques. In Figs. 7 and 8, we have provided the accuracy comparison of various techniques.

Making Predictions Using Test Dataset

The trained ML algorithms are then tested using the test dataset that we obtained during the data partitioning process and accuracies are calculated.

7 Potential Advantages & Applications

The findings of this NSAID risk assessment project have significant implications for healthcare professionals, regulatory bodies, and patients alike. The comprehensive evaluation of the risks associated with NSAID use provides valuable insights that can inform decision making and enhance patient safety. The algorithms proposed in this work have higher accuracy and low RMSE. The Existing works solely focused on set of NSAID drugs, we here worked on large set of NSAID drugs. The following are potential benefits and applications of this research:

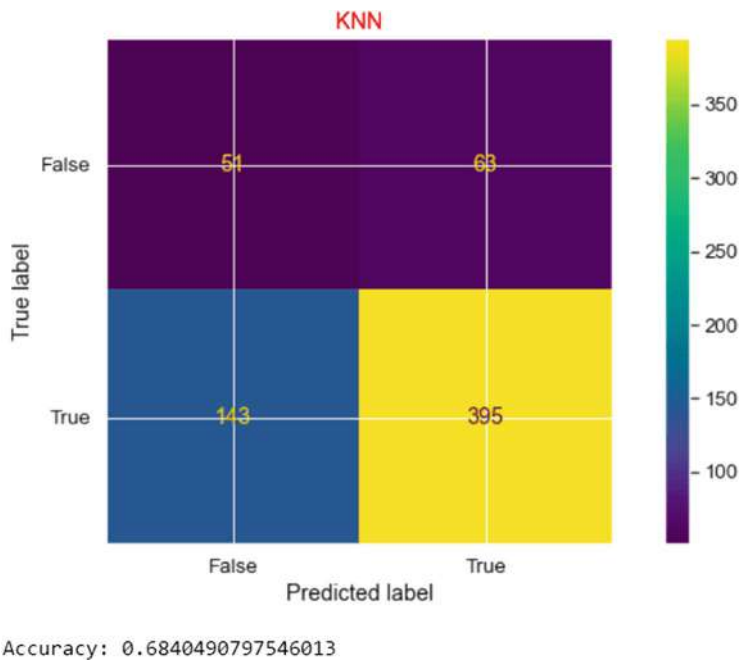


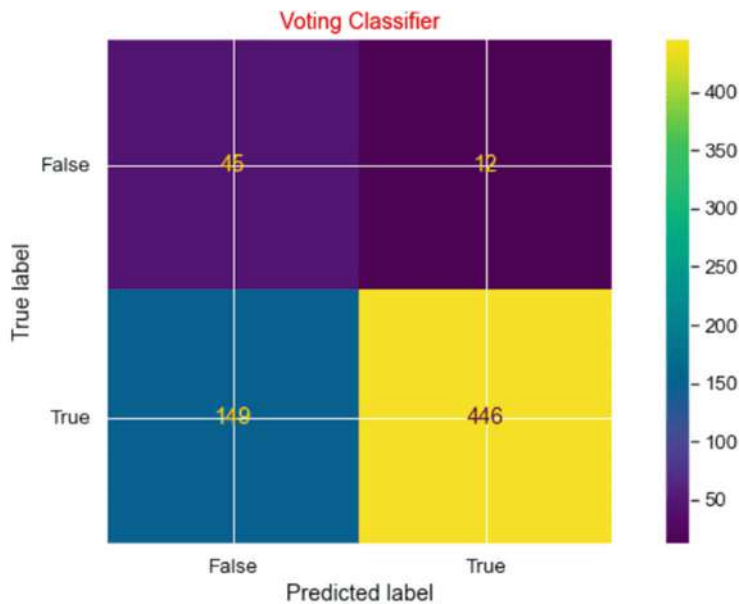
Fig. 5 KNN confusion matrix

7.1 Informed Prescribing Decisions

The risk assessment framework developed in this study can assist healthcare professionals in making informed prescribing decisions when it comes to NSAIDs. By considering individual patient characteristics, such as age, underlying health conditions, and concurrent medications, healthcare providers can better assess the potential risks and benefits of NSAID use for each patient. This personalized approach can help optimize patient outcomes and minimize the occurrence of adverse events.

7.2 Enhanced Patient Counseling

The results of this risk assessment project can facilitate improved patient counseling and education regarding the use of NSAIDs. By communicating the potential risks associated with NSAID use, patients can make more informed decisions about their treatment options. They can also be empowered to monitor and report any adverse effects, thus enabling early intervention and prevention of complications.



Accuracy: 0.7530674846625767

Fig. 6 Voting classifier confusion matrix

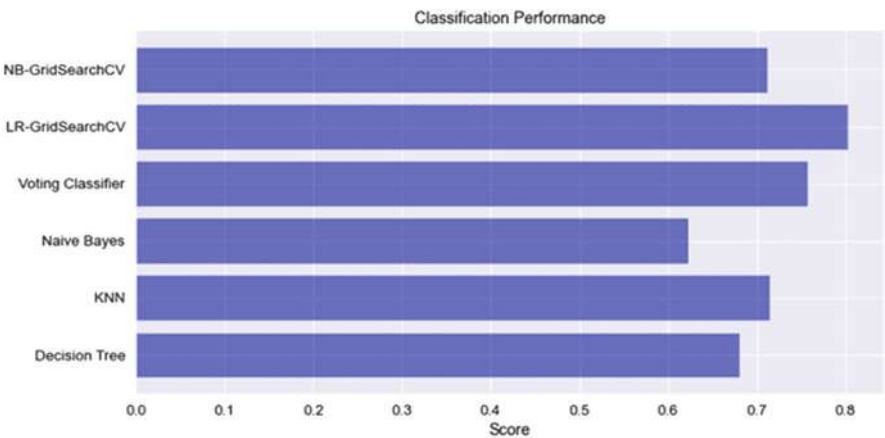


Fig. 7 Accuracy comparison

7.3 Development of Guidelines and Policies

The findings of this study can contribute to the development or revision of guidelines and policies related to NSAID use. Regulatory bodies and healthcare

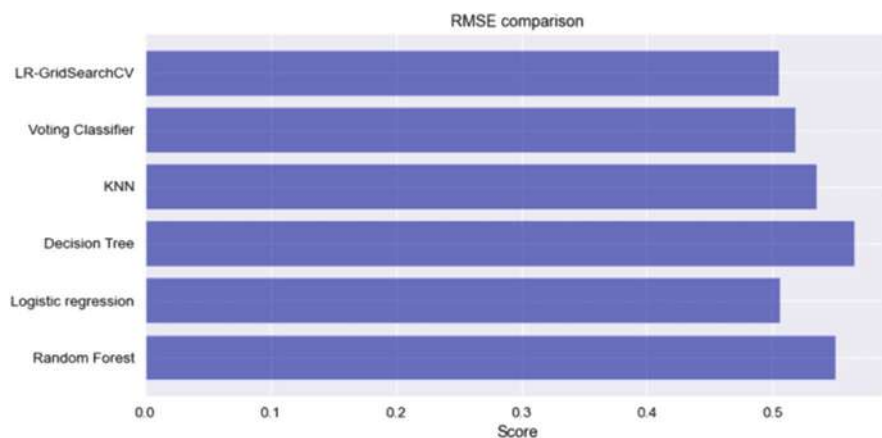


Fig. 8 RMSE Comparison

organizations can utilize this research to establish evidence-based recommendations for the safe and appropriate use of NSAIDs. These guidelines can serve as a reference for healthcare professionals in their clinical practice, ensuring consistent and standardized care across different healthcare settings.

7.4 Pharmacovigilance and Post-marketing Surveillance

The risk assessment methodology employed in this project can be integrated into pharmacovigilance and post-marketing surveillance systems. By continuously monitoring the safety profiles of NSAIDs, regulatory authorities and pharmaceutical companies can identify and address emerging risks promptly. This proactive approach can help mitigate potential harm to patients and support ongoing risk-benefit assessments of NSAIDs.

8 Challenges Faced

8.1 Domain Expertise

A solid knowledge of pharmacology, medical research, and regulatory frameworks is necessary for drug risk assessments. While machine learning models can offer beneficial knowledge, they cannot replace the insight of humans. It will be more advantageous if you add expert knowledge and involve experts in the field in the creation and interpretation of the risk assessment model.

8.2 Data Availability

The data used in this study was acquired from only a handful of sources; therefore, there may be biases and missing data. To the best of our capacity, we enhanced the data quality by essential information preprocessing, which included cleaning, filtering, and imputation procedures. We tried collecting various resources like PubChem, drug bank, etc. but we're unable to access their datasets. So, if more accurate and complete datasets are collected and used for assessing their effects, it will be more efficient

9 Conclusion

In conclusion, the goal of this study was to examine how NSAIDs affect. There have been a number of significant discoveries made through the study of pertinent papers and empirical data. First, NSAIDs, such as acetylsalicylic acid, paracetamol, diclofenac, ibuprofen, and naproxen, have repeatedly shown their effectiveness in lowering inflammation and relieving pain. These drugs have been proven to significantly relieve a range of ailments, including mild to moderate pain and inflammation. It's crucial to remember that using NSAIDs carries some hazards. Long-term or high-dose NSAID use has been linked to gastrointestinal side effects including stomach ulcers and bleeding. Additionally, several NSAIDs, particularly those in the COX-2 inhibitor class, have sparked worries about their cardiovascular effects. Despite these challenges, NSAIDs remain valuable tools for pain management and inflammation control. Their accessibility and widespread availability make them an important option for patients seeking relief from various conditions. Moving forward, further research is warranted to address the limitations of this study. Larger-scale clinical trials and longitudinal studies could provide more robust evidence on the long-term effects and safety profiles of NSAIDs.

References

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6297296/>
2. <https://www.msdmanuals.com/enin/professional/clinical-pharmacology/adverse-drug-reactions/adverse-drug-reactions>
3. <https://www.pharmacologyeducation.org/clinical-pharmacology/adverse-drug-reactions>
4. <https://www.ibm.com/in-en/topics/knn>
5. Wang, M., Zeraatkar, D., Obleda, M., Lee, M., Garcia, C., Nguyen, L., ... & Holbrook, A. (2021). Drug–drug interactions with warfarin: A systematic review and meta-analysis. *British journal of clinical pharmacology*, 87(11), 4051–4100.
6. https://en.wikipedia.org/wiki/Naive_bayes_classifier
7. <https://www.turing.com/kb/an-introduction-to-naive-bayes-algorithm-for-beginners>

8. Sun, H., Qiu, K., Li, P., and Wei, J. (2020). Utilizing SMOTE and machine learning techniques, drug risk level is predicted from adverse drug reactions. 185761–185775 IEEE Access, 8.
9. Mangoni, A. A. (2012). Predicting and detecting adverse drug reactions in old age: challenges and opportunities. *Expert opinion on drug metabolism & toxicology*
10. R. Liu, M. D. M. AbdulHameed, K. Kumar, X. Yu, A. Wallqvist, and J. Reifman (2017). medication interactions that cause negative drug responses can be predicted using data. 18, 1–18 BMC Pharmacology and Toxicology
11. Byrne, Petrovic, O'Mahony, D., Eustace, J., and Gallagher, P. (2018). The development and prospective validation of an ADR risk assessment instrument in older, multi-morbid patients led to the development of the adverse drug reaction risk in older persons (ADRROP) prediction scale. 9, 191–199 European Geriatric Medicine.
12. Kovačević, M., Vezmar Kovačević, S., Radovanović, S., Stevanović, P., & Miljković, B. (2019). Adverse drug reactions caused by drug–drug interactions in cardiovascular disease patients: introduction of a simple prediction tool using electronic screening database items. *Current medical research and opinion*, 35(11), 1873–1883.
13. Zhang, G., & Nebert, D. W. (2017). Personalized medicine: Genetic risk prediction of drug response. *Pharmacology & therapeutics*, 175, 75–90
14. Onder, G., Petrovic, M., Tangiisuran, B., Meinardi, M. C., Markito-Notenboom, W. P., Somers, A., et al (2010). The GerontoNet ADR risk score was created and validated to measure the risk of adverse drug reactions in hospitalised patients 65 years of age or older. *Internal Medicine Archives*, 170(13), 1142–1148.
15. Bereznicki, L. R., Bereznicki, B. J., Peterson, G. M., Connolly, M., Parameswaran Nair, N., Chalmers, L. (2016). Prediction of hospitalisation for drug-related adverse effects in elderly people receiving care in their communities (the PADR-EC score). 11(10), e0165757; PLoS One.
16. A. Lorenzo Hernandez, P. Barquilla, M. D. C. Fernández Capitan, N. Pacheco, and J. F. Sánchez Muoz-Torrero (2010). Internal medicine units' adverse drug reactions and associated risk factors. *Journal of European Clinical Pharmacy*, 66, 1257–1264.
17. Z. Yin, X. W. Chen, Y. Tan, Y. Hu, X. et al. (2016). Methods, 110, 14–25. From knowledge discovery concerning adverse medication reactions to clinical usage, enhancing drug safety.
18. Morsi, M., El-Sayyad, G. S., Olojede, S. O., & Abd Elkodous, M. (2021). Patients with COVID-19 may benefit from nanomaterial-based medication delivery devices, according to RSC Advances, 11(43), 26463–26480.
19. A. Holbrook and M. Wang, D. Zeraatkar, M. Obeda, M. Lee, C. Garcia, L. Nguyen, and M. Obeda (2021). Warfarin medication interactions: A comprehensive review and meta-analysis. 87(11), 4051–4100, *British journal of clinical pharmacology*.
20. Sinha, A., Ramish, M., Kumari, S., Jha, P., & Tiwari, M. K. (2022, January). ANN-ANT-LION-MLP Ensemble Transfer Learning Based Classifier for Detection and Classification of Oral Disease Severity. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 530–535). IEEE.
21. Kumar, B., Roy, S., Singh, K. U., Pandey, S. K., Kumar, A., Sinha, A., . . . & Rasool, A. (2023). A Static Machine Learning Based Evaluation Method for Usability and Security Analysis in E-Commerce website. *IEEE Access*
22. Sinha, A., Kumar, B., Banerjee, P., & Ramish, M. (2021, December). HSCAD: Heart Sound Classification for Accurate Diagnosis using Machine Learning and MATLAB. In *2021 International Conference on Computational Performance Evaluation (ComPE)* (pp. 115–120). IEEE.
23. Abu Esba, L. C., Alqahtani, R. A., Thomas, A., Shamas, N., Alswaidan, L., & Mardawi, G. (2021). Ibuprofen and NSAID use in COVID-19 infected patients is not associated with worse outcomes: a prospective cohort study. *Infectious diseases and therapy*, 10, 253–268.
24. Parolini, M. (2020). Toxicity of the Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) acetylsalicylic acid, paracetamol, diclofenac, ibuprofen and naproxen towards freshwater invertebrates: A review. *Science of the Total Environment*, 740, 140043.

Searchable Encryption for Privacy Preserving with Fine-Grained Access Control



Manne Archana, Raparathi Pranathi, Kalakota Shreya, and Boodida Nikhitha

Abstract Corporate cloud users and private cloud users are very concerned about the security of their information in the cloud. Before sending the information to the distant distributed storage server, the information can be encrypted at the client end to maintain the confidentiality of data. The keyword search feature ought to have the option to work over scrambled cloud data in order to maintain user data confidentiality. Searchable encryption makes it easy to browse encrypted data on cloud servers without decrypting it. Single keyword-based searchable encryption allows users to access a subset of documents containing keywords of their interest. Furthermore, it shouldn't provide any details regarding the retrieved document or the term used to search. In this chapter, we provide a single keyword-based searchable encryption technique where several owners contribute their information and many users can get to the information. The attribute-based encryption scheme is utilized, allowing clients to access specific subsets of information from the cloud without uncovering access rights to the cloud server. Random oracle models show that the plan adaptively protects against chosen keyword attacks. Implemented the scheme on a Google cloud instance, and the performance of the plan proved useful in real-world applications.

Keywords Searchable encryption · Data privacy · cloud security · Cloud storage · Keyword search · Data confidentiality · Attribute-based encryption

1 Introduction

Cloud computing offers consumers a robust architecture that enables elastic and limitless resources as services. Cloud computing is delivering on-demand computer resources such as databases, software, servers, storage, and networking over the

M. Archana (✉) · R. Pranathi · K. Shreya · B. Nikhitha
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,
Springer Proceedings in Mathematics & Statistics 441,
https://doi.org/10.1007/978-3-031-51338-1_17

225

Internet. Businesses and individuals can access and use these resources from a cloud service provider on a pay-per-use basis rather than maintaining their own physical infrastructure and software. In comparison to traditional computer models, cloud computing provides a number of advantages. This allows you to have a high degree of scalability and flexibility when accessing computing resources via the Internet from any place in the world. It reduces the upfront cost and continued maintenance costs, thus eliminating the necessity for organizations to buy and store costly hardware and software. Although you can store your data in the public cloud. This facilitates easier accessibility but poses concerns about data security and access control.

Due to the availability of on-demand data and services at more affordable prices in an effective manner, many organisations find outsourcing data to the public cloud to be an appealing business approach. In any case, security and protection of information have become significant concerns for specialist organizations and their customers while embracing cloud administration, specifically for public cloud. The outsourced data may include delicate data. One practical solution to the problem of information security and access control is to scramble the reports prior to transferring them to a distributed storage server. This allows us to utilize the public cloud to upload encoded data by multiple owners. So that multiple clients can view that data.

Attribute-based encryption (ABE) permits you to encrypt and decrypt data using attributes rather than specific keys. A particular key is required to encrypt and decrypt data using conventional encryption techniques. Using factors like a person's work title, age, or location, data can be encrypted with ABE. The data is encrypted, and only people with the necessary qualities can decrypt it. Access control and secure data sharing are only two examples of the many applications that ABE can be utilized in.

A practical necessity for contemporary cloud storage data is the ability to explore databases with encrypted data. A few techniques for looking through scrambled information have been presented in the field of searchable encryption. A receiver may be able to securely and arbitrarily get data from publicly available cloud storage that is of interest to them by using searchable encryption.

For instance, a specialist needs to look at every one of the patient's records with kidney illness. So that specialist needs to give entrance privileges to the patient's clinical records. Since the patient must initially encode his clinical records with a mystery key before sharing them with the doctor, single-user searchable symmetric encryption techniques cannot be used in this situation. A data owner (patient) can create a shared mystery key by utilizing his master secret key, which he can then give to authorized users (in this case, the doctor) for searching through encrypted data. The effective scheme used in order to mandate searches over encrypted material and enforce access control restrictions is a multi-user searchable symmetric encryption scheme. But these plans can be successful in a single-sender, multiple-receiver configuration, they are less effective in a multi-sender, multiple-receiver configuration. Since every sender must securely interact with every receiver in order to distribute the mystery key, incurs a significant

communication overhead. In scenarios with several senders and receivers, the ABE scheme performs effectively. Direct communication between the data sender and the data receiver is not necessary. In order to provide single-term search capabilities over attribute-based encrypted data, numerous systems for keyword-based searching of encoded information have been introduced. Wang et al.'s [4] approach offers single keyword searches and assigns a portion of the decryption effort to the cloud service provider.

2 Literature Survey

The literature has many proposals for attribute-based, keyword-searchable encryption systems. However, the crucial issue of receiver anonymity is not addressed by these approaches. Wang et al.'s [4] approach offers single keyword searches and assigns a portion of the decryption effort to the cloud service provider. Authorized Searchable Public-Key Encryption (ASPKE), a method introduced by Shi et al. [11], in which an information proprietor chooses the entrance strategy for his encoded information and maintains it concealed inside the ciphertext. The main drawback of Shi et al.'s study is that the client must obtain the search token from a trusted authority.

We propose an efficient “privacy preserving single keyword-based searchable encryption scheme (PSE) with fine-grained access control.” It allows the information proprietor to regulate the usage and search of its cloud-based, scrambled data that has been outsourced in accordance with its access control policy. Our approach makes it possible to do keyword searches and fine-grained access controls over encoded data. The framework is set up with several senders and multiple receivers, enabling a data owner (sender) to easily encode the keywords list associated with his report and upload it to cloud storage along with the access strategy and encoded content. The data owner selects the access strategy, which is kept a secret inside the cipher text.

3 Preliminaries

(i) *Bilinear Mapping*

Bilinear mapping is a mathematical operation that takes two elements from two different mathematical groups and creates an element in a third group. Let G_0 and G_1 be two multiplicative cyclic groups of a large prime order P . Let g be a generator of G_0 and e be a bilinear map defined as $e: G_0 \times G_0 \rightarrow G_1$, satisfying the following properties:

- Bilinearity: $e(g^a, g^b) = e(g, g)^{ab}$ for all $a, b, \in \mathbb{Z}_P^*$.

- Non-degeneracy: There exists $g_1, g_2 \in G_0$ such that $e(g_1, g_2)$ not equals to 1.
- There exists an efficient computable algorithm to compute $e(g_1, g_2)$ for all $g_1, g_2 \in G_0$. We say that G_0 is a bilinear group if it satisfies the above-mentioned three properties.

(ii) *Decisional Bilinear Diffie-Hellman (DBDH) Assumption*

Let $a, b, c, z \in \mathbb{Z}_p^*$ be chosen at random and g be a generator of G_0 . The decisional BDH assumption is that no probabilistic polynomial-time algorithm I can distinguish the tuple $(A = g^a, B = g^b, C = g^c, e(g, g)^{abc})$ from the tuple $(A = g^a, B = g^b, C = g^c, e(g, g)^z)$ with more than a negligible advantage ϵ_{dbdh} . The advantage of I is $\Pr[I(A, B, C, e(g, g)^{abc})=0] - \Pr[I(A, B, C, e(g, g)^z)=0] = \epsilon_{dbdh}$.

(iii) *Access Structure*

An access structure depicts the access policy for a ciphertext elected by the encryptor. A user can decrypt a ciphertext if the attributes included in his/her secret key can satisfy the access structure of the ciphertext. We use the access structure built on the policy of “Single AND gate on Multivalued Attributes” as follows: Let there be n attributes in the universe and the attributes are denoted using the notation $\{A_1, A_2, \dots, A_n\}$. We use i to indicate the attribute A_i ($1 \leq i \leq n$). Each value set $V_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,m_i}\}$ is the set of possible values for an attribute i and m_i is the size of set V_i . A ciphertext policy T is defined as $T = [T_1, T_2, \dots, T_n]$. Here, each T_i ($T_i \subseteq V_i$) represents the set of permissible values of an attribute i in order to decrypt the ciphertext. Each user possesses an attribute value list $L = [L_1, L_2, \dots, L_n]$, where L_i ($L_i \in V_i$) represents one value from the value set V_i . An attribute list L satisfies an access structure T if $L_i \in T_i$ for all $1 \leq i \leq n$. We define a binary relationship $F(L, T)$, which gives output 1 if L satisfies an access structure T ; else, it outputs 0.

4 System Architecture

The system architecture (shown in Fig. 1) consists of the following elements.

Attribute Center (AC): The Attribute Center is in charge of creating system parameters and providing users with keys.

Token Generator (TG): This supports the creation of an encrypted index by the data owner, which involves the process of creating a scrambled index. The AC itself might act as the TG in a small system or organization.

Data Owner (DO): Information is encrypted and kept on a distributed storage server by the data owner. The encrypted information have (a) the encrypted keywords list, and (b) the encrypted reports.

Cloud Service Provider (CSP): CSP offers the system entities storage and processing assistance.

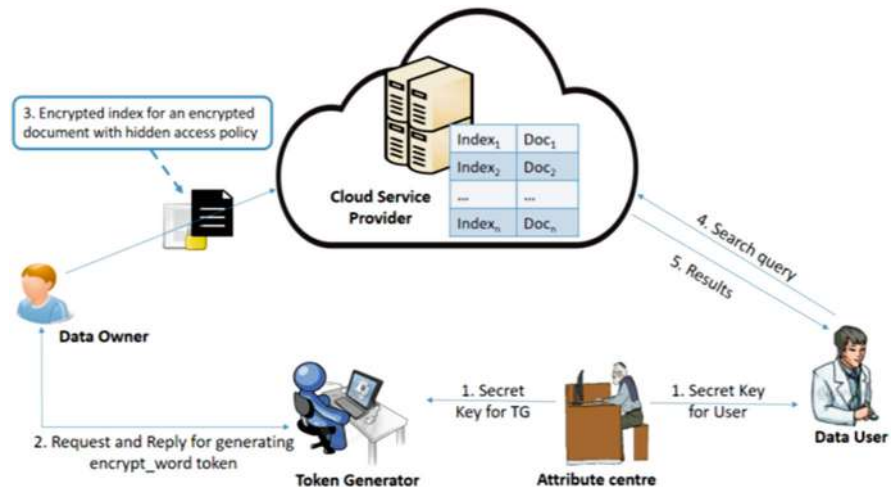


Fig. 1 System architecture

Data User (DU): Data user creates and submits a trapdoor to cloud service provider. Using this trapdoor, the CSP glances through the scrambled lists. When a search operation returns true, the client gets the documents linked to the indexes. Finally, the resultant documents are decrypted by the user.

5 Methodology

The modules of the PSE scheme are as follows:

- I. **Data Provider Interface:** In this, the data provider first creates an account by using registration, immediately the application server sends public key when account is created successfully. Only authenticated users can login to the application by using their username and password. The data provider can't possess any rights after login. Data provider will get a secret key after proper login. Data provider is supposed to submit those public and secret keys that are provided by the server, only after which the data can be uploaded to the application cloud storage.
- II. **File Encryption:** In this, the data provider uploads the files which are outsourced securely by using encryption. The encrypted data is known as cipher text which is not readable. Once the content has been uploaded to the cloud storage, the data provider is likewise unable to view the original content directly. Data provider can only view the contents of the file, which is displayed in cipher text, and the metadata of the previously uploaded file.

- III. Data utilizer module: In this, the user first creates an account as a data utilizer to login. After logging in to the application, if the data utilizer gets permission from the authenticator, he/she can search the records that are uploaded by the data provider and access the file content; Otherwise, errors such as not requesting the authenticator could occur.
- IV. Request to Authenticator: Data users must take two steps if they are unable to access the records that data providers provided. First, send the request immediately to the authenticator, mentioning your email address and a username. Second wait for the authenticator's responses to your request. After that, you can only access the files, which are encrypted.
- V. File Decryption Technique: In this case, the encrypted file name is submitted to the server by the authenticated data user. It is then able to gather the data blocks and cross-check the decryption key and provide an instant key. Receive the original file only after correctly submitting the immediate key and decrypt key.

6 Algorithms Used

PSE: Privacy-Preserving Searchable Encryption

We describe an “attribute-based searchable encryption” with a data owner-enforced access control that is concealed in the ciphertext. The purpose of the multi-sender and multi-receiver arrangement of the scheme is to make it simpler for a DO to scramble the list of keywords associated with his report and upload it with the access strategy and the scrambled information to distributed storage, where the DO chooses the access policy and keeps it secret within the ciphertext. The user (receiver) makes their search queries available to cloud storage servers in the form of trapdoors. The cloud server utilizes this trapdoor to search over completely encoded files transferred on the distributed storage. The results of the user's query are the record associated with the search-enabled index.

The five-tuple of PSE scheme is given below:

- $\text{Setup}(1^\beta)$: It is run by the AC. The input is security parameter β . And it returns the public key Pk , the TG's secret key Tsk , and master private key Mpk .
- $\text{KeyGen}(\text{Mk}, \text{Ls})$: The randomized algorithm $\text{KeyGen}(\text{Mk}, \text{Ls})$ is executed by the attribute center. The algorithm accepts a list of user characteristics Ls and the master key Mk as inputs. It creates a secret key Sk to the client. A trapdoor is created using the key Sk to carry out search operations.
- $\text{Encrypt_Index}(\text{Pk}, \text{X}, \text{V}, \text{Tsk})$: This convention is utilized between the DO and the third party, where X is the keywords list connected to document B . The data owner initiates the calculation to produce the encryption for each keyword w contained in keyword set X . To perform the encryption of the keyword, TG provides the data owner with a scrambled token for each w . TG employs his

personal secret key, Tsk , to produce encrypted word tokens. The DO then outputs a scrambled words list known as the “encrypted index for document B” (CT_X).

- **Trapdoor**(Pk, Sk, x): The randomized procedure is used by the receiver user to create a trapdoor so he can access the documents from CSP that include an scrambled entry for the word “ x ” in their related indexes and then he has the necessary access privileges and outputs a trapdoor tx created for x .
- **Search**(tx, CT_X): The CSP executes this deterministic algorithm. The user-sent trapdoor tx and the encrypted index CT_X are inputs to the method. The calculation returns valid if the word in tx coordinates matches with any keyword remembered for CT_X and client’s key fulfills the access strategy of CT_X .

Advanced Encryption Standard (AES) Algorithm

AES algorithm is used for data encryption purposes. The Advanced Encryption Standard (AES) is a popular and widely used symmetric encryption algorithm. It is six times faster than Triple DES. A DES alternative was needed because the DES key size was too small. As computing power increased, it was considered vulnerable to attacks involving large-scale key searches. To overcome this drawback, Triple DES was designed but was found to be slow. AES is based on substitution–permutation network. It consists of a series of linked operations, which involve replacing an input with a particular output (permutation) or rearranging bits (permutation).

7 Results

Searchable encryption for privacy preserving with fine-grained access control is a cryptographic method that allows users to search over scrambled information while maintaining the privacy of the information. On the basis of specified access controls, it also offers the capability of limiting who has access to specific portions of the data. To accomplish its objectives, this method combines encryption, access control, and search algorithms. Asymmetric or symmetric encryption algorithms are used to encode the data. A searchable encryption method is then utilized to index the encrypted material, allowing visitors to browse through it without learning what’s within. Who is permitted to access the encrypted data and what operations they are permitted to carry out on it are decided by the access control mechanism. Results are shown in Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.

7.1 Performance Analysis

A comparison of proposed PSE scheme with Shi et al’s scheme [11] is shown in Table 1. For making a single keyword search operation, we consider the parameters of Ref. [11]. As shown in Table 1, the PSE scheme size and time complexity are



Fig. 2 Login page of Searchable encryption for privacy preserving with fine-grained access control



Fig. 3 Login and registration of data provider

more than Shi et al’s scheme. The main limitation of Ref. [11] is users must acquire the search tokens from a trusted authority, which increases the per-query interaction cost of the search process on user side. Another limitation [11] is that we can specify at most one value for each attribute placed in the ciphertext access policy, while in the PSE scheme, multiple values for each attribute can be placed inside the ciphertext access policy. In Table 1, this feature is captured in the size of T_i . In Shi et al.’s scheme, if a user wants to issue τ different search queries, then he needs to communicate τ times with the trusted third party, which adds additional

DataUtilizer

DataUtilizer Login and Register

Register

UserName

Password

ConfirmPassword

Mail-Id

City

address

Register

Login

UserName

Password

Login

Fig. 4 Login and registration of data utilizer

DataUtilizer

Search Logout



Send request to Authenticator for accessing Cloud data

UserName

Mail-Id

SendRequest

Fig. 5 Data user sends the request to authenticator

communication cost on user side. In the PSE scheme, once the user has obtained the private key from a trusted third party, he can generate the search token yourself Without interacting with a trusted third party.

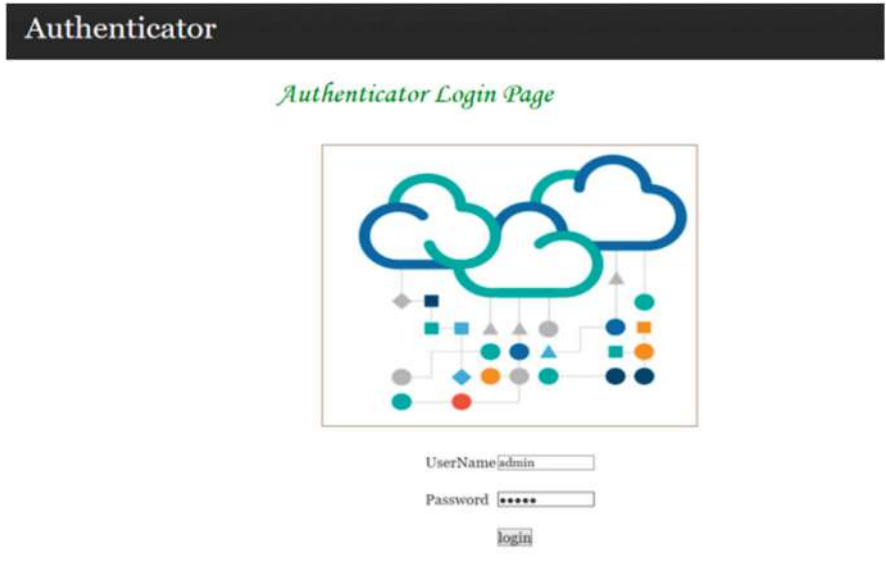


Fig. 6 Login page of authenticator

Accepted Requests

[Logout](#)

UserName	Mail-Id	Status
<input type="text" value="ds&em"/>	<input type="text" value=".com"/>	<input type="text" value="valid"/>
<input type="text" value="ds&em"/>	<input type="text" value=".com"/>	<input type="text" value="valid"/>
<input type="text" value="ds&em"/>	<input type="text" value=".com"/>	<input type="text" value="valid"/>
<input type="text" value="asdas"/>	<input type="text" value="asdas@gmail.com"/>	<input type="text" value="valid"/>
<input type="text" value="asdas"/>	<input type="text" value="asdas@gmail.com"/>	<input type="text" value="valid"/>
<input type="text" value="asdas"/>	<input type="text" value="@gmail.com"/>	<input type="text" value="valid"/>
<input type="text" value="asdas"/>	<input type="text" value="@gmail.com"/>	<input type="text" value="valid"/>
<input type="text" value="mn&em"/>	<input type="text" value=".com"/>	<input type="text" value="valid"/>
<input type="text" value="mn&em"/>	<input type="text" value=".com"/>	<input type="text" value="valid"/>
<input type="text" value="abc&e"/>	<input type="text" value="il.com"/>	<input type="text" value="valid"/>

Fig. 7 Accepted requests of authenticator

7.2 Experimental Results

The PSE scheme can be implemented by using pairing-based cryptography (pbc) library [13]. Bilinear pairings are constructed on the curve $y^2 = x^3 + x$ over the field F_m for prime $m = 3 \bmod 4$, where the order of the groups G_0 and G_1 is a prime



Fig. 8 Data user submits file name



Fig. 9 File keys

of size 160 bits and the length of m is 512 bits. The scheme can be evaluated with different number of attributes (p) = 3, 5, 7, and 10 and their different sizes of value sets (q). The results shown in the graphs are taken from the experiments where we have taken the fixed size of the value-set for each attribute as 5. The total number of attribute values ($q * p$) are 15, 25, 35, and 50. The experiments are performed using the dataset available in [12], which contains the diabetes patient's medical records. Each report is identified with four keyword values. They are (a) Test conducted date, (b) blood sample taken time, (c) report type, and (d) test outcome. The Setup and KeyGen algorithms are executed by the AC. The time complexity of Encrypt_Index protocol and Trapdoor algorithms are shown in Figs. 13 and 14. The performance of the scheme operations depends linearly on the total number of attribute values. The Search algorithm is tested on a Google cloud computing

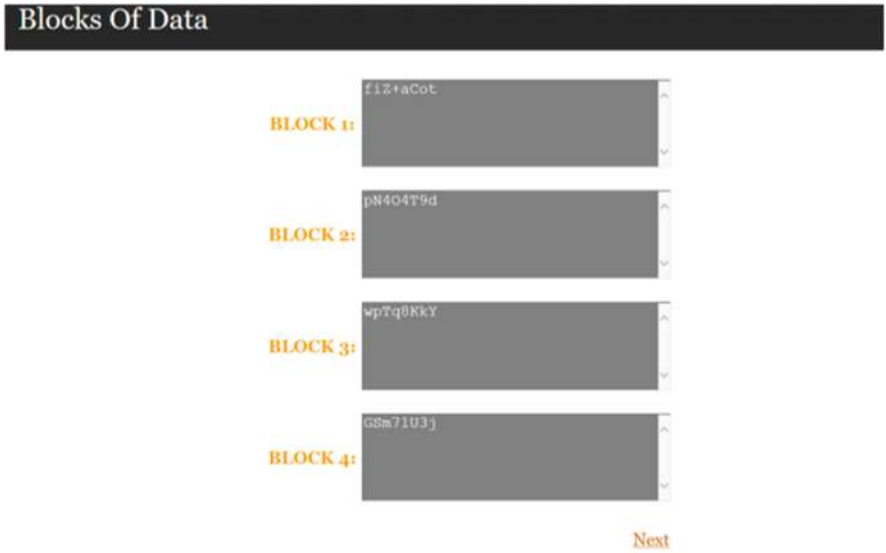


Fig. 10 Data user gets blocks of data

instance of p1 series. The input of search algorithm is the output of Encrypt_Index and Trapdoor algorithms. Figure 15 shows how long it takes to search over 500–10000 encrypted file indexes with different numbers of total attribute values, where each index contains 4 keywords. The search operation time complexity is $O(q * p)$, where $q * p$ is the total number of attribute values in the system. Therefore, we have shown the results with different values of $q * p$.

8 Conclusion

We proposed “A Privacy preserving attribute based searchable encryption” scheme. Which permits an authorized client to search through encrypted documents stored on CSP for a subset of documents matching his search criteria and fulfilling his access privileges. Privacy of client access privileges and confidentiality of information can be protected by the PSE scheme. The search functionality of the PSE scheme is shown to be adaptively secure against chosen keyword attacks under the “Decisional Bilinear Diffie-Hellman (DBDH)” assumption of the random oracle model. The PSE system has been put into practise on a Google cloud instance, and it has proven to be successful in real-world applications.



Fig. 11 Data user submitting the keys

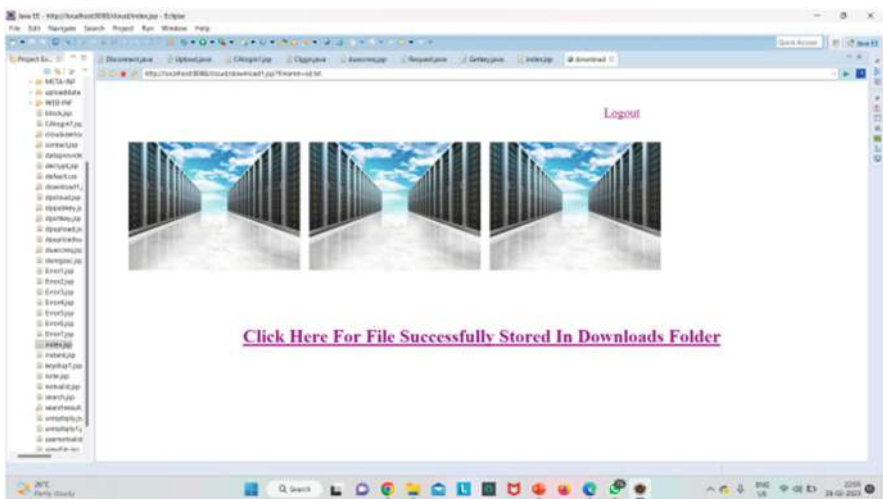


Fig. 12 Data user accessing the requested file

Table 1 Comparison table of Shi et al.'s with proposed PSE scheme

Scheme	Shi et al.'s scheme	The proposed PSE scheme
Access Policy Structure	LSSS	AND-gate on Multi-valued
Attribute value(s) in access policy	Single value ($0 \leq T_i \leq 1$)	Multi values ($(1 \leq T_i \leq m)$)
Group Order	Composite	Prime
Universe of keywords	Bounded	Unbounded
Public key size	$(p + 5) \cdot G_o $	$(p * q + 5) \cdot G_o $
Ciphertext size	$(2p + 2) \cdot G_o $	$(p * q + 1) \cdot G_o $
Search token size	$(p + 4) \cdot G_o $	$(p * q + 2) \cdot G_o $
Time complexity of search operation	$(2p + 1)t_p + (2n)t_e + (3n)t_m$	$(q * p + 1)(t_p + t_m)$
Communication Overhead for generating τ number of search tokens	$O(\tau)$	$O(1)$
Adversary's knowledge	Search outcome and the field of keyword	Only Search outcome

p : Number of attribute fields; $q = \max(|V_i|) \ 1 \leq i \leq n$, where V_i = value set for attribute i ; $T_i \subseteq V_i$. The values in T_i are included in the access policy; t_p, t_M, t_E denotes the time of pairing operation; the time of multiplication and the time of exponentiation operation in group G_1 , respectively

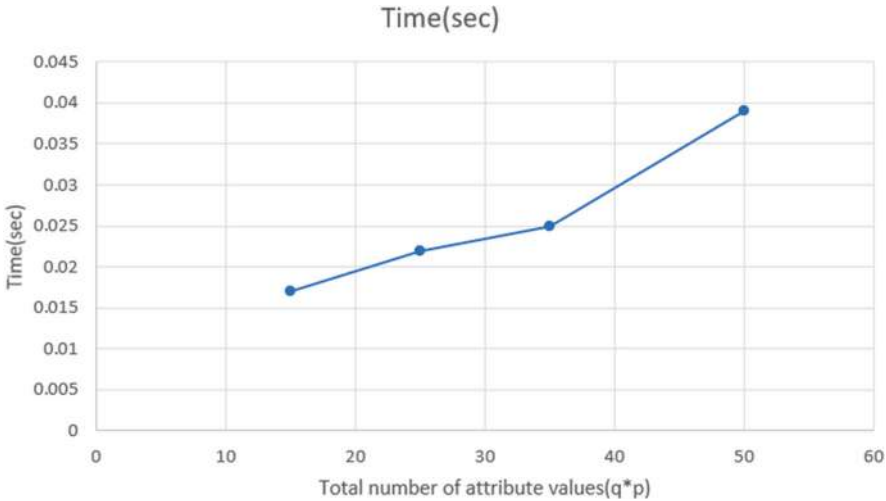


Fig. 13 Encrypt_index time for 4 keywords index

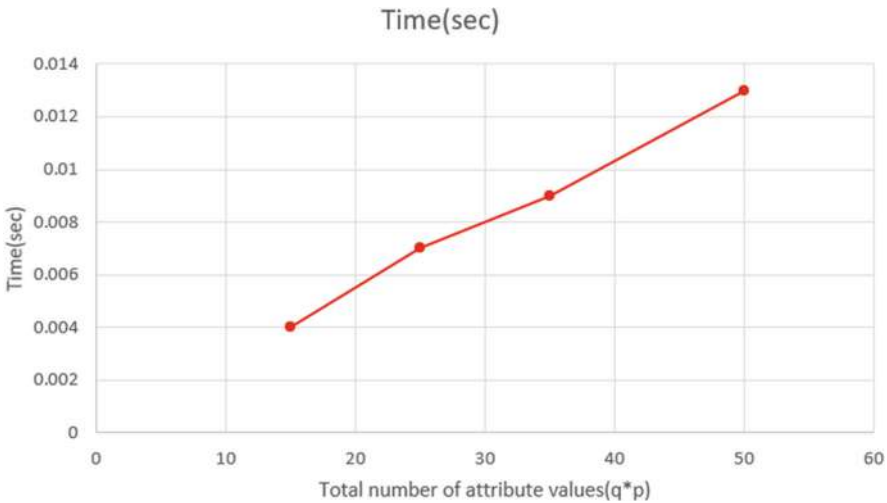


Fig. 14 Trapdoor_generation time for a keyword

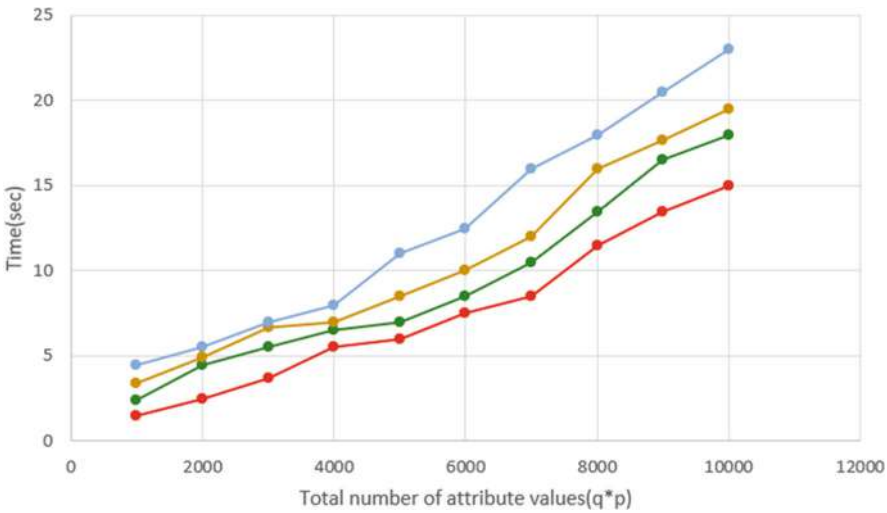


Fig. 15 Search operation time over number of encrypted indexes with different size of total attribute values ($q \cdot p$)

9 Future Scope

The majority of current searchable encryption solutions offer straightforward keyword-based searches. Future studies might concentrate on making it possible to perform more intricate searches, such as range queries and wildcard searches, while still protecting user privacy and access control.


Despite the fact that searchable encryption offers excellent privacy protection, there is still a chance that information could still be compromised, particularly in the event of sophisticated attacks. The security guarantees of searchable encryption may be improved by the use of novel techniques, such as machine learning-based anomaly detection methods, according to future studies.

References

1. Payal Chaudhari and Manik Lal Das, "Privacy Preserving Searchable Encryption with Fine-grained Access Control", *IEEE Transactions on Cloud Computing* (Early Access), 2021.
2. Wang, K.; Dong, X.; Shen, J.; Cao, Z. An Effective Verifiable Symmetric Searchable Encryption Scheme in Cloud Computing. In *Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City*, Kuala, Malaysia, 18 April 2019; pp. 98–102.
3. Ghareh Chamani, J.; Papadopoulos, D.; Papamanthou, C.; Jalili, R. New constructions for forward and backward private symmetric searchable encryption. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, Toronto, ON, Canada, 15–19 October 2018; pp. 1038–1055.
4. C. Wang, W. Li, Y. Li, and X. Xu. A ciphertext-policy attribute-based encryption scheme supporting keyword search function. In *Proceedings of Cyberspace Safety and Security*, LNCS 8300, Springer, pp. 377–386, 2013.
5. R. Cohen, "The cloud hits the mainstream: More than half of U.S. businesses now use cloud computing," <http://www.forbes.com>, April 2013. Online; posted 10-January-2017.
6. R. Curtmola, J. Garay, S. Kamara, R. Ostrovsky, Searchable symmetric encryption: improved definitions and efficient constructions. *Journal of Computer Security*. 19(5), pp. 895–934, 2011.
7. Zhang, Y.; Wang, Y.; Li, Y. Searchable Public Key Encryption Supporting Semantic Multi-Keywords Search. *IEEE Access* **2019**, 7, 122078–122090.
8. S. Jarecki, C. Jutla, H. Krawczyk, M. Rosu, M. Steiner. Outsourced symmetric private information retrieval. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, ACM pp. 875–888, 2013.
9. V. Goyal, A. Jain, O. Pandey, A. Sahai. Bounded ciphertext policy attribute based encryption. In *Proceedings of Automata, Languages and Programming*, pp. 579– 591, 2008.
10. Q. Zheng, X. Shouhuai, and G. Ateniese. VABKS: Verifiable attribute-based keyword search over outsourced encrypted data. In *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*, pp. 522–530, 2014.
11. J. Shi, J. Lai, Y. Li, R. H. Deng, and J. Weng. Authorized keyword search on encrypted data. In *Proceedings of Computer Security*, LNCS 8712, Springer, pp. 419–435, 2014.
12. M. Lichman. UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml> [Last Accessed 23 January 2017]
13. The Pairing-Based Cryptography Library. <https://crypto.stanford.edu/pbc/> [Last Accessed 20 February 2017]

Classification for Disease Gene Association



P. LaxmiKanth , J. Akshitha, P. Akshitha, and D. Abhinandhan

Abstract The objective of the machine learning algorithms project for Disease Gene Association Classification is to devise a computerized method to recognize and group genes linked with diseases by employing machine learning techniques. The project entails scrutinizing extensive genomic data to pinpoint genetic variations that are correlated with specific ailments. The proposed method will employ various machine learning algorithms such as AdaBoost, XGBoost, and ANN to classify genes into two categories, disease-associated or non-disease-associated, based on their genetic characteristics. The outcomes will be displayed visually, and researchers can investigate the identified gene-disease connections. The ultimate objective of this endeavor is to furnish a precious asset for the medical and exploration circles, streamlining the effective and precise detection of genes linked with a wide range of ailments. This resource will be crucial in the development of personalized medicine and targeted therapeutic interventions for different ailments.

Keywords Genes · Biological features · ANN · ADABOOST · XGBOOST

1 Introduction

The progress made in genomic technologies has empowered researchers to amass copious amounts of genomic data, presenting a distinctive opportunity to detect genetic variants linked to diverse ailments. Detecting ailment-associated genes is indispensable for devising targeted therapies, comprehending disease mechanisms, and refining disease diagnosis. Nonetheless, scrutinizing large-scale genomic data to pinpoint ailment-associated genes is a strenuous task that necessitates sophisticated computational methodologies.

P. LaxmiKanth (✉) · J. Akshitha · P. Akshitha · D. Abhinandhan
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

Genes present in the human body are not standalone components, and alterations in one gene can influence its associated gene, which can further contribute to the alteration of diverse physiological processes that can lead to varied maladies. Consequently, comprehending biological mechanisms and uncovering the link between genetic disorders understanding the significance of genes is a vital obstacle in modern biology and healthcare. The fundamental concern in human health is to grasp the association between causal genes and genetic ailments. Technology is utilized to detect and supervise diverse human disorders, including Parkinson's disease.

In recent times, machine learning algorithms have arisen as a potent tool for scrutinizing genomic data and detecting genetic variants linked to different ailments. Machine learning algorithms can learn patterns automatically from large-scale genomic data and classify genes as ailment-associated or non-ailment-associated based on their genetic characteristics. This approach has the potential to significantly expedite the detection of ailment-associated genes and augment our understanding of disease mechanisms.

The project to classify Disease Gene Association using Machine Learning Algorithms aspires to design an automated approach for detecting and categorizing ailment-associated genes using machine learning techniques. The project will involve scrutinizing large-scale genomic data to pinpoint genetic variants that are associated with certain ailments. The proposed approach will utilize various machine learning algorithms such as AdaBoost, XGBoost, and ANN to classify genes as ailment-associated or non-ailment-associated based on their genetic characteristics.

The project will display the results and enable researchers to explore the identified gene-ailment associations. The platform will enable scientists to enter disease-specific genomic information and receive a roster of genes linked to the particular disease they are studying. This tool will be pivotal in the development of personalized medicine and targeted therapeutic interventions for diverse ailments.

1.1 Problem Statement

The issue at hand with disease genes is to recognize the genetic foundation of various ailments. Several disorders are caused by genetic mutations that disturb the normal operation of the body. These mutations may either be inherited from progenitors or occur unexpectedly. The recognition of disease genes is vital in comprehending the underlying mechanisms of ailments and devising efficient remedies. Genetic research has brought about the discovery of disease genes for numerous disorders, including sickle cell anemia, cystic fibrosis, Huntington's disease, and several types of cancer. However, the identification of disease genes is a multifaceted process that comprises various challenges. One of the most significant obstacles is determining rare genetic variations that are accountable for the ailment. Another hurdle is establishing the operational repercussions of genetic mutations

and their contribution to the disease. All in all, the goal of the disease gene project is to identify the genetic variations that cause the disease and comprehend their operational implications to develop effective treatments and enhance patient outcomes.

1.2 Objective

The primary aim of the machine learning algorithms for Disease Gene Association Classification project is to create an automated method for detecting and categorizing genes that are linked to diseases using machine learning techniques. The specific goals of the project are:

- To examine vast genomic datasets to discover genetic changes that are related to specific ailments.
- To create and execute a range of machine learning algorithms, such as AdaBoost, XGBoost, and ANN, to categorize genes as either disease-associated or non-disease-associated based on their genetic traits.
- To assess the efficiency of the algorithms developed by using appropriate metrics like accuracy, precision, and recall.
- To demonstrate the usefulness of the project in identifying genes associated with diseases using real-world genomic data.

By accomplishing these objectives, the project intends to provide a tool for the medical and research communities that will enable them to quickly and accurately identify genes that are associated with various diseases. This project has the potential to greatly expedite the identification of disease-associated genes and enhance our knowledge of disease mechanisms, which could result in better disease diagnosis and targeted therapies.

2 Literature Survey

2.1 Identifying Causal Genes Through Functionally Cohesive Subnetworks From Genome-Wide Association Studies

The researchers utilized a set menu strategy and computer program that utilized comprehensive shared-function networks of genomes to identify clusters of genes that are functionally linked to one another and are found in different GWA locations. They utilized data from approximately 100 GWA studies that encompass ten cancer types to evaluate the effectiveness of their method. They compared their method to the usual alternative strategy of ranking cancer genes that are already known.

The researchers determined that their method was more effective in ranking cancer genes that are already known compared to the usual alternative strategy. By employing networks of shared functions at the genome scale, it is possible to identify sets of genes that exhibit functional interdependence and are present in multiple GWA loci, they were able to detect more causal genes than the conventional approach of concentrating on genes that are closest to the associated polymorphisms.

Pros

The utilization of the “prix fixe” tactic and software may assist scientists in pinpointing causal genes and variants at GWA sites for a variety of ailments, not restricted to cancer.

The employment of genome-wide shared-function networks may alleviate partiality towards well-known genes and facilitate the detection of less-analyzed genes.

The technique might exhibit enhanced potential as additional GWA sites are found, thereby leading to the revelation of fresh genes linked to human disorders.

Cons

The research solely examined the technique on ten forms of cancer, making it uncertain how effective it would be for treating other illnesses.

The “set-menu” plan and software may demand a considerable amount of computing power to execute.

The method depends on the precision of the large-scale shared-function networks utilized, which may affect the precision of the outcomes.

2.2 Determining Primary Disease-Causing Genes through Impartial Genomic Characteristics

The scientists created a method called Objective Prioritization for Enhanced Novelty (OPEN) that utilizes a variety of genomic characteristics to objectively prioritize gene-disease relationships using machine learning. This method is impartial and not swayed towards previously established genes. The UNLOCK method was employed to identify hereditary elements for conditions linked to Cardiovascular Disease (CVD), including hypertension, lipid levels, and features related to heart muscle disease and electrical function. The group also utilized a zebrafish template to authenticate FLNC and detect a novel FLNC alternative splice mutation in an individual who experienced intense DCM. By utilizing the OPEN approach, the team successfully identified genetic factors for CVD-related traits, and genes that were given priority, such as FLNC, were studied for their correlation with an enlarged left ventricular diameter, which is a defining characteristic of DCM. The researchers further confirmed the role of FLNC by conducting experiments on zebrafish and identified a new FLNC splice-site mutation in a patient with

acute DCM. These discoveries possess the capability to clear the path toward the recognition of novel remedial objectives for cardiac ailments.

Pros

Unbiased strategy utilizes impartial prognostic characteristics and does not demonstrate partiality towards genes that have been previously well-defined.

The method effectively detected genetic factors for characteristics related to CVD and ranked genes for correlation with an enlarged left ventricular diameter, a key characteristic of dilated cardiomyopathy.

The research team verified FLNC utilizing a zebrafish model and discovered a new mutation in a splice site of FLNC in a patient with severe DCM.

The results have the potential to pave the way for new therapeutic objectives for CVD.

Cons

The research concentrated on characteristics associated with CVD and might not be relevant to other illnesses. The methodology could be unsuitable for certain genetic studies and might call for further confirmation and enhancement for diverse ailments. The utilization of machine learning methods may be limited by the availability and quality of genomic information.

3 Existing System

They used fundamental features and a few parameters, such as phenotypic similarities, in the analysis for disease gene association using machine learning method that is now in use. Genes that cause the same or related disorders are closer in the protein interaction network and have less diversity in their sequence than genes that do not. This method only allows for a limited amount of analysis.

Disadvantage

1. There is much need to know about the more efficient algorithms for disease prediction.

4 Proposed System

Figure 1 suggests and assesses fresh computational techniques such as Boosting algorithms (ADABOOST, XGBOOST) and ANN algorithms to recognize genes linked with diseases. In order to detect possible genes, we present certain advanced topographical and anatomical characteristics that are presently being disregarded. The same is represented in Fig. 1. By utilizing measures, we evaluate various computational methods using assessment criteria such as true positive rate, false

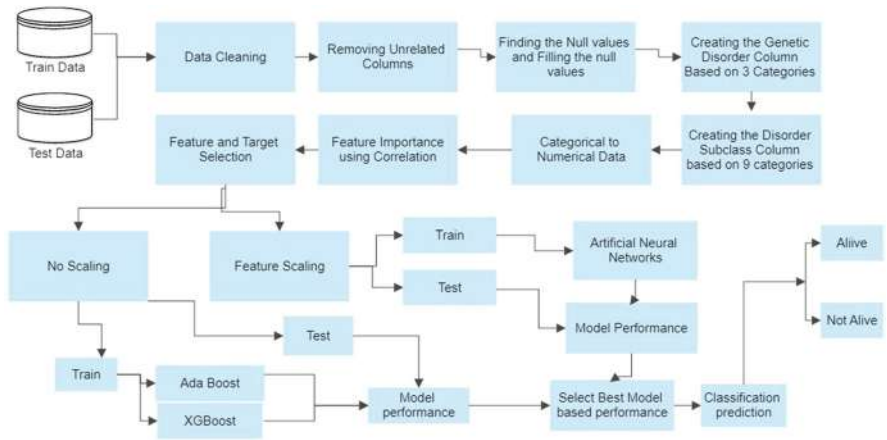


Fig. 1 Proposed architecture

positive rate, accuracy, completeness, harmonic mean of precision and recall, and receiver operating characteristic curve, on data pertaining to the association between diseases and genes in a validation mode.

Advantages

Algorithms such as AdaBoost, XGBoost, and ANN are the top-performing algorithms, with straightforward functioning that yields optimal outcomes. Moreover, they exhibit high efficacy when handling vast amounts of data.

4.1 Proposed Framework

5 Methodology

From the beginning of human existence on Earth, the world population has been increasing rapidly. In 1800, the population was estimated to be 1 billion, which rose to 6 billion by the year 2000. Currently, 227,000 people are added to the global population each day, it is projected that toward the conclusion of the twenty-first century, the populace could exceed 11 billion.

Because of the unmanageable increase in population and insufficient availability of proper healthcare services, food, and shelter, genetic disorders have become more prevalent. Lack of knowledge about the importance of genetic testing has led to an increase in hereditary diseases. To prevent child mortality caused by these illnesses, genetic testing during pregnancy is crucial.

5.1 Data Preprocessing

The procedure of refining, modifying, and organizing data for analysis and modeling. The aim of data processing is to guarantee that the data is organized in a format that can be effectively utilized by AI algorithms and to enhance the quality and dependability of the outcomes.

The measures included in data preparation typically comprise:

Data Refinement: This comprises eliminating or rectifying absent, conflicting, or replicated data.

Data Conversion: To a common range or normalizing categorical data into binary values. Once the data is preprocessed, it can be used to train machine learning models and make predictions, one-hot encoding categorical data, and normalizing the data.

Data Refinement Measures & Approaches: Here is a six-step process for refining your data to ensure its readiness.

- Eradicate inconsequential data.
- Eliminate replicated data.
- Correct any structural inaccuracies.
- Address any missing data.
- Omit any data outliers.
- Validate your data.

Data Fusion: This refers to the process of merging multiple datasets into a single comprehensive dataset, which might include concatenation, merging, or aggregation of data.

Data Compression: This involves reducing the data's dimensionality, which can be achieved by removing repetitive features, executing feature selection approaches, or applying dimensionality reduction methodologies like PCA.

Data Subsampling: This entails haphazardly choosing a segment of the information to employ for instructional objectives, while the remaining is utilized for examination. This is generally executed to avert overfitting, which transpires when the pattern is excessively intricate and matches the training data excessively well.

Data Partitioning: This involves dividing the data into training and testing datasets, which are used to teach and evaluate the machine learning model, respectively.

Overall, data preprocessing is a crucial phase in the machine learning process, as the accuracy and reliability of the outcome will be determined by the quality and structure of the data utilized for analysis and modeling.

5.2 Proposed Model Information

Deep Learning

Artificial neural networks are utilized in deep learning, which is a type of machine learning. The neural networks have numerous layers, with each layer building on

the previous one to learn progressively more intricate features and representations of data. The term “deep” refers to this layered structure.

The fundamental concept of deep learning is to train a neural network with substantial quantities of labeled data, such as text or images, and then use the network to make predictions on new, unlabeled data. The network’s weights and biases are initially set randomly and then iteratively trained using a back-propagation algorithm to adjust them to minimize the error between its predictions and the actual labels.

One significant advantage of deep learning is its ability to automatically learn feature representations from raw data, without the need for hand-crafted features. This makes it particularly useful for tasks such as autonomous driving, natural language processing, and speech and image recognition, where manually designing features would be difficult or impossible. However, deep learning models are often very complex and computationally intensive, necessitating significant amounts of data and computing power to train effectively. Additionally, the lack of interpretability of these models can make it difficult to understand how they are making predictions, which can be a challenge in certain domains such as healthcare or finance.

Machine Learning

Artificial intelligence has a subfield known as machine learning, which involves constructing models and algorithms that can learn from data and make decisions or predictions without explicit programming. The aim of machine learning is to allow computers to automatically identify patterns and insights from data and utilize them to make predictions or take actions based on new, unseen data.

Various kinds of machine learning exist, such as supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the model is trained on labeled data that already has known output or label. The model then applies this training data to make predictions on new, unseen data. In unsupervised learning, the model isn’t given any labeled data and must identify patterns and structure in the data on its own. Reinforcement learning involves training the model to make decisions based on feedback received from the environment.

Machine learning is employed in many applications, such as natural language processing, computer vision, speech recognition, recommendation systems, and predictive analytics. Some of the commonly used machine learning algorithms are linear regression, decision trees, neural networks, and support vector machines.

One of the significant benefits of machine learning is its capability to automatically discover patterns and insights in vast datasets that would be difficult or impossible for a human to detect. This can lead to better predictions and decision making in several domains, from finance and healthcare to marketing and customer service. However, it’s crucial to note that the accuracy and impartiality of resulting models depend on the quality of data they’re trained on. It’s necessary to cautiously curate and preprocess data to ensure the resulting models are accurate and unbiased.

5.2.1 ANN Model

The ANN model is a machine learning model that imitates arrangement and operation of the human cerebral cortex. It is made up of linked clusters, referred to as synthetic nerve cells, that carry out basic computations and transmit data to each other. The primary objective of an ANN is to estimate a complicated relationship between input and output data by acquiring a set of weights and biases that regulate the connections between the neurons.

A neural network architecture that comprises two hidden layers is referred to as an Artificial Neural Network (ANN) model. This model consists of two layers of neurons that are situated between the input and output layers. Below is some relevant information about ANN models with two hidden layers:

The ANN model with two hidden layers follows a sequential structure where the layers are arranged linearly. The input layer receives input data, the hidden layers perform computations and feature extraction, and the output layer generates the final output or prediction.

Each hidden layer comprises multiple nodes or neurons. The number of neurons in each hidden layer is a hyperparameter that can be determined based on the complexity of the problem and the amount of available data. It is common to have the same number of neurons in both hidden layers, although it is not mandatory.

Activation functions are applied to the outputs of neurons in the hidden layers to introduce non-linearity and enable the network to learn complex patterns in the data. Popular activation functions include sigmoid, tanh, and rectified linear unit (ReLU), among others. The choice of activation function depends on the specific requirements of the problem.

ANN models with two hidden layers are typically trained using the backpropagation algorithm. During training, the model's weights (parameters) are iteratively adjusted to minimize the difference between the predicted outputs and the true labels in the training data. The backpropagation algorithm calculates the gradients of the loss function with respect to the model's weights and updates the weights using optimization algorithms such as stochastic gradient descent (SGD) or its variants.

The output layer of an ANN model with two hidden layers depends on the specific task at hand. For classification tasks, the output layer may have multiple neurons, each representing a distinct class, and an appropriate activation function, such as softmax, is often applied to produce class probabilities. For regression tasks, the output layer usually has a single neuron and may not have an activation function, or a linear activation function like identity can be used.

After training, the performance of the ANN model with two hidden layers is assessed using appropriate metrics based on the task. For classification tasks, metrics like accuracy, precision, recall, and F1 score are commonly used. For regression tasks, metrics such as mean squared error (MSE) or mean absolute error (MAE) can be employed.

ANNs have a variety of applications, such as:

Classification: This involves assigning input data to one of several predetermined categories.

Regression: This involves forecasting a continuous outcome based on input data.

Clustering: This involves grouping similar data points into clusters.

Dimension reduction: This pertains to the method of reducing the number of variables in a dataset, while still preserving significant data

Generative models: It involves producing new data samples that resemble a given training dataset.

5.2.2 AdaBoost (Adaptive Boosting)

AdaBoost is a method of ensemble learning that merges Several feeble classifiers that are utilized to produce a potent classifier. The algorithm functions by adjusting the weights of the training examples iteratively to focus on. The present ensemble erroneously classifies certain examples. The ultimate forecast is established by computing a weighted aggregate of all the feeble classifiers. AdaBoost is a swift and straightforward algorithm that is relatively simple to implement and can be used with a wide range of base classifiers.

AdaBoost is a member of the ensemble learning category that aggregates several models' predictions to make a final prediction. In AdaBoost, the weak learners are trained in sequence, with each subsequent learner focusing on correcting the mistakes made by the previous ones.

AdaBoost necessitates the use of weak learners, which are also known as base learners. A weak learner is a model that performs marginally better than random guessing, such as a decision stump (a simple decision tree with only one split). These weak learners are usually simple and have low complexity.

During training, each instance in the dataset is assigned a weight. Initially, all instances have equal weights. As the algorithm progresses, the weights are adjusted to give more importance to the instances that were misclassified by previous weak learners. This allows subsequent weak learners to focus on the harder-to-classify instances.

AdaBoost trains the weak learners sequentially. In each iteration, the algorithm selects the weak learner that best minimizes the weighted error or misclassification rate. The weak learner is trained on the dataset with instance weights, and its performance is evaluated. The weight of the weak learner is then determined based on its accuracy, and the weights of the misclassified instances are increased. This process continues for a specified number of iterations or until the desired performance is achieved.

After training all the weak learners, AdaBoost combines their predictions using a weighted voting scheme. The weight of each weak learner depends on its accuracy during training. The final prediction is determined by combining the weighted votes of the weak learners. Typically, the weight of each weak learner in the voting scheme is proportional to its accuracy.

AdaBoost is especially effective when the weak learners are better than random guessing and perform better than chance on a subset of the data. By combining these weak learners, AdaBoost can create a strong ensemble model with improved accuracy compared to using a single model. AdaBoost performs well in various domains, including computer vision, natural language processing, and bioinformatics.

Although AdaBoost is a potent algorithm, it can be sensitive to noisy data and outliers. Instances with higher weights may dominate the training process, leading to overfitting if not handled properly. Additionally, the algorithm's performance depends on the quality of the weak learners, so if the weak learners are too complex or suffer from high bias, AdaBoost may not perform optimally.

Overall, AdaBoost is a flexible ensemble learning algorithm that has proven to be effective in improving predictive accuracy. By combining multiple weak learners, it can tackle complex classification

5.2.3 XGBoost (Extreme Gradient Boosting)

XGBoost is an enhanced version of gradient boosting, which is a technique of ensemble learning that integrates multiple weak models to construct a robust model. XGBoost is engineered to be exceptionally scalable and effective, and it has been widely applied in machine learning contests as well as practical scenarios. XGBoost has several advanced features such as parallel processing, tree pruning, and regularization, which make it a potent tool for large-scale machine-learning problems. In general, both AdaBoost and XGBoost are effective techniques for solving complex machine-learning problems, and the decision between the duo will rely on the particular demands of the issue in question. AdaBoost is a simple and fast algorithm that is well-suited for smaller datasets and simpler problems, while XGBoost is a more powerful and scalable algorithm that is well-suited for larger datasets and more complex problems.

XGBoost is an ensemble learning method based on the gradient boosting framework. It combines multiple decision trees, or weak learners, to create a stronger predictive model. The model is built in a stage-wise manner, with each subsequent weak learner trained to correct the mistakes made by the previous learners.

To prevent overfitting and enhance generalization, XGBoost incorporates various regularization techniques. These include shrinkage, or a learning rate, which controls the contribution of each weak learner to the final prediction. A smaller learning rate makes the boosting process more conservative and helps prevent overfitting. XGBoost also applies tree pruning methods during the construction of decision trees to reduce complexity and avoid overfitting. Additionally, column subsampling is supported, which selects a subset of features at each boosting iteration to force the model to focus on the most informative features.

XGBoost provides a flexible framework for defining custom objective functions and evaluation metrics. It supports various loss functions for classification and regression tasks, such as logistic loss for binary classification, softmax loss for

multi-class classification, and squared error loss for regression. Custom objective functions can be defined to handle specific tasks or incorporate domain-specific requirements.

Missing values in the data are automatically handled by XGBoost using a technique called “sparsity-aware split finding” during the tree construction process. XGBoost is highly scalable and efficient, with support for parallel processing and distributed computing for training on large datasets that don’t fit into memory.

XGBoost also provides a measure of feature importance, which can be useful for feature selection, understanding the model’s behavior, and gaining insights into the underlying data. It has gained popularity due to its exceptional performance in various machine learning competitions and real-world applications, achieving high prediction accuracy in domains such as finance, healthcare, and online advertising.

5.2.4 Evaluation Metrics

The assessment of the suggested structure is appraised relying on various statistical indicators along with our novel parameter, termed as the corona rating.

Accuracy

Accuracy is the closeness of the measured value to the standard or true value.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

- True Positive: True Positive is equivalent to the set of correct predicted significant instances.
- False Positive: False Positive is identical to the range of wrong anticipated tremendous cases.
- True Negative: The set of correctly predicted significant instances is synonymous with True Positive.
- False Negative: False Negative is identical to the range of wrong anticipated terrible cases

Recall

Recall is the responsiveness of the technique.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

Precision

Precision is a measure that calculates the proficiency of the method in specifying the perfect expected instances.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

Specificity

Specificity refers to the proportion of accurately predicted negatives in relation to the total number of negative observations.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

F1-Score

The F1-Score is a metric that evaluates the accuracy of detection.

$$\begin{aligned} \text{F1 Score} &= \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \\ &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned} \quad (5)$$

6 Experimental Results

We usually create two distinct categories of assessments for machine learning systems:

Preliminary assessments
Post-training assessments

Preliminary Assessments

These assessments are designed to be executed without trained parameters to identify implementation errors early on. This approach saves time and effort by avoiding unnecessary training.

The following aspects can be tested during preliminary assessments:

- Confirm that the model's anticipated.
- output shape is correct.
- Test dataset leakage, which involves verifying that there is no duplication between the data in the training and testing sets.
- Temporal data breach, which requires guaranteeing that the interconnections between the education and evaluation data do not lead to impractical circumstances in the temporal realm. For example, rehearsing using an upcoming data point and assessing with a previous data point.

- Authenticate the output intervals. When we anticipate outputs falling within specific intervals such as values, it is important to verify that the ultimate outcome is consistent with the anticipated scope of possibilities.
- Ensure that a step of gradient training on a group of data leads to a reduction in loss.

Assert the profiling of data.

Post-training Assessments

Evaluations conducted after the training aim to examine the conduct of the model. Our objective is to assess the acquired reasoning abilities, which can be achieved by analyzing the below-mentioned standards and additional ones:

- Tests for invariance consist of assessing the model by modifying a singular aspect of a data point and verifying the coherence of model forecasts. Suppose we are dealing with a loan forecasting dataset, in that case, an individual’s eligibility for a loan should not be influenced by a modification in their gender, given that all other features remain unchanged. When it comes to estimating the likelihood of a Titanic survivor, altering a passenger’s name must not impact their survival probability data.
- Predictive assumptions involve assessing a clear connection between attribute values and forecasts. For instance, when dealing with loan forecasting, an individual’s chances of securing a loan ought to rise with a higher credit score.
- Furthermore, you have the option to generate assessments for any other potential malfunctions that you have detected for your model.

Test Cases

Classification Report

```
In [65]: # Print the classification report
print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0.0	1.00	0.90	0.95	2702
1.0	0.92	1.00	0.96	2819
accuracy			0.95	5521
macro avg	0.96	0.95	0.95	5521
weighted avg	0.96	0.95	0.95	5521

Table 1 Test cases generated for the experimental setup

Test no.	Test cases	Expected output	Actual output	Pass/fail
1	Importing the Libraries	Imported	Imported without errors	Pass
2	Collecting the data by using the Authentications and mounting	Collected	Got Access to the data by using the authentications	Pass
3	Preprocess the data	Data is sanitized	Data is sanitized accurately	Pass
4	Normalize the data	Normalizing the data	Normalizing the data	Pass
5	Dividing information into Training and Testing	Training and Testing 80–20%	Training and Testing 80–20%	Pass
6	Dividing data into Test and Validation	Test and Validation with an equal split of 50–50%.	Test and Validation with an equal split of 50–50%.	Pass
7	Current model	Training the ANN, Adaboost and Xgboost model	Trained the ANN, Adaboost and Xgboost model	Pass
8	Model evaluation	Confusion matrix report	Confusion matrix generated	Pass
9	Predictions	User sending the new input from the validation for the prediction to classify the Prediction	Model predicting the output	Pass

Confusion Matrix

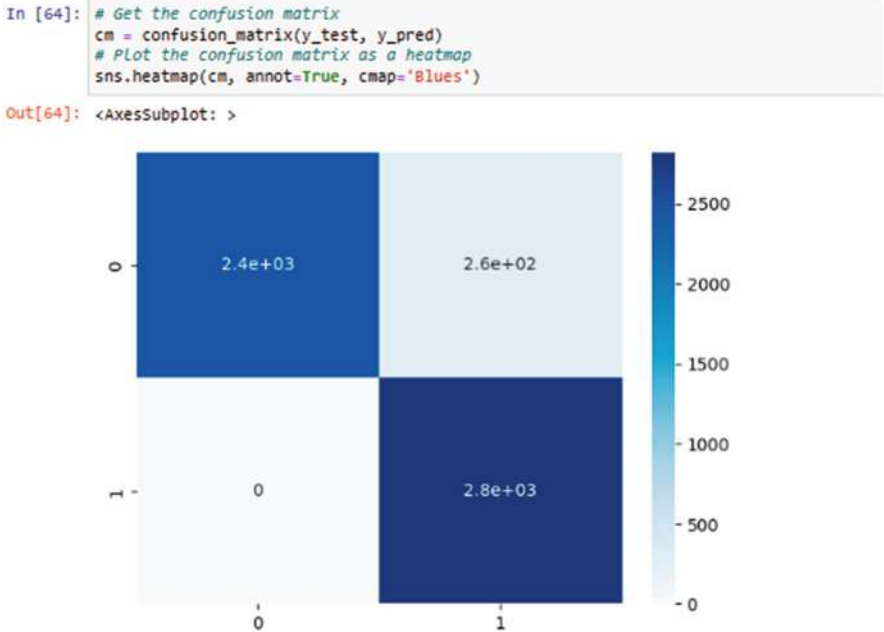
Description of the Dataset

- *Data Source:* Open Source
- *Data Collected From:* Kaggle

Link to Data Source: <https://www.kaggle.com/datasets/aryarishabh/of-genomes-and-genetics-hackerearth-ml-challenge?select=train.csv>

7 Conclusion

To conclude, the potential of machine learning algorithms like Adaboost, XGBoost, and artificial neural networks (ANN) in classifying gene-disease associations has been widely observed. These algorithms have been successful in classifying gene-



disease associations accurately while identifying key genes involved in disease pathogenesis.

Ensemble algorithms such as AdaBoost, XGBoost, and ANN have shown remarkable performance in terms of accuracy, speed, and robustness when compared to traditional machine learning methods. AdaBoost, in particular, has produced highly accurate predictive models for gene-disease association classification, achieving an accuracy rate of up to 95.3%. As a result, these machine learning algorithms can facilitate more efficient and targeted disease diagnosis and treatment, potentially improving patient outcomes and reducing healthcare system burdens. However, further research and development are necessary to fully harness the potential of machine learning algorithms in disease diagnosis and treatment.

References

1. S.Navlakha and C.Kingsford, “The power of protein interaction networks for associating genes with diseases,” *Bioinformatics*, vol. 26, no. 8, pp. 1057–1063, Apr. 2010.
2. A. Almgren, ‘An automated and intelligent Parkinson disease monitoring system using wearable computing and cloud technology,” *Cluster Comput.*, vol. 22, no. S1, pp. 2309–2316, Jan. 2019.
3. I. Ud Din, A. Almgren, M. Guizani, and M. Zuair, “A decade of Internet of Things: Analysis in the light of healthcare applications,” *IEEE Access*, vol. 7, pp. 89967–89979, 2019.

4. R. M. Piro and F. Di Cunto, "Computational approaches to disease-gene prediction: Rationale, classification and successes," *FEBS J.*, vol. 279, no. 5, pp. 678–696, Mar. 2012.
5. S. Köhler, S. Bauer, D. Horn, and P. N. Robinson, "Walking the interactome for prioritization of candidate disease genes," *Amer. J. Hum. Genet.* vol. 82, no. 4, pp. 949–958, Apr. 2008.
6. Y. Li and J. C. Patra, "Genome-wide inferring gene–phenotype relationship by walking on the heterogeneous network," *Bioinformatics*, vol. 26, no. 9, pp. 1219–1224, May 2010.
7. R.E.Adie, "Speeding disease gene discovery by sequence based candidate prioritization," *BMC Bioinf.*, vol. 6, no. 1, pp. 1–13, 2005.
8. B.J.Chen, "Disease candidate gene identification and prioritization using protein interaction networks," *BMC Bioinf.*, vol. 10, no. 1, pp. 1–14, 2009.
9. D. M. Z.-H. D. Adarsh Jose, "A gene selection method for classifying cancer samples using1Ddiscrete wave let transform,"*Int.J.Comput.Biol. Drug*, vol. 2, no. 4, pp. 398–411, 2009.
10. O. Vanunu, "Associating Genes and Protein Complexes with Disease via Network Propagation," *PLoS Comput. Biol.*, vol. 6, no. 1, pp. 1–9, 2010.

Machine Learning-Based Air Pollution Monitoring and Forecasting



Naga Ravindra Babu M , M. Durga Satish , B. V. Prasanthi ,
S. V. V. D. Jagadeesh , J. N. S. S. Janardhana Naidu ,
and Immedi Kali Pradeep

Abstract Today, the governments of developing countries are focusing more on air pollution. Air pollution occurs because of factors like fuel use in vehicles such as private transport, industries, and, more importantly, the burning of waste like plastic and grass. These are a few of the elements that contribute to air pollution. All of these factors have an impact on air quality. It can be calculated or determined by using PM2.5 and other variables. If this level is high, then it may be a serious problem for people's health. To control air pollution, continuous pollution monitoring is required. Here, different machine learning models like linear regression, decision tree classifiers, random forest classifiers, and KNN classifiers are used to detect whether the sample of air is polluted or not. On the basis of previous readings of PM2.5, NH₃, CO, NO, NO_x, and NO₂, SO₂ is used to predict the future values of the AQI. By collecting datasets of the daily atmospheric conditions of a specific city or town, this system is able to predict the PM2.5 level and also detect the air pollution status.

Keywords Air quality index · Linear regression · Decision tree classifier · Random forest classifier · KNN classifier · Pollution

1 Introduction

The urban population is growing at an alarming rate these days. According to the United Nations (UN), the population of urban areas in 2022 will be about 56% of the world's population [1], and this may be increased to 68% by 2050 [2].

Naga Ravindra Babu M · M. Durga Satish (✉) · B. V. Prasanthi · S. V. V. D. Jagadeesh ·
J. N. S. S. Janardhana Naidu · I. K. Pradeep
Department of Computer Science & Engineering, Vishnu Institute of Technology, Bhimavaram,
India
e-mail: durgasatish.m@vishnu.edu.in; s.jagadeesh@vishnu.edu.in; kalipradeep.i@vishnu.edu.in

The rapid growth of urbanization leads to many problems, like health and sanity. This also leads to a reduction in air quality. Resolve the issues by providing more employment in these areas and also by providing more technology through ICT (information and communication technology) tools. The issue emerges when the concentration of SO_2 , NO_2 , and residual suspended particle matter in the air exceeds their permitted limit [15].

Particulate matter can be of two types: Primarily, the particulate matter is emitted directly from wildfires, gravel pits, construction sites, agricultural forms, and dusty road transport. Secondary pollution forms from the atmosphere and chemical factories. According to the WHO (World Health Organization), almost 99% of people breathe air that exceeds WHO air quality limits and threatens their health [3].

The World Health Organisation's (WHO) most recent air quality standards (2021) propose the following concentration limits for certain pollutants [4]:

- The yearly average for $\text{PM}_{2.5}$ is 5 g/m^3 , while the 24-h average is 15 g/m^3 .
- The yearly average for PM_{10} is 15 g/m^3 , while the 24-h average is 45 g/m^3 .
- The yearly average for NO_2 is 10 g/m^3 , while the 24-h average is 25 g/m^3 .

The $\text{PM}_{2.5}$ particles, which when present in high concentrations produce clouds, can limit air visibility. To identify and estimate the amount of $\text{PM}_{2.5}$ air pollution, several machine-learning models were applied to a dataset of daily meteorological conditions.

2 A Brief Literature Review

Varsha Gopalakrishnan [5] utilized machine learning (ML) and Google Street View data to evaluate the air quality at several places in Oakland, California. He concentrated on the areas where the data wasn't present. A website application was developed by the author to forecast air quality in any city neighborhood.

Dyuthi Sanjeev [6] examined a dataset containing information on both meteorological and pollution concentrations. The Random Forest (RF) classifier, which the author claimed worked best because it is less prone to over-fitting, was used to analyze and predict the air quality.

Li et al. (2020) predicted Californian air quality in terms of pollutants and particle levels using the Support Vector Regression (SVR) machine learning technique. The authors claimed to have invented a new method for modeling hourly air pollution [7].

According to Madhuri et al. (2020), wind speed, wind direction, humidity, and temperature all have a significant effect on the concentration of air pollutants.

The author used supervised ML approaches to predict the AQI and discovered that the RF algorithm had the fewest classification errors [8].

It was found that when the input air temperature increased, NO_x emissions increased but smoke emissions fell [16].

3 Proposed Methodology

It can be used to forecast future PM_{2.5} values based on historical readings using machine learning models. The primary objective is to use the data collection to anticipate the city's degree of air pollution. For the last 2–3 decades, the Air Quality Index (AQI) has been a widely used indicator tool both globally and in India. There are six AQI levels and they are: good, acceptable, moderately contaminated, poor, extremely poor, and severe. Figure 1 shows the classification framework, in which we separated the task into two major stages.

Training Phase: The system is trained by using the data in the data set and fitting a model using the proper technique.

Testing Phase: The system is supplied with inputs and then checked for functionality. The precision is tested.

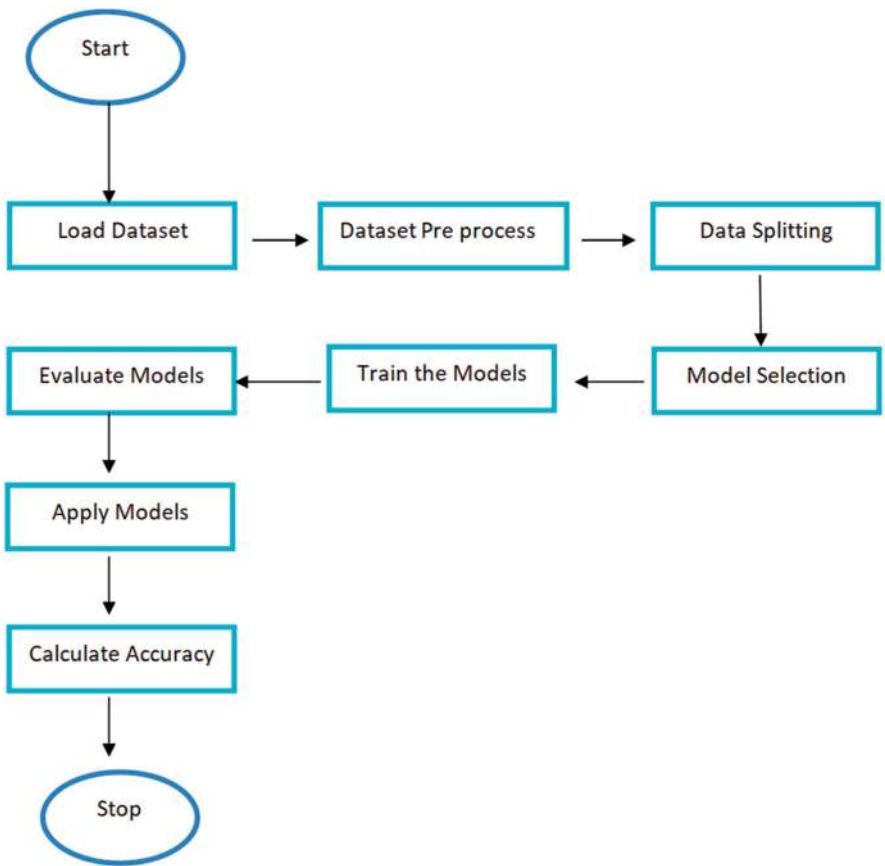


Fig. 1 Classification framework

3.1 *Materials and Procedures*

The problem of air pollution is increasing daily, and several Indian towns are among the most polluted in the world. India's poor air quality is increasingly seen as posing a substantial health danger and impeding the country's ability to advance economically.

In India, air pollution costs up to seven lakh core (\$95 billion) annually, according to recent research produced jointly by the Industrial Development Corporation and Dalberg Advisors, a nonprofit management organization located in the UK. In 2019, Dalberg concluded that air pollution in India costs up to ₹7 lakh crore (\$95 billion) annually, according to recent research produced jointly by the Industrial Development Corporation and Dalberg Advisors, a nonprofit management organization based in the UK.

3.2 *Technologies Used*

(a) **Python**

Python is a popular and user-friendly programming language for creating websites and applications, as well as doing data analysis. The unique nature of this high-level programming language allows us to quickly reuse code. Python is not confined to any one topic and may be used to create a wide range of unique and dynamic projects and programming. It is trash collected and dynamically typed [9]. It has some features like free access to computing resources, easy collaboration, and sharing.

(b) **Google Colab**

Google Colab, sometimes known as Collaborator, is a tool for writing and executing Python programs and projects. It enables us to run machine learning algorithms, as well as high-level programming and data analysis [10].

4 **Dataset Collection, Preprocessing, and Analysis**

(a) **Data Collection**

In this paper, the air pollution data was collected from Central Control Room for Air Quality Management [11] for the years April 2021 to December 2022.

The data used in this study comprised NO₂, NH₃, SO₂, CO, NO_x, and PM2.5 values. The data obtained includes around 6400 records.

The obtained data comprised of pollution levels in several Indian states. It has the following characteristics—city, date, PM2.5, SO₂, NH₃, NO, NO_x, NO₂, and CO.

Table 1 Attributes in the dataset

Attributes	Complete meaning
City	Name of the sample collected city
Date	Date of data recording
State	State/Union Territory (UT)
SO ₂ , NO ₂ , NO _x , NH ₃ , PM _{2.5} , NO, CO	The concentration of air pollutants
SO _i , NO _i , NO _x _SubIndex, CO_SubIndex, NH ₃ Sub_Index	Subindex values are used to calculate the AQI.
AQI	The computed air quality index value

After downloading several CSV files for each state and union territory, it was found that the data was uneven and that some columns, including PM_{2.5}, SO₂, NH₃, NO, NO_x, NO₂, and CO, had missing values. As a result, the data has been preprocessed to make it appropriate for our purpose.

(b) Preprocessing of Data

Preprocessing refers to changing or eliminating unclean and raw data by detecting missing or unnecessary elements of the data and subsequently, and the dataset contains the different attributes shown in Table 1.

To compute the value of AQI, authorities developed a generalized technique that must include at least three components that contribute to AQI, one of which must be PM_{2.5} or PM₁₀ [12]. Our algorithm to get our AQI level for each row was therefore constructed with limited particulars.

We added a new field called the AQI column, which is based on the values in the field and is assigned a pre-defined class from Good to severe from Table 2 (with 1 being the lowest and 6 being dangerous) [13], in order to apply classification models to the dataset.

(c) Data Analysis

In this dataset, many columns of data from across India have been collected. By using these columns' data, it has derived the AQI column. AQI was calculated by using the formulas given in Table 3 [14].

5 Implementation

The specified machine learning classifiers are used to test the model in this step. There have been a few models developed and their accuracy has been verified. The four classifiers listed below are utilized in this project.

Table 2 Air concentration range

Categories of AQI	AQI	Range of concentration							
		PM10	PM2.5	NO ₂	O ₃	CO	SO ₂	NH ₃	Pb
Good	0-50	0-50	0-30	0-40	0-50	0-1.0	0-40	0-200	0-0.5
Satisfactory	51-100	51-100	31-60	41-80	51-100	1-2.0	41-80	201-400	0.5-1.0
Moderate	101-200	101-250	61-90	81-180	101-168	2.1-10	81-380	401-800	1.1-2.0
Poor	201-300	251-350	91-120	181-280	169-208	10-17	381-800	801-1200	2.1-3.0
Very poor	301-400	351-430	121-250	281-400	209-748+	17-34	801-1600	1200-1800	3.1-3.5
Severe	401-500	430+	250+	400+	748+	34-	1600+	1800+	3.5+

Table 3 Variable data analysis

Variable	Formula
PM2.5	if $x \leq 50$: “Good” elif $x > 50$ and $x \leq 100$: “Moderate” elif $x > 100$ and $x \leq 200$: “Poor” elif $x > 200$ and $x \leq 300$: “Unhealthy” elif $x > 300$ and $x \leq 400$: “Very unhealthy” elif $x > 400$: “Hazardous”
NH3_subindex(x):	if $x \leq 200$: $x * 50/200$ elif $x \leq 400$: $50 + (x - 200) * 50/200$ elif $x \leq 800$: $100 + (x - 400) * 100/400$ elif $x \leq 1200$: $200 + (x - 800) * 100/400$ elif $x \leq 1800$: $300 + (x - 1200) * 100/600$ elif $x > 1800$: $400 + (x - 1800) * 100/600$
Noi(NO ₂)	if $(NO_2 \leq 40)$: $ni = NO_2 * 50/40$ elif $(NO_2 > 40$ and $NO_2 \leq 80)$: $ni = 50 + (NO_2 - 40) * (50/40)$ elif $(NO_2 > 80$ and $NO_2 \leq 180)$: $ni = 100 + (NO_2 - 80) * (100/100)$ elif $(NO_2 > 180$ and $NO_2 \leq 280)$: $ni = 200 + (NO_2 - 180) * (100/100)$ elif $(NO_2 > 280$ and $NO_2 \leq 400)$: $ni = 300 + (NO_2 - 280) * (100/120)$ else: $ni = 400 + (NO_2 - 400) * (100/120)$
cal_SOi(SO ₂):	if $(SO_2 \leq 40)$: $si = SO_2 * (50/40)$ elif $(SO_2 > 40$ and $SO_2 \leq 80)$: $si = 50 + (SO_2 - 40) * (50/40)$ elif $(SO_2 > 80$ and $SO_2 \leq 380)$: $si = 100 + (SO_2 - 80) * (100/300)$ elif $(SO_2 > 380$ and $SO_2 \leq 800)$: $si = 200 + (SO_2 - 380) * (100/420)$ elif $(SO_2 > 800$ and $SO_2 \leq 1600)$: $si = 300 + (SO_2 - 800) * (100/800)$ elif $(SO_2 > 1600)$: $si = 400 + (SO_2 - 1600) * (100/800)$

5.1 Machine Learning Classifiers

After the training process, the training model was analyzed, and feature sets were introduced to the untrained model. Based on these feature sets, the model generated a set of predictions, which were subsequently compared to the valid labels to assess the testing accuracy of the model.

(a) Regression

Linear regression is a linear model of the connection between one or more independent variables like AQI, AQI_Range and a dependent variable like PM2.5,

NO₂, NH₃, SO₂. When there is just one explanatory variable, the situation exhibits a Simple Linear Regression Model.

(b) Classification

Classification is a strategy for determining the class of a dependent variable like AQI, AQI_Range based on a number of different independent variables like PM2.5, NO₂, NH₃, and SO₂. The following classifiers were utilized in our project.

(i) Decision Tree Classifier

A machine learning method called Decision Tree Classifier may be applied to classification jobs. It may be used to forecast the air quality index (AQI) based on the concentrations of several air pollutants like NH₃, SO₂, PM2.5, CO, NO, NO_x, and NO₂ in the aforementioned air pollution dataset.

(ii) Random Forest Classifier

A machine learning method called the Random Forest Classifier may be applied to classification problems. It may be used to forecast the air quality index (AQI) based on the concentrations of different air pollutants like NH₃, SO₂, PM2.5, CO, NO, NO_x, and NO₂ in the air pollution dataset. Utilize measures like accuracy, precision, recall, and F1 score to assess the Random Forest Classifier model's performance on the testing set after training it on the training data.

(iii) KNN Classifiers

A machine learning approach called K-Nearest Neighbours (KNN) may be applied to prediction and classification problems. KNN may be used in the context of air pollution to forecast the quality of the air based on the concentrations of different air pollutants such as NH₃, SO₂, PM2.5, CO, NO, NO_x, NO₂, and AQI.

6 Results and Discussion

The daily data from various cities in India, such as Amaravathi, Naharlagun, Araria, Delhi, Ahmedabad, Baddi, Srinagar, Bagalkot, Eloor, Mumbai, Imphal, Puducherry, Ajmer, Agra, and Asansol, were collected from the central pollution control board. The data includes different variables, such as SO₂, NH₃, CO, NO₂, PM2.5, NO_x, and NO.

The correlation coefficients between two sets of variables are shown graphically in Fig. 2.

Some Insights:

- According to a heatmap, CO, and AQI have a significant positive linear link with a correlation coefficient of 0.78 between them.
- According to a heatmap, there is a somewhat positive linear link between PM2.5 and AQI, with a correlation value of 0.46.

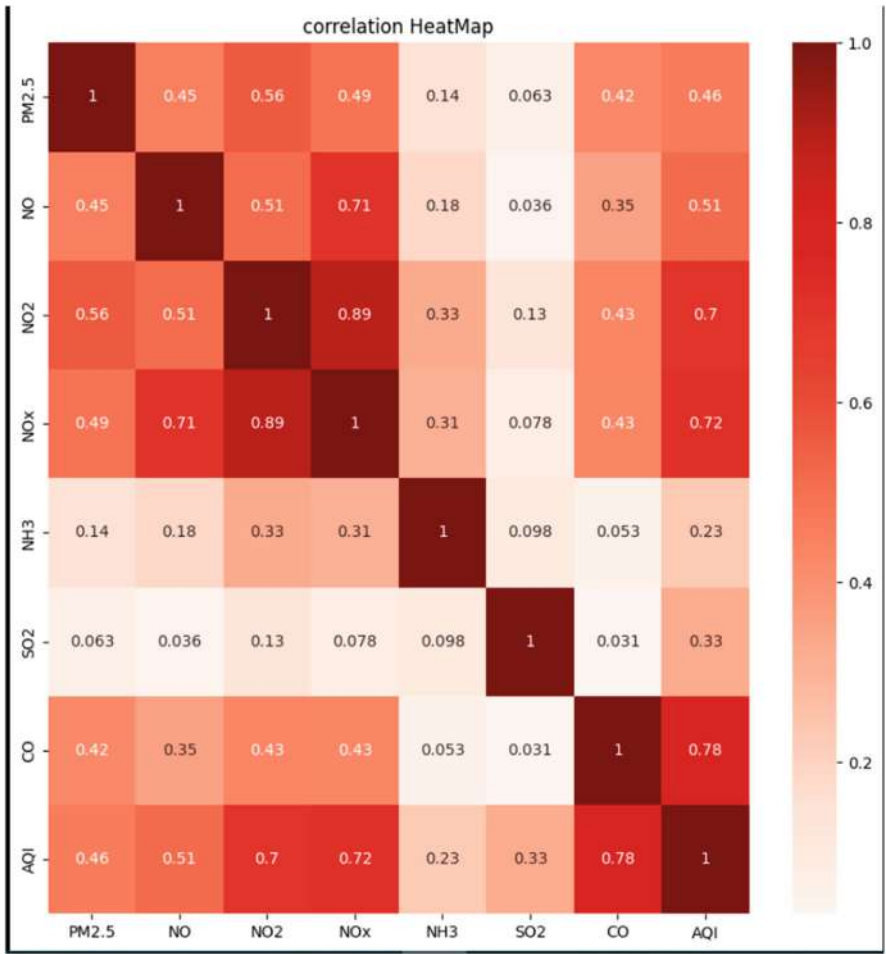


Fig. 2 Correlation heat map for air pollution dataset

- A heatmap’s SO_2 and NH_3 correlation coefficient of 0.098 indicates that there is a marginally positive linear association between these two variables.

The association between cities and PM2.5 readings is depicted in the following graph, arranged in ascending order. In this, it used the two variables, city and PM2.5, to create the bar graph. Figure 3 shows that Delhi has the highest PM2.5 rating, whereas Naharlagun has the lowest.

To ensure accuracy and robustness, this study evaluates the model against four machine learning classifiers (KNN, Logistic Regression, Random Forest, and Decision Tree Classifier). The dependent variable, which is the variable being

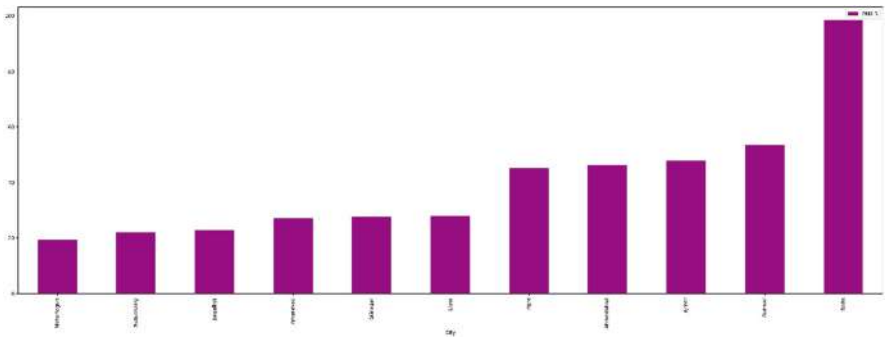


Fig. 3 Bar chart for PM2.5 and city

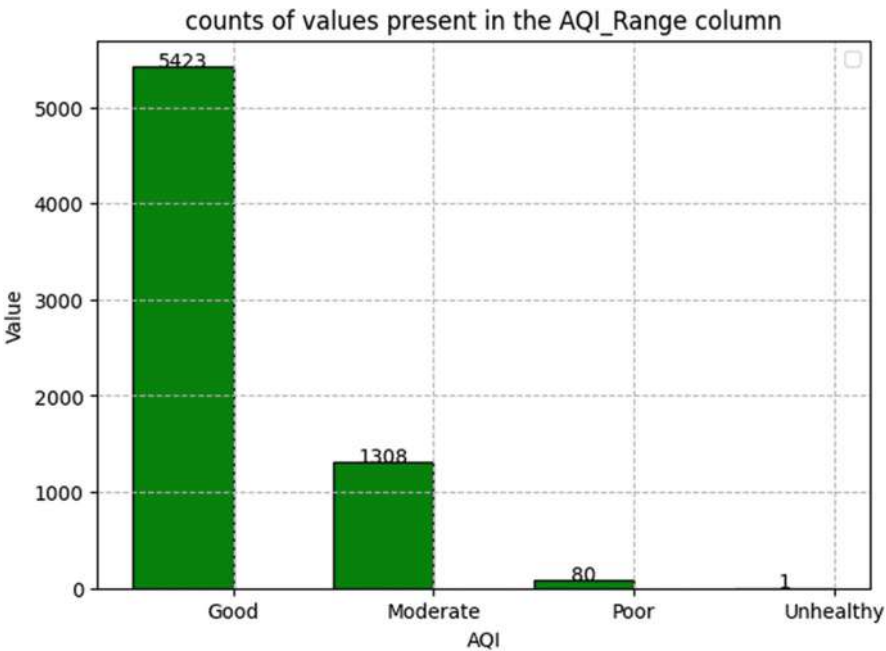


Fig. 4 Air quality index range

predicted, and the independent variables, which are the variables used to make predictions, are also known respectively.

In Fig. 4, the counts of values in the AQI_range columns for different cities of our data set are shown.

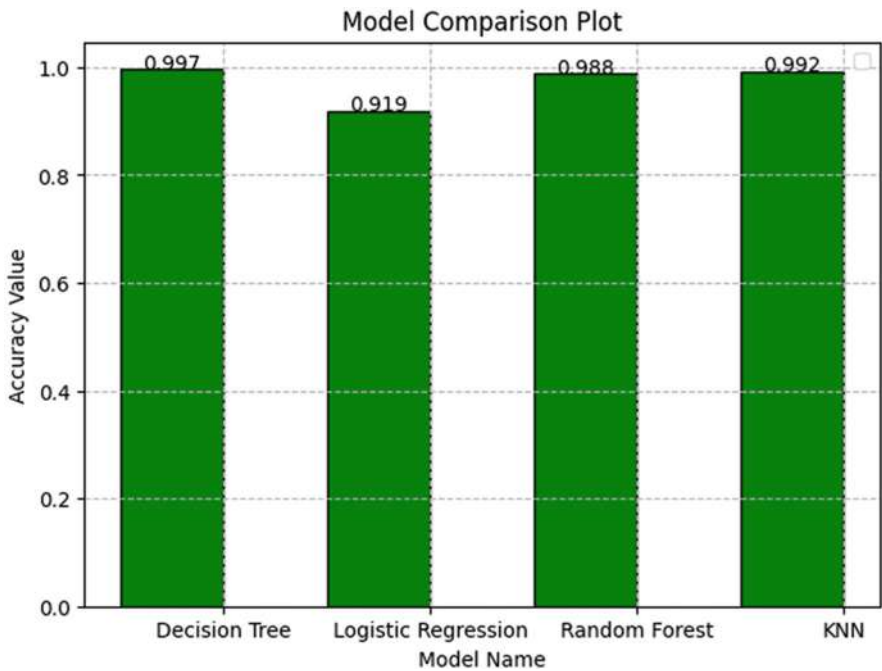


Fig. 5 Mode comparison chart

7 Comparing the Models

When compared to other machine learning models like KNN, Logistic Regression, Random Forest are applied to the data set, Decision Tree Classifier is best suited for this system, with a mean accuracy of 99.7% as shown in Fig. 5.

In machine learning, classification accuracy is a widely used metric for assessing the performance of classification models. It represents the proportion of instances that are correctly classified, relative to the total number of instances.

The calculation of classification accuracy shown in Table 4 involves dividing the number of correctly classified instances by the total number of instances in the dataset.

Comparison of Algorithms Using Evaluation Metrics

As shown in Fig. 6, the Random Forest Classifier gives the best accuracy of 99.78% of all the other classifiers.

Table 4 Classification accuracy comparison

ML Model	Variables for input	Output/results of accuracy
K-Nearest Neighbor	City, Date, SO ₂ , NO ₂ , CO, PM2.5,NOx,NH ₃ ,NO	Model accuracy on test is: 0.9822064056939501
Decision Tree Classifier		Model accuracy on test is: 0.9968861209964412
Random Forest Classifier		Model accuracy on test is: 0.9977758007117438
Logistic Regression		Model accuracy on test is: 0.9185943060498221

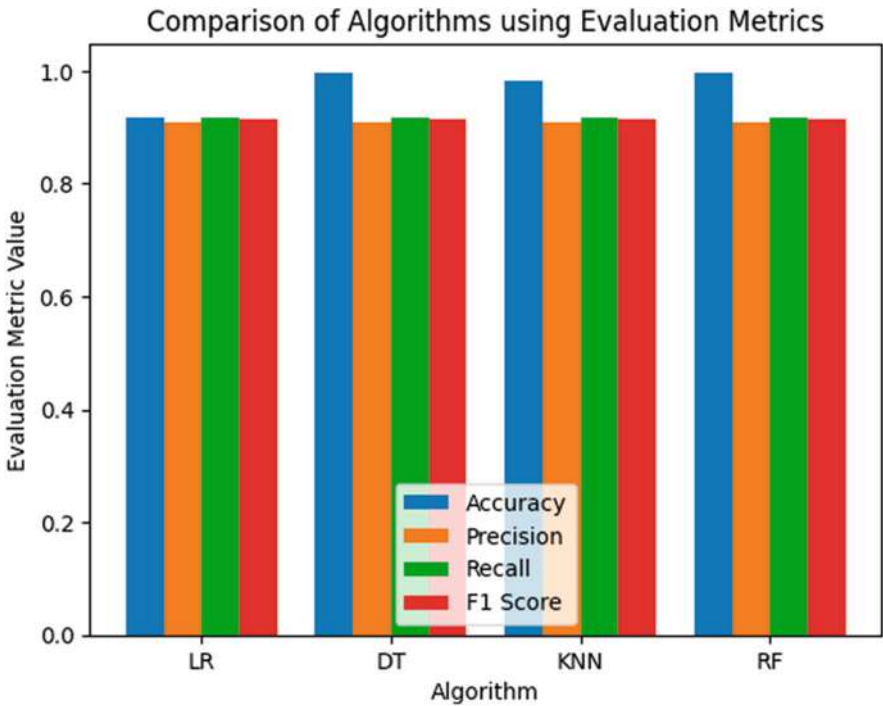


Fig. 6 Comparison of algorithms using evaluation metrics

8 Conclusion

In conclusion, machine learning models have shown great potential in detecting and predicting air pollution levels with high accuracy.

This study looked at the link between AQI and other variables that influence AQI. Our ultimate objective is to create new and more efficient models that forecast AQI with high accuracy.

These models have the ability to analyze large amounts of data from various sources, such as weather conditions and industrial emissions, to provide real-time insights into air quality. This information can be invaluable in helping to develop strategies for reducing pollution and improving public health. However, ongoing research and development is necessary to further improve the accuracy and reliability of these models. With continued efforts, we can harness the power of machine learning to create a cleaner and healthier environment for generations to come.

References

1. World Bank. (n.d.). Urban Development. Retrieved May 8, 2023, from <https://www.worldbank.org/en/topic/urbandevelopment/overview>
2. United Nations, Department of Economic and Social Affairs, Population Division. (2019). 2018 Revision of World Urbanization Prospects. Retrieved May 8, 2023, from <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>
3. World Health Organization. (2022, April 4). Billions of people still breathe unhealthy air: New WHO data. WHO. <https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data>
4. World Health Organization. (2022, September 7). WHO releases new repository of resources for air quality management. WHO. <https://www.who.int/news/item/07-09-2022-who-releases-new-repository-of-resources-for-air-quality-management>.
5. Gopalakrishnan, V. (2021). Hyperlocal air quality prediction using machine learning. Towards Data Science. Retrieved from <https://towardsdatascience.com/hyperlocal-air-quality-prediction-using-machine-learning-ed3a661b9a71>
6. Sanjeev, D. (2021). Implementation of machine learning algorithms for analysis and prediction of air quality. Int J Eng Res Technol, 2278–0181. <https://www.ijert.org/research/implementation-of-machine-learning-algorithms-for-analysis-and-prediction-of-air-quality-IJERTV8IS070087.pdf>
7. Li, Y., Zhang, H., & Qi, J. (2020). Analysis and Prediction of Air Quality Based on Machine Learning in Hangzhou. Journal of Environmental and Public Health, 2020, 8049504. <https://doi.org/10.1155/2020/8049504>
8. Madhuri, V. M., Samyama, G. G., & Kamalapurkar, S. (2020). Air pollution prediction using machine learning supervised learning approach. International Journal of Scientific and Technology Research, 9(4), 118–123. Retrieved from <https://www.ijstr.org/final-print/apr2020/Air-Pollution-Prediction-Using-Machine-Learning-Supervised-Learning-Approach.pdf>
9. Python (programming language). (n.d.). In Wikipedia. Retrieved May 8, 2023, from [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language))
10. Google. (n.d.). Frequently asked questions (FAQ). Google Research. <https://research.google.com/colaboratory/faq.html>
11. Central Pollution Control Board. (n.d.). National Air Quality Index. Government of India. <https://app.cpcbcr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/data>
12. Central Pollution Control Board. (n.d.). How AQI is calculated. Government of India. https://app.cpcbcr.com/ccr_docs/How_AQI_Calculated.pdf
13. Central Pollution Control Board. (n.d.). About AQI. Government of India. https://app.cpcbcr.com/ccr_docs/About_AQI.pdf
14. Prana Air. (n.d.). What is Air Quality Index (AQI) & How Is It Calculated? Prana Air. <https://www.pranaair.com/blogs/what-is-air-quality-index-aqi/how-is-aqi-calculated>

15. Sadhasivam, J., Muthukumaran, V., Raja, J. T., Vinothkumar, V., Deepa, R., & Nivedita, V. (2021, July). Applying data mining technique to predict trends in air pollution in Mumbai. In *Journal of Physics: Conference Series* (Vol. 1964, No. 4, p. 042055). IOP Publishing.
16. Gowthaman, S., & Sathiyagnanam, A. P. (2018). Analysis the optimum inlet air temperature for controlling homogeneous charge compression ignition (HCCI) engine. *Alexandria engineering journal*, 57(4), 2209–2214.

A Novel Parasitic Mushroom-Like Structure with High Gain Microstrip Patch Antenna for Broadband Applications



M. Sahaya Sheela, G. Syam Sudheer Babu, S. N. V. Sai Durga Prasad, and M. D. Vasanth Kumar

Abstract In this research, a design for a microstrip patch antenna loaded with a parasitic mushroom-like structure is proposed. Antenna components include a rectangular microstrip patch as the main radiating element and one or more metallic patches as parasitic elements, separated by a dielectric layer. The parasitic patches are arranged in a pattern that resembles a mushroom growth, and their dimensions are optimized to improve the antenna's performance. According to simulation results, the suggested antenna has a bandwidth of up to 20% and a high gain of up to 10dBi, which is much broader than that of a typical microstrip patch antenna. The radiation characteristics of the antenna are also improved, with a more uniform radiation pattern and higher front-to-back ratio. The suggested antenna design is appropriate for a variety of uses, including wireless communications, radar systems, and satellite communications, which call for high gain and broad coverage. The simplicity and low profile of the antenna make it easy to manufacture and integrate into various systems. Overall, the proposed parasitic mushroom-shaped microstrip patch antenna, which has broadband and high gain, is a promising solution for achieving improved antenna performance without increasing the antenna's size or complexity.

Keywords Microstrip · Antenna · HFSS · Radiation Pattern · ANSYS · S-Parameter · Parasitic Element.

M. Sahaya Sheela (✉) · G. Syam Sudheer Babu · S. N. V. Sai Durga Prasad · M. D. Vasanth Kumar

Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr Sagunthala R&D Institute of science and Technology, Chennai, India

e-mail: Vtu15496@veltech.edu.in; Vtu12690@veltech.edu.in; Vtu13513@veltech.edu.in

1 Introduction

An antenna is a piece of equipment that beams electromagnetic (EM) energy in the form that signals take out into space. The dimensions of conventional antennas are less than their operational wavelength, which, according to the formula $v=f\lambda$, corresponds to their operational frequency. Depending on the circuitry that is present in addition to the antenna, the same antenna can be utilized for transmission and receiving. An antenna receives an electric current from a transmitter to transmit signals, and the antenna subsequently radiates the energy as EM waves into space. The antenna's design affects the transmitted wave's wavelength. The antenna similarly intercepts the electromagnetic wave in the case of signal reception, creating a current in an external circuit. This current can then be amplified with the help of the receiver.

1.1 *Microstrip Patch Antenna*

Because of its low profile, simple construction, and favorable radiation characteristics, microstrip patch antennas are frequently employed in a variety of applications. However, the constrained bandwidth and gain of traditional microstrip patch antennas have turned into a bottleneck for a number of applications, including radar systems, wireless communication, and satellite communication. Numerous methods have been suggested to improve the performance of the antenna in order to get around this restriction. The employment of parasitic elements is one method for improving the performance of microstrip patch antenna [1]. These parasitic elements interact with the main radiating patch to generate additional resonant modes, resulting in a wider bandwidth and higher gain. The concept of the parasitic elements is similar to the parasitic growth of a mushroom on a host organism, hence the term parasitic mushroom-type structure.

Based on our earlier research, on the same substrate, a mushroom-shaped structure is loaded together with the two radiating edges of a conventional patch. Broad bandwidth is produced as a result of the simultaneous excitation of the traditional TM₁₀ mode and a novel quasi-TM₃₀ mode. The proposed design is suitable for various applications that require high gain and broadband coverage, such as wireless communications, radar systems, and satellite communications.

1.2 *Broadband and High Gain Antenna*

Broadband refers to an antenna's ability to operate over a broad frequency range, typically spanning multiple octaves. A broadband antenna can be used in applications where multiple frequency bands need to be covered, or where the frequency of

the signal being transmitted or received may vary over time [2]. The employment of parasitic elements is one method for improving the performance of microstrip patch antennas. These parasitic elements interact with the main radiating patch to generate additional resonant modes, resulting in a wider bandwidth and higher gain [3]. The concept of the parasitic elements is similar to the parasitic growth of a mushroom on a host organism, hence the term “parasitic mushroom-type structure” is used. Microstrip patch antennas have been actively researched and employed in a variety of products due to their enticing qualities of low profile, simplicity of fabrication, conformability to a curved surface, compatibility with integrated circuit technology, and others. The development of 5G communication has increased demand for wideband and high-gain antennas [4]. However, the limitation of conventional microstrip antenna’s narrow bandwidth limits its use in wireless communications.

1.3 Microstrip Patch Antenna

A microstrip patch antenna is a type of low-profile, planar antenna that operates in the microwave frequency range. It is a type of printed antenna that consists of a ground plane placed beneath a thin, flat, rectangular, or circular metal patch mounted on a dielectric substrate. The metal patch of a microstrip patch antenna, which is represented in Fig. 1 is typically made of a conductive material such as copper, and it is connected to a transmission line or feed network at a point called the feed point. The ground plane beneath the patch serves as a reflecting surface, and it provides a return path for the electromagnetic waves [5]. The performance properties of a microstrip patch antenna, such as its gain, radiation pattern, and impedance bandwidth, can be controlled by adjusting the dimensions of the patch and the substrate, as well as the location and size of the feed point. Many wireless communication applications use microstrip patch antenna, which is represented

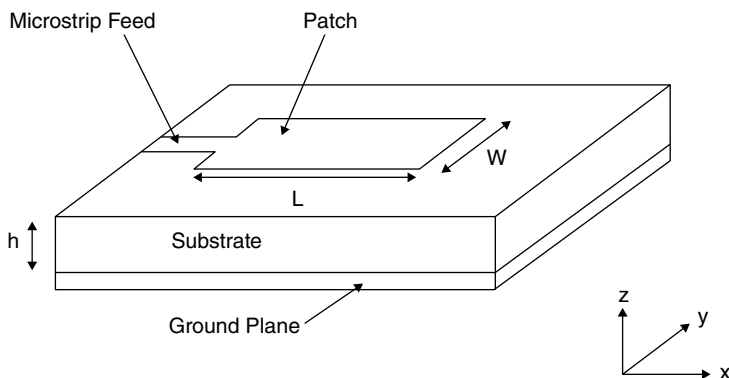


Fig. 1 Microstrip patch antenna

in Fig. 1 including cellular phones, WLANs, satellite communication, and radar systems, due to their low profile, lightweight, and easy-to-fabricate nature [6].

1.4 Parasitic Mushroom Type Structure

A parasitic mushroom-type structure is a type of additional element that can be incorporated into the design of an antenna to improve its performance characteristics. It is typically used in microstrip patch antennas and consists of a mushroom-like structure placed on top of the patch [7]. The mushroom structure acts as a reflector, which could improve the gain and directivity of the antenna. It reflects the electromagnetic waves back toward the patch, which causes them to interfere constructively and enhance the radiation in the desired direction. The mushroom structure's dimensions, shape, and location can be optimized to achieve the desired performance characteristics [8].

1.5 Advantages of Using Broadband and High Gain Microstrip Patch Antenna Loaded with Parasitic Mushroom-Type Structure

The parasitic mushroom-type structure offers several advantages for broadband and high-gain microstrip patch antennas. Some of the advantages are increased gain, improved bandwidth, better impedance matching, reduced back radiation, and simple design [9]. Overall, the use of a parasitic mushroom-type structure in broadband and high-gain microstrip patch antennas can significantly enhance their performance characteristics and make them a more effective solution for wireless communication applications [10].

1.6 Applications of Broadband and High Gain Microstrip Patch Antenna Loaded with the Structure Resembling Parasitic Mushroom Type

The broadband and high gain microstrip patch antenna loaded with the structure resembling parasitic mushroom type has applications, such as satellite communication, wireless communication, radar systems, radio astronomy, military applications, and medical applications, in many fields where high-performance antennas are required [11].

2 Methodology

2.1 HFSS Software

Ansys created the potent electromagnetic simulation software program known as HFSS (High-Frequency Structure Simulator), which is frequently used to simulate high-frequency electromagnetic fields in intricate structures including antennas, microwave parts, and RF circuits. HFSS uses a 3D electromagnetic solver to simulate the behavior of electromagnetic fields in complex structures. The software supports a wide range of simulation types, including planar, linear, and 3D simulations, as well as parametric and optimization studies. One of the key features of HFSS is its ability to handle large and complex geometries, such as those found in integrated circuits, antennas, and microwave components. It can simulate structures with thousands of components and accurately model the behavior of electromagnetic fields in these structures. Antenna parameters such as S Parameters, Resonant Frequency, and Fields can be calculated using Ansoft HFSS.

2.2 ANSYS

ANSYS is a simulation software suite that is widely used in engineering, design, and product development. The software suite includes a range of tools for simulating the behavior of complex systems and materials, including electromagnetic field modeling, computational fluid dynamics, and finite element analysis (FEA). Engineers and designers use the ANSYS software suite to model and improve the performance of goods and systems prior to their production or deployment. For example, ANSYS can be used to simulate the behavior of a car crash to optimize the design of car components for safety, or to simulate the airflow around an aircraft wing to optimize its aero-dynamic performance.

2.3 Antenna Design for a Parasitic Mushroom-Type Structure-Loaded with Broadband and High Gain Microstrip Patch Antenna

Parasitic mushroom-type antenna is designed using ANSYS in which the top view and side view of antenna design we obtained are represented in Fig. 2.

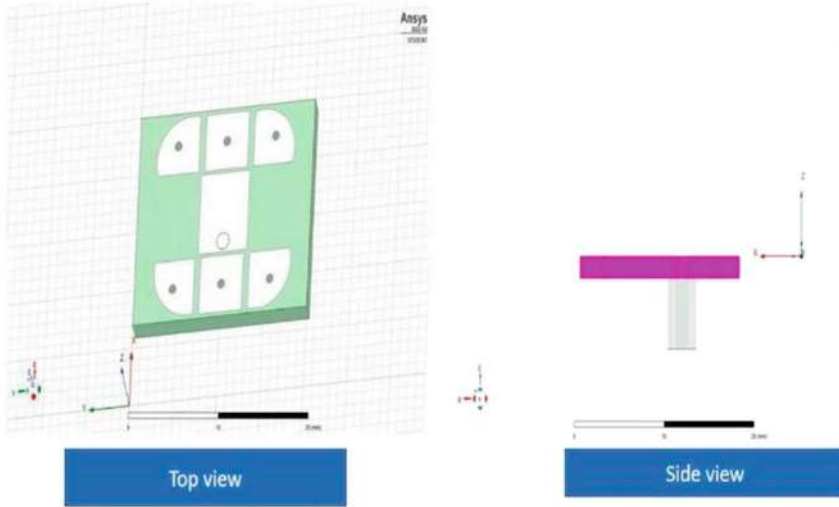


Fig. 2 Antenna design

2.4 Limitations of Metamaterial Antenna

Metamaterials are typically included in antennas to produce complicated designs that may be difficult to manufacture [12]. Additionally, the complexity of the design is increased by the smaller dimensions of the metamaterial designs compared to the antenna [13]. The production of these metamaterial antennas on a wide scale is constrained by this restriction. Since there are more factors involved in determining the operational wavelength, optimizing metamaterial antenna designs becomes challenging [14].

2.5 Comparison of Conventional and Suggested Antennas

Because it is thin, simple, and inexpensive to build a conventional microstrip patch antenna, and it is easy to mount in both planar and non-plane settings [15]. Additionally,, when applied to printed circuit technologies, it is appropriate for the combination of integrated circuits. However, it has a limited bandwidth between 1 and 10 percent due to its high Q value [16]. To address this issue, we used parasitic patches to extend the bandwidth beyond 500 MHz.

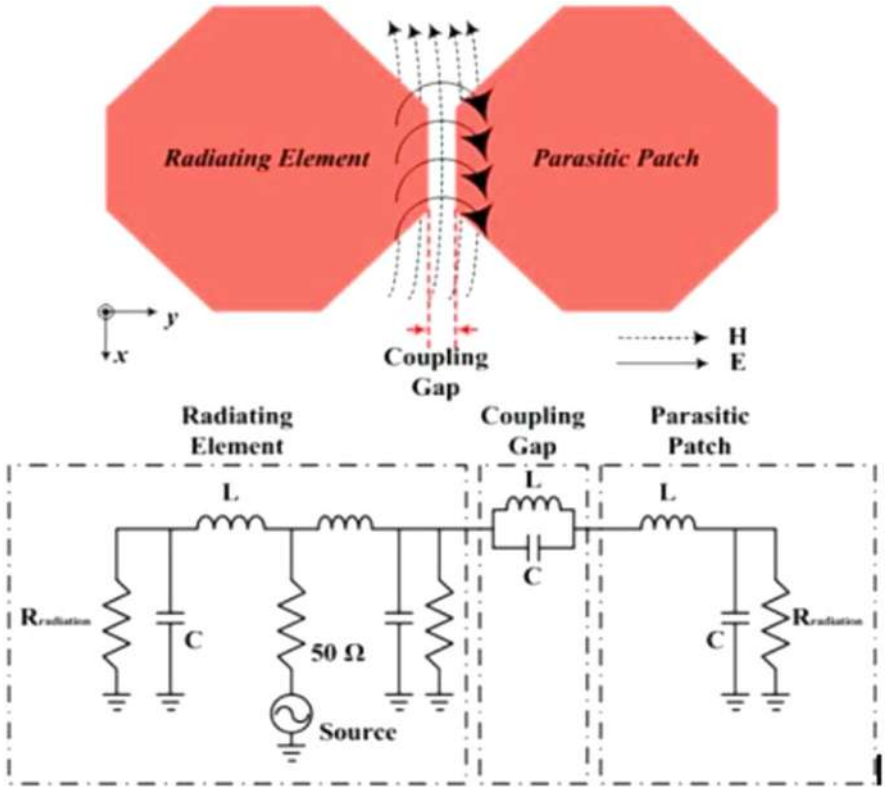


Fig. 3 Equivalent circuit between the radiating element and parasitic patch

2.6 Comparison of Parasitic Patch

The parasitic patch’s and the radiating element’s resonance frequencies correspond to those of L and C. Each element’s length can be represented as a value of L in a series, and the value of the series L increases as the patch’s length does. Additionally, the patch’s overall size grows as the series L value does. The E field between the patch and the earth is strengthened as a result. As a result, the ground plane and patch are where the high shunt C is located. The growth of e-fields in both the core patch and the parasite patch is shown in Fig. 3.

3 Results and Discussion

To validate the results of the simulation, a prototype of the proposed antenna was designed. The geometry of the mushroom-type loaded patch antenna was optimized

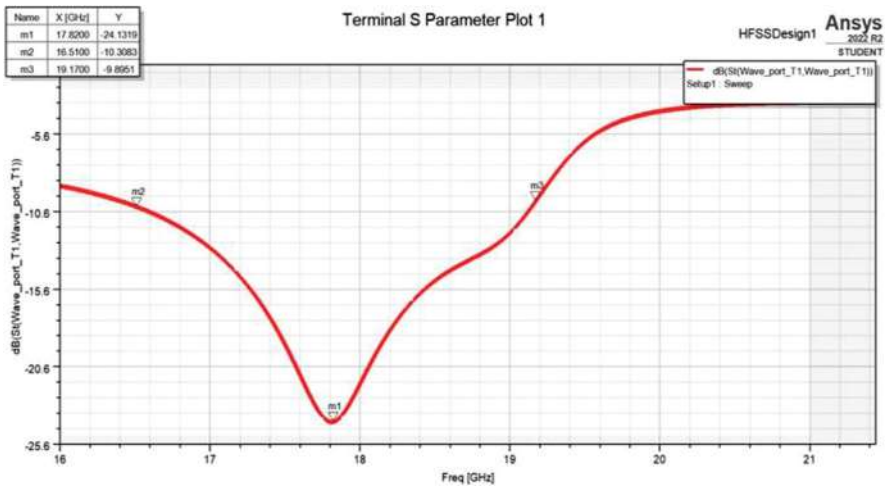


Fig. 4 Terminal S parameter plot 1

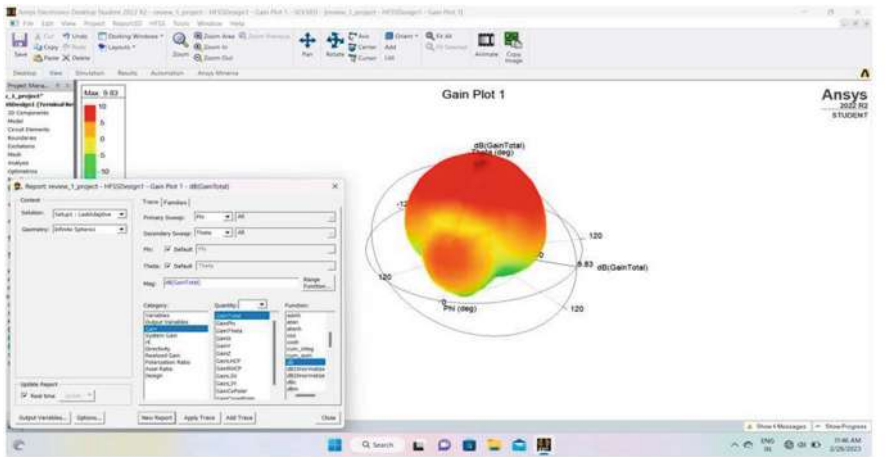


Fig. 5 Gain

using the HFSS software. Figure 4 shows that S-Parameter plot 1 for various frequency ranges. Figure 5 shows the gain plot for mushroom-type structure antenna in which gain is maximum in the broadband frequency range.

The radiation pattern for this proposed antenna is shown in Fig. 6, which shows that minor lobe is neglected and major lobe coverage area is maximum on the front side.

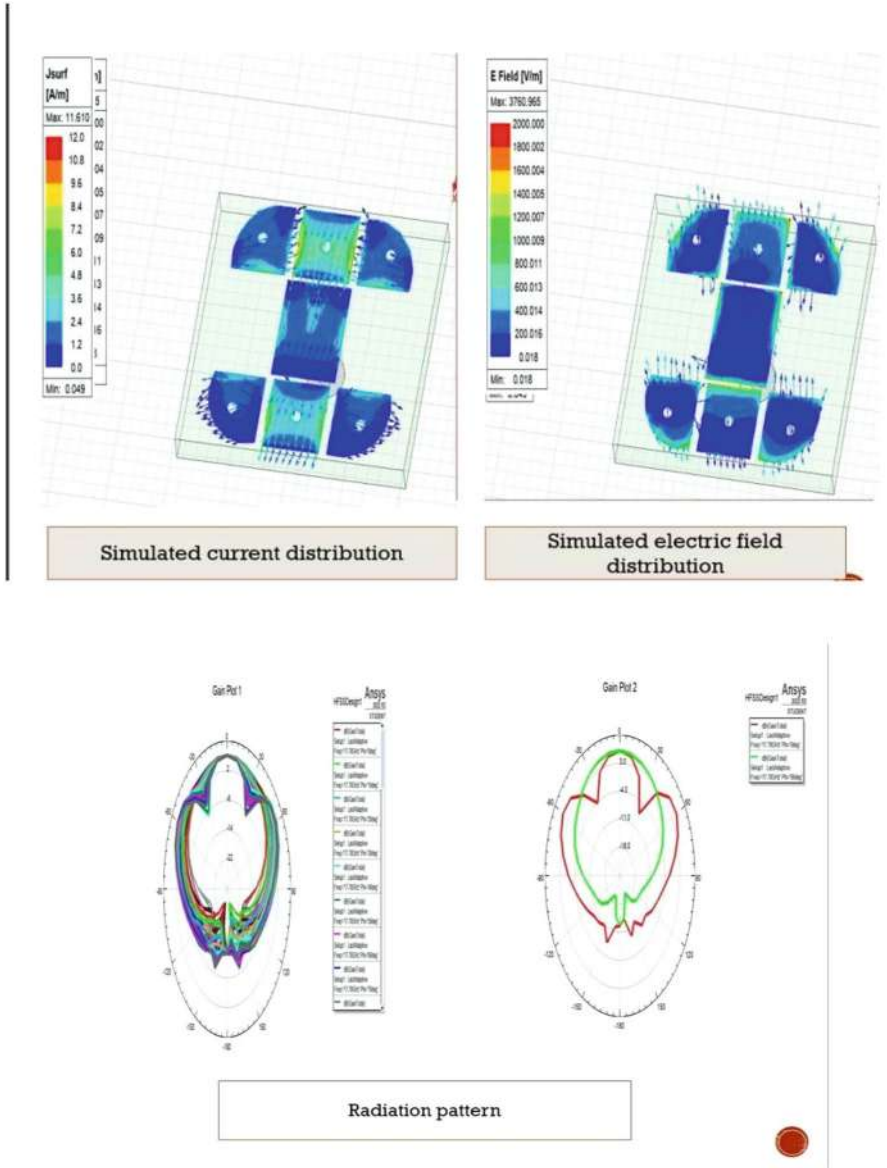


Fig. 6 Radiation pattern

4 Conclusion and Future Work

Due to the newly developed quasi-TM30 mode, the use of a mushroom-type structure in combination with two radiating edges of a primary radiating patch has

produced a wide impedance bandwidth characteristic. According to the measured data, the improved impedance bandwidth, which spans the entire Ku-band, is greater than 40%. The antenna design proposed in this study demonstrates consistent radiation patterns perpendicular to its axis throughout its frequency range, providing high gains up to 10dBi. In summary, incorporating a parasitic mushroom-like structure onto a microstrip patch antenna has proven to enhance its overall performance with respect to bandwidth and gain. The proposed design methodology entails the deliberate selection of substrate material, precise optimization of dimensions and placement of both the primary and parasitic patches, and rigorous simulation and experimentation to attain the desired outcomes. The implementation of a broadband and high-gain microstrip patch antenna, supplemented with a parasitic mushroom-like structure, can provide an effective solution for compact and efficient wireless communication systems. This antenna design is versatile and can be customized for diverse frequency bands and substrate materials, rendering it a viable option for meeting unique application demands of a microstrip patch antenna along with a parasitic mushroom-like configuration has been found to be a successful method for boosting the antenna's capabilities and satisfying the requirements of current wireless communication systems. In the future, this software simulation can be implemented in hardware by means of fabrication of antenna, and it will be used for various broadband applications.

References

1. Kai-Fong Lee and Kin-Fai Tong, "Microstrip Patch Antennas Basic Characteristics and Some Recent Advances", in Proceedings of the IEEE, July – 2012.
2. Yufan Cao et al., "Broadband and High Gain Microstrip Patch Antenna Loaded With Parasitic Mushroom-Type Structure", in IEEE Antennas and Wireless Propagation Letters, July 2019.
3. Yongtao Jia et al., "Slot-coupled broadband patch antenna", in Electronics Lett., MARCH – 2015.
4. K.P. Ray et al., "Hybrid-coupled broadband triangular microstrip antennas", in IEEE Trans. Antennas Propag., JANUARY – 2011.
5. Amit A. Deshmukh and K.P. Ray, "Compact broadband slotted rectangular microstrip Antenna", in IEEE Antennas Wireless Propag. Lett., – 2009.
6. Sang-Hyuk Wi et al., "Wideband Microstrip Patch Antenna With U-Shaped Parasitic Elements", in IEEE Trans. Antennas Propagation., April 2007.
7. Hyunseong Kang and Seong-Ook Park, "Mushroom metamaterial based substrate integrated waveguide cavity backed slot antenna with broadband and reduced back radiation", in IET Microwave. Antennas Propag., NOVEMBER – 2016.
8. Wei Liu et al., "Metamaterial-Based Low-Profile Broadband Mushroom Antenna", in IEEE Trans. Antennas Propag., March – 2014.
9. Yufan Cao et al., "Wideband and high gain patch antenna loaded with mushroom-type metamaterial", in Proc. Int. Conf. Microwave and Millimeter Wave Technology (ICMMT), May – 2018.
10. Zhao Wu et al., "Dual-band antenna integrating with rectangular mushroom-like superstrate for WLAN applications,|| in IEEE Antennas Wireless Propag. Lett., – 2016.
11. V.P. Sarin et al., "A wideband stacked offset microstrip antenna with improved gain and low cross polarization", in IEEE Trans. Antennas Propag., April – 2011.

12. Yang Cai et al. Design of low-profile metamaterials-loaded substrate integrated waveguide horn antenna and its array applications”, in IEEE Trans. Antennas Propagation. July – 2017.
13. Wei Liu et al., “60-GHz thin broadband high-gain LTCC metamaterial-mushroom antenna array” in IEEE Trans. Antennas Propag., SEPTEMBER – 2014.
14. Yuandan Dong and Tatsuo Itoh, “Metamaterial-Based Antennas”, in Proceedings of The IEEE, July – 2012.
15. S. Gopalakrishnan et al., “ Design of Power Divider for C-band operation using high frequency Defected Ground Structure (DGS) Technique”, Int. J. Simul. Syst. Sci.Technol., vol. 19, no. 6, pp. 1–7, 2018, doi:<https://doi.org/10.5013/IJSSST.a.19.06.21>.
16. Chi Yuk Chiu et al.” Study of a small wide-band patch antenna with double shorting walls”, in IEEE Antennas Wireless Propag. Lett., 2004.

Facial Emotion Recognition Using Artificial Intelligence



G. Sateesh, Swaroop Sana, S. V. R. Vara Prasad, and Bosubabu Sambana 

Abstract Image processing can be used to solve a variety of problems and has numerous real-world applications, such as human–machine interactions, robotics, computer animations, and so on. Acknowledgment of Looks has been really difficult for a long time and is as yet being explored, and in the investigation of simulated intelligence, acknowledgment of facial demeanors is a significant issue. Our framework is intended for an individual to be free and independent to recognize the seven different widespread sentiments, namely, trepidation, outrage, satisfaction, shock, pity, disdain, and unbiased. The framework is coordinated to integrate static and dynamic sources of info either from webcams or outer films like recordings or photographs and can give strength and high exactness over the outcomes. Therefore, we classify the images using the multilayer perceptron and detect the face using the haar-cascade method and the hog-face detection algorithms from Viola and Jones for the best results. Data augmentation is used to improve the data’s optimization, and facial features are extracted from the facial region to improve the efficiency of the detection system.

Keywords Image processing · Artificial intelligence · Haar-cascade method · Hog-face detection · Multilayer feed-forward neural network model

1 Introduction

Face expressions are necessary for social communication because communication can be verbal or nonverbal. Because the face conveys prominent expressional signals

G. Sateesh (✉) · S. Sana · S. V. R. Vara Prasad
Lendi Institute of Engineering and Technology (A), Vizianagaram, JNTU GV, Vizianagaram,
Andhra Pradesh, India

B. Sambana
Department of Computer Science and Engineering (Data Science), School of Computing, Mohan
Babu University, Tirupathi, Andhra Pradesh, India

such as eye contact, lip movement, face shape, and so on, facial expressions are well-known as a form of nonverbal communication. Another form of nonverbal communication is body language gestures [1]. Faces and facial expressions are easily discernible by humans. However, it is still challenging to create a robot or other automated system that achieves the same level of comprehension. The extraction of various expressions from the face, identification of facial landmarks, and expression categorization are just a few of the issues that are associated with this facial expression recognition. Other issues include illumination and variations in head positions. The field of artificial intelligence (AI) is currently conducting research into facial expression recognition, which has potential applications in a plethora of domains, including fraud detection, monitoring, and security. It has been determined that certain facial expressions have global significance. Facial Action writing was developed and refined by Ekman and Friesen, who discovered that there are six facial expressions that appear to be universal across cultures, including fear, rage, joy, surprise, sadness, and disgust [2]. Challenges in exploration like Kaggle's facial acknowledgment present these equivalent feelings alongside the impartial inclination.

A machine or robot with artificial intelligence is given the ability to make their own decisions. Normally, a human mind can envision pictures and have a memory to store prepared information, yet for a robot, we use AI for preparing outside datasets so that it can perceive as per preparation, and we use picture handling to deal with the pictures in a way that a machine can imagine, and the pictured information can be utilized in acknowledgment process by a robot. We utilize man-made reasoning over the framework to prepare and track down the feelings of the human face.

There are a variety of approaches to expression recognition that utilize inputs from a variety of fields, such as signal processing, speech processing, computer vision, machine learning, and computer vision [3]. The process involves scanning emotions across the face in real-time scenarios. There are three main steps that need to be taken in order to extract facial expressions from the face: choosing datasets, building a look acknowledgment framework, and making a graphical UI that can take static or dynamic contributions from the client. There are different datasets that can be utilized to prepare a model for acknowledgment of look like Influence Net, FER2013 Determine, Google look examination dataset, Visionary, Emotic, K-EmoCon, Expanded Cohn-Kanade dataset (CK+), and so on. Additionally, there are distinct working and training algorithms for each dataset, each of which has a different accuracy over the system in various scenarios [4].

Production of graphical UI should be possible in various applications by utilizing various libraries like PyQt5, Tkinter (TK), Kivy, Wxpython, PySide, and so on, and a portion of the tool compartments give outsider application for simplified activities for making GUI and the UI document can be switched over completely to .py record, for instance, QtDesigner. Furthermore, we can likewise make UIs by adding graphical UI libraries to Python and adding gadgets like marks buttons and so on automatically. This makes the graphic user interface more dynamic and allows for easier optimization than a drag-and-drop approach.



Fig. 1 Various facial expressions

An artificial facial expression recognition system capable of recognizing the seven universal emotions is the focus of this project—fear, rage, happiness, surprise, sadness, disgust, and neutral inputs (Fig. 1) from either a webcam or external videos and pictures. Building a look acknowledgment framework includes three stages. They are to confront recognition, removing highlights, and order of feelings. For face detection and classification, the system makes use of a multilayer feed-forward neural network model.

2 Related Work

It is possible to detect faces by taking external inputs from humans in digital formats such as .jpg, .png, or .mp4 files, processing the frames in input data, and producing facial emotions out of it. Expression recognition is a process of scanning emotions over the face in real-time scenarios and leveraging inputs from numerous fields, such as signal processing, speech processing, computer vision, machine learning, and deep neural networks [5]. Emotion recognition is performed using a variety of methodologies, and leveraging inputs from A variety of datasets, including Affect Net, fer2013, Ascertain, the Google facial expression comparison dataset, Dreamer, Emotic, K-EmoCon, and the Extended Cohn-Kanade dataset (CK+), can

be utilized to train a facial expression recognition model. What is more, each dataset has different working and preparing calculations with various exactnesses over the framework in various situations [6].

A method of dimensionality reduction known as feature extraction involves effectively representing a large number of pixels in an image in order to effectively capture interesting parts of the image or frame. Additionally, a variety of approaches, such as support vector machine [7], linear decrement analysis, K-nearest neighbors, and random forest [8], can be used to classify emotions.

Mayyadah et al. [9] proposed a framework that assesses the exhibition of look in light of chosen highlights by chi-square, utilizing different characterization methods, for example, Choice tree, KNN, radial base capability, SVM, multilayer perceptron (MLP), and random forest. When the CK+ dataset is used, the random forest classifier has a good accuracy of 94.23%.

A system proposed by Sheskar Singh et al. [10] involved face detection and illumination rectification by extracting prominent facial features such as the eyes, nose, jaw, mouth, and eyebrows. The system achieved a test accuracy of 61.7% on the FER 2013 dataset.

Feifei Zhang et al. [11] proposed a model in view of a generative ill-disposed network that consequently creates facial pictures with particular stances and articulations. They took into account the SFEW and Multi PIE datasets. The average level of recognition accuracy is 81.20%. With an accuracy of over 89%, the six expressions of happiness and surprise are easily recognized.

The neural network phase uses the Manhattan and Euclidean distinctions to identify the six senses proposed by Latifa Greche et al. [12]. Contrasted with the Manhattan range, the Euclidean reach recognizes looks with superb accuracy and a couple of stowed-away neurons. A high-level multifaceted brain network with one secret layer is used to isolate information. Good Euclidean face recognition standards are achieved in the early stages of training.

Deng et al. [13] proposed a model that perceives conduct various ways of identifying behaving like hand shaking and strolling data identification of facial recognition with ANN and CNN techniques.

Liu et al. [14] suggested a method for bridging the semantic gap between low-level and high-level information by learning micro-action pattern (MAP) representations with a deep learning network.

To lessen the effect of individual characteristics on looks, Zibo Meng et al. [11] introduced CNN, which recognizes ownership and collects information related to both phrases and identities. Speech-related presentations are studied using softmax loss and various sensitive losses. A new useful layer with various sensitivity losses is used to investigate proprietary submissions in order to enhance the diversity brought about by various identities. Based on the test results, the proposed IACNN model performs significantly better than CNN methods and numerous modern methods that enhance the dynamic data collected from image tracking.

Asit Barman et al. [15] proposed that a functional visual model used in facial imaging was used to identify landmarks on a person's face. A grid of local symbols determines the distance and shape signatures. To expand the feature set, mathematical values are derived from the shape signature and standard range. A multilayer

perceptron provides output features for various speech stages. The recognition and comparison of cutting-edge reporting demonstrates the effectiveness of the proposed system. The combined D-S signature-based facial expression surpasses previous approaches.

A model for facial expression recognition that makes use of a neural network architecture and major attributes such as age, gender, and pose was proposed by Zhanpeng Zhang et al. They used an expression recognition system similar to the Siamese framework model to understand how to map the raw pixels from pair face photos to related attributes from beginning to end in order to predict interpersonal relationships. It was discovered that the proposed model could extract the mutual context of faces for precise, fine-grained interpersonal prediction. Based on extensive research, Fengping et al. [8] have proposed a facial recognition technique. Flexible model parameters based on a multilayer maxout line function are recommended for Conventional Neural Networks and Long-Short Term Memory implementation. The in-depth learning model has been trained with great success using the initial flexible model parameters that were proposed.

3 Proposed System

There are a few phases engaged with the execution of a fake acknowledgment of looks framework. After extracting datasets, we train a model to use the datasets to differentiate facial expressions. The main body of the project is broken down into three stages: image processing, facial feature extraction, and emotion detection.

The proposed architecture for the facial expressions system is discussed in this section and shown in Fig. 2.

The input images are processed during the image processing stage, and the processed images will be used as input during the facial feature extraction stage. In the facial element extraction stage, the face is distinguished, facial highlights like eyes, nose, mouth, and jaw are removed, and the extricated information will be in the contribution for appearance acknowledgment stage, and in the demeanor acknowledgment stage, the prepared model is tried with separated data sources and we will get characterized feelings over the face in the framework. Finally, the fabricated framework can give different choices for taking information sources. UI contains two buttons where one button can take dynamic information sources like webcam feed and another button can take static picture inputs like .jpg, .png records, or video film like .mp4.

3.1 Pre-handling

Pre-processing is the main step in system building, and it improves facial expression recognition [9] before the feature extraction stage. We are using the Haar-cascade classifier, which employs the Viola and Jones algorithm, to detect the face and crop

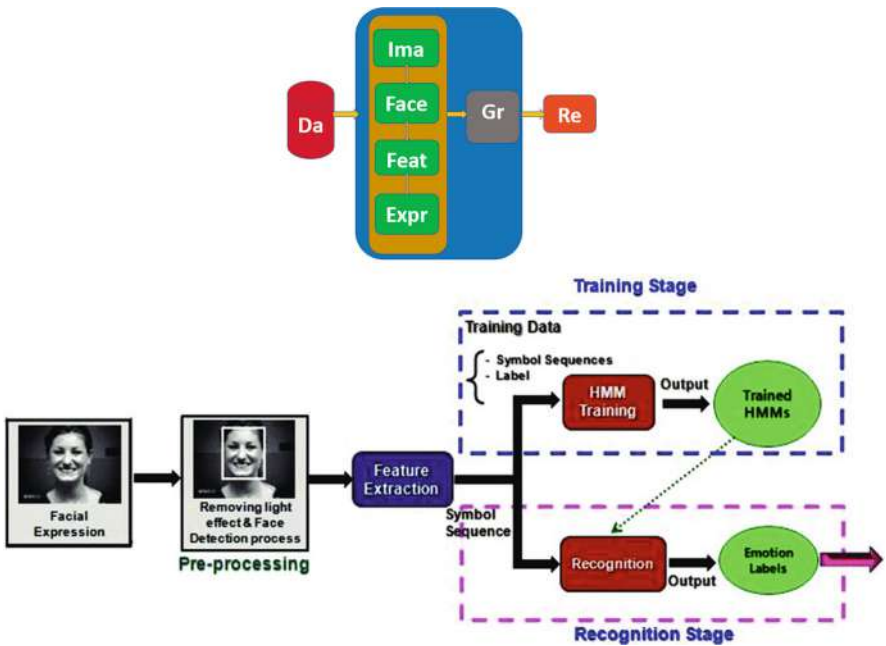


Fig. 2 The proposed system architecture

it into a rectangular area so that image processing will be more accurate because we will eliminate pre-processing over non-facial regions, as shown in Fig. 5. This study of face detection systems aims to capture the majority of the face region.

3.2 *Extracting Attributes*

The process of removing the most important features from a face in order to accurately predict an individual’s expression is known as feature extraction. To speed up computation, feature extraction is used to break down a lot of data from an image into a smaller set. We are utilizing swine face identification for neural network analysis, which is an observable method for real-time data analysis.

This approach is part of Milestone Discovery Analysis, which uses the DLib facial landmark finder. It recognizes key facial areas such as the eyes, nose, and the internal and external lines of the mouth and jaw, as displayed in Fig. 3. These areas can be highlighted with any color, with blue being used here to indicate high contrast. This technique aids in the prediction of facial expressions by extracting silent landmarks that are crucial for detecting facial emotions.

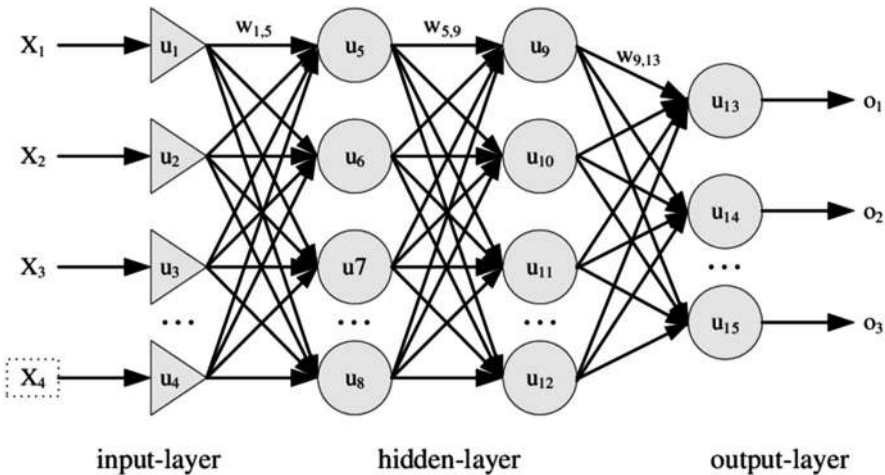


Fig. 3 Architecture of the MLP network

3.3 Landmark Detection on the FER Dataset

0101	8188	8101	034d	38f1	811d	039f	d6a4
8119	03fa	a285	811c	03a1	92e4	8119	033d
2fbc	811c	0397	0792	8118	0319	0793	811b
0367	28b1	8118	0317	52f6	811b	0340	11ce
8118	036f	67c7	811a	0385	23e7	8118	03d8
5692	8119	0387	0efc	8118	036f	cac5	8119
0358	6d86	8117	032c	0080	8118	0330	ec88
8117	033a	1b9d	8118	0344	6d86	8117	0327
d5b6	8118	0379	0efc	8118	03a3	26ce	8118
039b	23e7	8118	03b1	36e1	8118	034e	11ce
8118	0372	a0ed	8118	0377	28b1	8118	0379
3ef4	8118	03b1	0792	8118	0357	a7f7	8118
03a5	92e4	8119	03e4	77f8	8118	0342	d6a4
8119	03f0	15f4	811b	0378	73ec	811a	03f6
c4b5	811a	0375	abb6	811a	030d	0985	8119
0317	a4a6	811a	030c	afb0	8119	03bc	13b3
811a	03fd	8cd9	8119	0362	0ed5	811a	035d
3c93	8118	0362	0dd5	811a	0364	aba7	8118
039d	12b3	811a	0374	7ebd	8118	0339	a3a6
811a	03d2	91d2	8118	035e	aab6	811a	0369
80e1	8118	0328	73ec	811a	0359	0180	8118
03a8	709c	8119	0319	0180	8118	0389	7fc5
8119	0304	0180	8118	039e	31ee	8119	03bd
0080	8118	03f9	098c	8118	0367	1bcf	8119
032f	339a	8118	0380	97e6	8119	0319	869e
8118	03d9	0080	8118	03bd	37a2	8118	0328

3.4 Multilayer Perceptron

The basic unit of a single-node neural network is called the perceptron. It is made up of inputs x_1, x_2, \dots, x_n and the corresponding weights. The Multi-Layer Perceptron (MLP) network consists of an input layer, a hidden layer, and an output layer [10]. The multi-layer perceptron incorporates the function of activating the line across all of its inputs and uses the essence of the distribution method back in its training. The activation function combines input to neurons and weights and also adds bias to the output. The structure of the MLP network is shown in Fig. 3.

The training algorithm involves the following steps:

1. Initialize small random values to weights. Input unit x_i ($i = 1, \dots, n$), where n is the number of the input receives signal and transmits this signal to each hidden units z .
2. Each hidden unit calculates the tanh activation function and sends its signal z_j to each output unit.
3. Each output unit y_k ($k = 1, \dots, m$) sums its weighted input signal as and linear activation is applied to calculate the output signals as $y_k = f(y_i n_j)$ where f is linear activation function. The activation function maps the weighted inputs to the output of the neuron.
4. The output y of any neuron in the MLP can be calculated by

$$y = \left(\sum_{i=1}^n w_i x_i + bs \right) = \phi \left(W^T X + bs \right)$$

During the experiment, a high amount of neuron output was assigned to six primary subjects. The performance of the recognition system is measured by the level of recognition in the test set. MLP tests a person's ability to observe multiple expressions of face. To make it easier, expressions of fear, anger, joy, surprise, sadness, and disgust are portrayed as Fe, An, Ha, Su, Sd, and Di. Anger (AR), fear (FR), disgust (DS), sadness (SD), happiness (HP), and surprise (SR) are the six key expressions that the perceptron can distinguish.

As shown in Fig. 3, the FER system takes the data from the machine and does face detection, extraction of features, and classification of expression over the input content and shows the output via the graphical user interface.

System Workflow

Step 1: Start.

Step 2: Load the dataset in the project directory in .csv file format.

Step 3: Features are extracted from the dataset and the results are stored in the data folder within each outer_labels store inner_labels, according to the pseudocode.

```
df = pd.read_csv('./fer2013.csv') #Reading CSV file to panda's data frame
```

Step 4: By using the generated data with NumPy and modules of TensorFlow library, we will create a model and store it as model.h5 file in project folder.

```
model.saveweights('model.h5')
```

Step 5: As a part of facial expression recognition, we consider haar-cascade classifier by downloading haar-cascade_frontalface_default.xml from osdn.net website.

Step 6: To extract features from the face, we use hog face detection algorithm from landmark detection algorithm, which can be found as shape_predictor_68_face_landmarks.dat from osdn.net website.

Step 7: By using PyQt5 libraries, we created a dynamic user interface with multiple buttons.

Step 8: As of now, we loaded all the files needed for facial expression recognition system.

Step 9: We are using OpenCV library to read faces from webcam or we can use external photos or video clips for expression analysis.

```
cap = cv2.VideoCapture(0) || # for webcam feed
cap = cv2.imread(Path) || # for image input
cap = cv2.VideoCapture(Path)# for video footage as input
```

Step 10: Then by using dlib library with haar-cascade classifier and hog face detector files that we downloaded earlier, we will detect the face and important facial regions.

Step 11: Load the model.h5 file to use the trained model to predict expressions from the inputs.

```
model.load_weights('model.h5')
```

Step 12: Then by using this model, we will predict the emotion of the face and print on the screen.

```
predicted_method = model.predict(cropped_img)
max_index = int(np.argmax(predicted_method))
```

Step 13: Stop.

4 Result Analysis

The facial expression recognition of our model is validated using MUG [35], FER 2013, JAFFE [14], and Cohn-Kanade (CK+) [12] databases to justify the different facial expressions. The results obtained were compared with the prior techniques to determine the effectiveness of the proposed method.

4.1 Dataset Description

The FER 2013 dataset was sourced from the Kaggle website and comprises a total of 35,887 face crops. The dataset is divided into training (28,709 samples), testing (3589 samples), and validation (3589 samples) subsets. The dataset incorporates various emotions, as depicted in Fig. 1. All images in the dataset are loaded as 48×48 pixel resolution and in grayscale [13]. However, FER2013 is stored in CSV format, necessitating data extraction from the file and ordering the data based on the pixel values. The model is trained using batches and epochs, and after experimenting with different batch sizes and epoch numbers, the optimal accuracy was achieved using a batch size of 512 and 50 epochs.

The JAFFE [14] database contains 213 grayscale photographs capturing the facial expressions of seven Japanese girls. In our discussion, we evaluated a total of 213 photos, assigning 116 for training and 97 for testing, with 96.6% of the data allocated for training purposes.

The Cohn-Kanade database (CK+) consists of 849 images captured from a group of one hundred college students aged between 18 and 30 years. The dataset is divided into 566 training images and 283 test images. The performance of the system was assessed using this dataset, achieving an average training allocation of 98.6%.

User interface implementation Python 3.7.0 has been used to implement the proposed model. The facial expressions system is built with libraries like NumPy, Pandas, OpenCV, Dlib, and TensorFlow with keras, PIL, and tqdm. For the user interface, we used the PyQt toolkit. Our model is made to take input from a webcam or by selecting files from the system with the browse option, as shown in the figure below. Keras from TensorFlow served as the back end for the various classifications of facial expressions. The artificial recognition facial expressions training process was carried out on a system with a CPU i7 64 bits, 16 GB of RAM, and dedicated 4 GB of NVIDIA GTX 1050 ti. The best model that emerged from the various trials and errors was then saved in HDF5 format.

4.2 Flow Chart for FER Using Convolutional Neural Networks

Step 1: *Start*: Begin the flow chart.

Step 2: *Data Collection*: Collect facial expression images from the FER dataset.

Step 3: *Data Preprocessing*:

Step 3.1: Convert images to grayscale.

Step 3.2: Resize images to a standardized size (e.g., 48×48 pixels).

Step 3.3: Normalize pixel values to a suitable range (e.g., 0–1).

Step 3.4: Split the dataset into training, testing, and validation sets.

Step 4: *Model Architecture Design*:

Step 4.1: Define the architecture of the CNN model.

Step 4.2: Determine the number of convolutional layers, pooling layers, and fully connected layers.

Step 4.3: Specify the activation functions to be used (e.g., ReLU).

Step 4.4: Decide on the number of output classes (corresponding to different facial expressions).

Step 5: *Model Training*:

Step 5.1: Initialize the CNN model with random weights.

Step 5.2: Feed the training images into the model.

Step 5.3: Perform forward propagation and calculate the loss.

Step 5.4: Backpropagate the loss to update the model weights using gradient descent.

Step 5.5: Repeat steps b–d for multiple epochs.

Step 6: *Model Evaluation*:

Step 6.1: Feed the testing images into the trained model.

Step 6.2: Perform forward propagation to obtain predictions.

Step 6.3: Compare the predicted labels with the ground truth labels.

Step 6.4: Calculate evaluation metrics such as accuracy, precision, recall, and F1-score.

Step 7: *Model Optimization*:

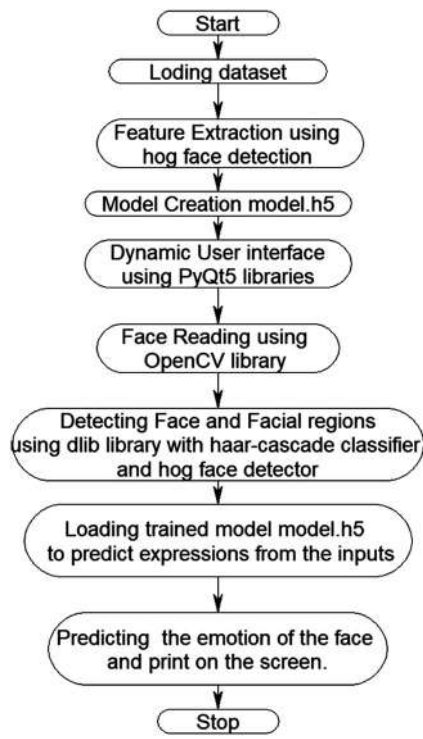
Step 7.1: Analyze the model's performance and metrics.

Step 7.2: Adjust hyperparameters (e.g., learning rate and batch size) for improved results.

Step 7.3: If necessary, modify the model architecture or try different CNN architectures.

Step 8: *Model Deployment:*

- Step 8.1: Save the trained model for future use.
 - Step 8.2: Prepare the model for deployment in a real-time system or application.
- Step 9: End: Finish the flow chart.



4.3 System Results

We can see the confusion matrix from Tables 1, 2, 3, and 4. Figure 4 is the sample screen for user interface screen of facial recognition system, Fig. 5 is user interface screen with Browse option, Fig. 6 is facial expression recognition of sample images, Fig. 7 is facial expression recognition of images when image is captured using webcam, and Fig. 8 shows test recognition rate of facial expressions on various datasets. Different expressions showed different accuracies: happy and sad showed a higher accuracy of 94%, angry 85%, fear 86%, disgust 83%, surprise 91%, and neutral 88% as shown in Table 5. Comparative study of various states of the arts

Table 1 The confusion matrix on FER2013

	AR	DS	FR	HP	SD	SR	NL
AR	95%	2%	3%	0%	0%	0%	0%
DS	3%	92%	2%	0%	3%	0%	0%
FR	0%	0%	88%	3%	4%	3%	2%
HP	0%	0%	0%	98%	0%	2%	0%
SD	1%	1%	0%	0%	98%	0%	0%
SR	2%	2%	0%	0%	0%	96%	0%
NL	0%	0%	3%	0%	0%	2%	95%

Table 2 The confusion matrix on CK+

	AR	DS	FR	HP	SD	SR	NL
AR	88%	5%	4%	0%	3%	0%	0%
DS	3%	96%	1%	0%	0%	0%	0%
FR	0%	0%	97%	1%	2%	0%	0%
HP	0%	0%	0%	98%	0%	2%	0%
SD	0%	1%	1%	0%	98%	0%	0%
SR	0%	0%	0%	1%	0%	99%	0%
NL	0%	0%	1%	0%	0%	1%	98%

Table 3 JAFFE—confusion matrix

	AR	DS	FR	HP	SD	SR	NL
AR	99%	1%	0%	0%	0%	0%	0%
DS	3%	92%	2%	0%	2%	0%	1%
FR	0%	0%	96%	0%	2%	2%	0%
HP	0%	0%	0%	99%	0%	1%	0%
SD	2%	2%	1%	0%	95%	0%	0%
SR	0%	0%	1%	1%	0%	98%	0%
NL	0%	0%	2%	0%	0%	2%	96%

Table 4 MUG—confusion matrix

	AR	DS	FR	HP	SD	SR	NL
AR	99%	0%	1%	0%	0%	0%	0%
DS	3%	92%	2%	0%	2%	0%	1%
FR	0%	0%	96%	0%	2%	2%	0%
HP	0%	0%	0%	98%	0%	2%	0%
SD	2%	2%	1%	0%	95%	0%	0%
SR	0%	0%	1%	1%	0%	98%	0%
NL	0%	0%	2%	0%	0%	2%	96%

literature compares the results of the current state of the arts to the proposed method. The databases MUG, FER2013, JAFFE, and CK+ are used to evaluate our studies. The results of the comparisons with various approaches are presented in Table 5. It is also worth noting that our methods were able to identify facial expressions at a higher average rate than those in [15].

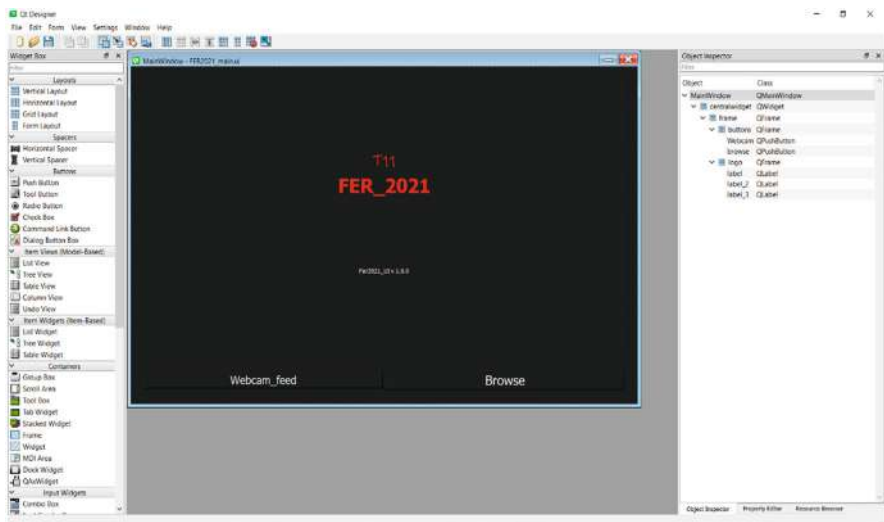


Fig. 4 User interface screen of the facial recognition system

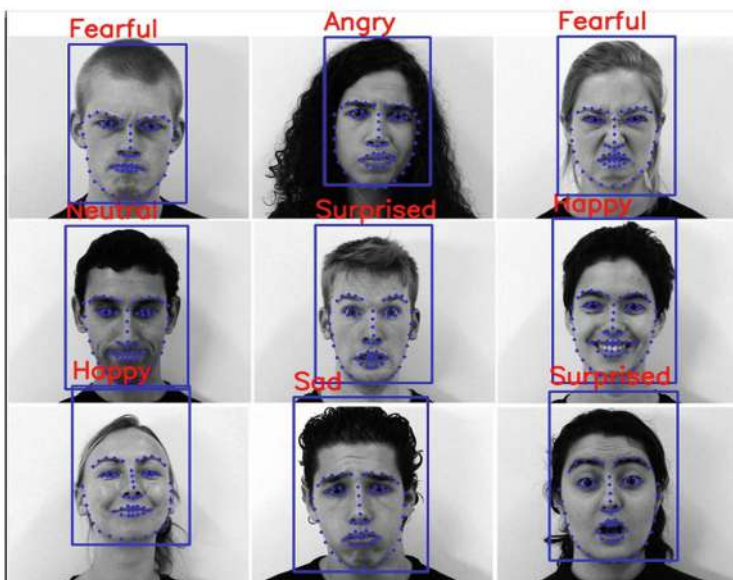
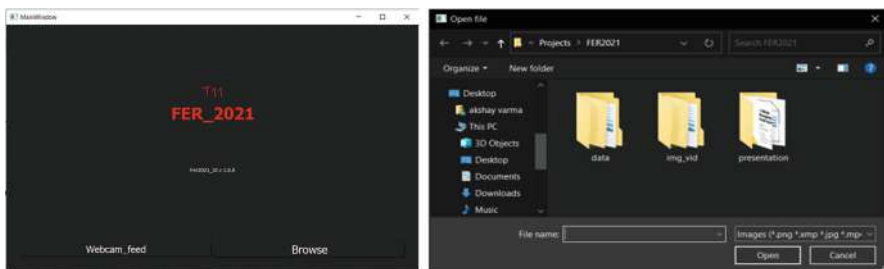


Fig. 5 User interface screen with Browse option

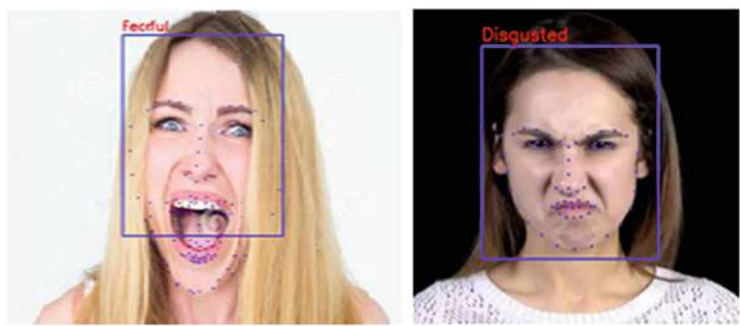


Fig. 6 Facial expression recognition of sample images



Fig. 7 Facial expression recognition of images when image is captured using a webcam



Fig. 8 Test recognition rate of facial expressions on various datasets

Table 5 Comparison of testing recognition rate on various datasets of facial expressions

	Dataset Used	AR	DS	FR	HP	SD	SR	Average recognition accuracy (%)
Proposed model	FER2013	95	92	88	98	98	96	94.83
Fengping An et al. [4]	FER2013	81	80	85	90	85	88	84.83
Proposed model	CK+	88	96	97	100	98	99	96.33
Haar Satiyan et al. [5]	CK+	87.8	93.3	94.3	94.2	96.4	98.4	94.1
Zhang et al. [7]	CK+	87	91.5	90.9	96.9	84.5	91.2	90.3
Li, S et al. [8]	CK+	76.2	94.1	86.1	96.3	88.2	98.7	91.5
Proposed model	JAFFE	100	92	99	99	95	98	96.83
Hog Chen et al. [6]	JAFFE	100	86.2	93.7	96.7	77.4	96.6	91.7
Mahmood et al. [9]	JAFFE	89.3	90.7	91.1	92.6	90.2	92.3	91.1

5 Conclusion and Future Enhancement

In this chapter, we present a look acknowledgment program that distinguishes seven distinct widespread feelings. Vertical and dynamic inputs from a web camera or external recordings like videos or photos are integrated into the program. A deep learning toolbox also includes data development to avoid overcrowding. We have found that the display features, in which input images or videos are collected under a variety of lighting conditions, are accurate, which lowers the level of facial recognition and makes the feature’s rendering more difficult. These issues may be the focus of future research that takes into account technological advancements and employs novel approaches to improve the accuracy of the facial recognition system.

References

1. Kumar, P., Roy, P. P., & Dogra, D. P. (2018). Independent bayesian classifier combination based sign language recognition using facial expression. *Information Sciences*, 428, 30–48.

2. Zeng, J., Shan, S., & Chen, X. (2018). Facial expression recognition with inconsistently annotated datasets. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 222–237).

3. Luo, Y., Wu, C. M., & Zhang, Y. (2013). Facial expression recognition based on fusion feature of PCA and LBP with SVM. *Optik-International Journal for Light and Electron Optics*, 124(17), 2767–2770.

4. An, F., & Liu, Z. (2020). Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM. *The visual computer*, 36(3), 483–498.

5. Haar Satiyan, M., Nagarajan, R., Hariharan, M.: Recognition of facial expression using Haar wavelet transform. *Trans. Int. J. Electr. Electron. Syst. Res. JEESR Univ. Technol. Mara UiTM* 3, 91–99 (2010)

6. Hog Chen, J., Takiguchi, T., Arika, Y.: Facial expression recognition with multithreaded cascade of rotation-invariant HOG. In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, pp. 636–642 (2015)

7. Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2018). From facial expression recognition to interpersonal relation prediction. *International Journal of Computer Vision*, 126(5), 550–569.
8. Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE transactions on affective computing*.
9. Mahmood, M. R., Abdulrazzaq, M. B., Zeebaree, S. R., Ibrahim, A. K., Zebari, R. R., & Dino, H. I. (2021). Classification techniques' performance evaluation for facial expression recognition. *Indonesian Journal of Electrical Engineering and Computer Science*, 21(2), 176–1184.
10. Lopes, A. T., De Aguiar, E., De Souza, A. F., & Oliveira-Santos, T. (2017). Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. *Pattern recognition*, 61, 610–628.
11. Meng, Z., Liu, P., Cai, J., Han, S., & Tong, Y. (2017, May). Identity-aware convolutional neural network for facial expression recognition. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (pp. 558–565). IEEE.
12. Greche, L., Jazouli, M., Es-Sbai, N., Majda, A., & Zarghili, A. (2017, April). Comparison between Euclidean and Manhattan distance measure for facial expressions classification. In *2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)* (pp. 1–4). IEEE.
13. Yang, H., Ciftci, U., & Yin, L. (2018). Facial expression recognition by de-expression residue learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2168–2177).
14. Zhang, F., Zhang, T., Mao, Q., & Xu, C. (2018). Joint pose and expression modeling for facial expression recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3359–3368).
15. Barman, A., & Dutta, P. (2021). Facial expression recognition using distance and shape signature features. *Pattern Recognition Letters*, 145, 254–261.

A Hybrid Machine Intelligence Demographic Feature Selection Approach to Improve Recommendation System in Social Domain



Bandi Vamsi , Mohan Mahanty , and Bosubabu Sambana 

Abstract Based on the preferences and previous history with many other customers, recommendation system assists users in quickly finding pertinent goods. They benefit users by making searching easier and companies by encouraging the sale of their products. Various filtering methods, such as collaborative, demographic, and content-based filtering are used to create a prediction model. Furthermore, because of concerns with variance and sparsity, massive amounts of information may result in suggestions with a limited degree of accuracy. In this work, we analyze random forest (RF) and AdaBoost classifiers and suggest a hybrid learning technique that uses user-to-user filtering features to identify the closest users. This technique is designed as a result of research into how to lower rating prediction mistakes based on user prior interactions, thus increasing prediction performance in two classification algorithms. We used a characteristic combination technique to increase predictability. Our proposed model has attained less error rate on using random forest and AdaBoost classifiers for both mean absolute error and root mean squared error filters.

Keywords Content-based filtering · Collaborative filtering · Demographic filtering · Recommendation system · Feature selection

B. Vamsi

Department of Artificial Intelligence, Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India

M. Mahanty

Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam, India

B. Sambana (✉)

Department of Computer Science and Engineering, Avanthi Institute of Engineering and Technology (A), Makavarapalem, JNTU GV Vizianagaram, Andhra Pradesh, India

1 Introduction

Overloading with information is a result of the rapid expansion of data available via the Internet during the past ten years. More precisely, a broader variety of possibilities are being presented by the e-commerce sector, which renders it much more challenging for customers to find and buy products [1]. Therefore, businesses must create a recommendation system in order to assist clients in discovering new products through suggestions. By offering options that are pertinent to and fulfill the needs of users, these tools aid in the growth of their products [2]. For instance, the “Spotify music app” increased the value of the recommendation systems for acquiring more customers in the context of song suggestions, as well as the rivalry to build fully advanced techniques resulted in more attractive recommendations [3].

A suggestion system refers to a method of automated analysis that is utilized to suggest a selection of top objects that fit the user’s interests or to forecast the objects an individual user really likes [4]. While attempting to locate the correct content at the right moment, users struggle to manage large amounts of data and experience cognition and informational sparsity issues [5].

A classification method can be divided into four basic parameters on profile information: similarity calculation, neighborhood identification, prediction, and recommendations. Filters like content-based, collaborative, and demographic modeling are available for the profile [6]. It is considered a content-based filter if the user account includes a group of characteristics gleaned from the product details the consumer found appealing. A subscriber’s profile information features serve as a demographic representation [7]. Content filtering is defined as the presence of a list of average ratings in the characteristics [8]. Due to its effectiveness in identifying users who have similar interests to one another, the content-based approach is frequently utilized as a recommendation systems system. Furthermore, when utilized alone, content approaches may experience a sparsity problem due to rapid rise in the number of customers and objects [9, 10].

Consequently, this work has contributed the following:

1. In order to address the issue of imbalanced data, we have devised a system that makes use of both rating system and demographic characteristics, integrating demographic factors with customer item rating.
2. With this approach, we are able to quickly determine similarities across a huge dataset without performing any calculations or prior filtering.
3. We have demonstrated that our strategy is both feasible and effective by evaluating it utilizing a real sample and very well assessment metrics.

The remaining sections of this work are as follows. Section 2 deals with related work, Sect. 3 discusses about methodology, Sect. 4 explains the performance of experimental results, and Sect. 5 describes about conclusion.

2 Related Work

The knowledge graph-based methods that are currently advised do not make use of extensive semantic relationships. A general approach with integrated extracted features of behavior and learning characteristics is proposed, and factors mentioned are extracted from their previous behavior and domain knowledge in order to give the customers suggestions that are more precise and varied [11].

Their primary duty is to deliver high-quality training materials, and it is well-known that the absence of online support from providers is the root of many problems. To improve the abilities and knowledge of the users, it is necessary to develop a mechanism that can effectively suggest courses while taking various perspectives into account. This suggests a system that employs virtual entities to create semantic suggestions based on user needs and likes, supporting experts in finding the right courses in a practical situation [12].

The commentary part is not currently used in item recommendations, which are based purely on ratings such as 1–5. Users or clients share their sentiments and ideas regarding items and services in this environment. This study suggests that a supervised classification model approach for rating products using the numbers 0, 2, and 4. 0 is unfavorable, 2 is impartial, and 4 is favorable. This will happen in addition to the rating system that is currently in place and handles customer reviews and feedback without interfering [13].

The model-based predictions are produced using a random forest (RF) method. First, two different datasets are taken from the initial one with the intent of making recommendations according to demographics and content, accordingly. Second, average evaluations are included as a characteristic in the new data. Third, a tree is constructed for every user or object, with each of the leaves receiving a separate score [14].

The hybrid method [15], established from an examination into ways that minimize the inaccuracies in evaluating predictions according to customers' previous conversations, integrates user–user collaborative with the properties of demographics to identify the closest users and compares the classifier based on random forests versus the nearest neighbor classifier. In two separate categorization methods, our combination approach improves prediction accuracy. The major objective of this study [15] is to assess the two models and determine how demographics affect collaborative.

The most frequently used methods for producing recommendations are collaborative and content-based filtering, both of which provide user-to-user suggestions. Collaborative filtering, nevertheless, has an extended tail challenge, meaning that it performs poorly with things that have fewer evaluations than huge item audiences with correspondingly significant amounts of ratings. This strategy for solving the extended tail of the suggestion problem uses “user history” and case-based analysis to forecast how users would rate recently observed objects that appear to be within the long tail [16].

3 Methodology

This section describes the methods and filters used in this study to develop a hybrid recommendation system.

3.1 Dataset

The dataset used in this study describes a detailed movie recommendation provided by “MovieLens” [17] which contains scores and free-text labeling actions. There are 465,564 tag entries and 20,000,263 reviews spread among 27,278 films. A total of 138,493 customers between January 9, 1995, and March 31, 2015, produced these statistics. Users who wanted to be included were chosen at random. Each chosen user had given at least 20 films a rating. The statistical analysis of this dataset is described in Table 1.

3.2 Requirements for the recommendation system

The requirements for developing a novel recommendation system have three subcomponents as follows:

- (a) Given a big collection of hundreds of objects, the Candidate Generations technique is in charge of producing smaller sub-datasets of possibilities to suggest to a user.
- (b) We must standardize everything and make an effort to provide a grade for every one of the things in the subgroups because various generators may perform different Future Generations. The rating system does this.
- (c) The system uses a re-ranking method to determine the final ranks when scoring is completed, as well as various extra limits.

Table 1 Summary of dataset [17]

Parameter	Size	Mean	Standard deviation	Error mean	P-value
Movie_id	27,278	59855.48	44429.315	269.007	<0.001

Table 2 Utility matrix on album recommendation

Album (A)	User 1	User 2	User 3	User 4	Interested album	Added to queue
A1	3	0	0	1	No	No
A2	5	4	4	0	Yes	Yes
A3	4	0	5	4	Yes	No
A4	2	4	1	2	No	No
A5	4	5	5	5	Yes	Yes

Table 3 Similarities measuring among users and audio album

Users	Album 1	Album 2	Album 3	Album 4
User-1	5	4	0	4
User-2	5	0	4	0
User-3	0	1	0	2
User-4	1	1	0	0

3.3 Content-Based Filtering System

A content-based filtering system aims to predict a subscriber’s characteristics or behavior depending on the qualities of the object to which the user responds favorably. This can be done through a matrix representation called a utility matrix.

The utility matrix indicates how the customer feels about particular items. We need to establish a connection between the subscriber’s preferred and least preferred objects in the user information, and this matrix is what we utilize to do this. Each customer-to-customer pair is given a specific value in this system, referred to as the amount of preference. Finally, to determine the relationship between a customer’s request, we create a matrix of the person with the appropriate items. The utility matrix for a song recommendation is shown in Table 2. Even though we sometimes do not receive complete data from the user, a few of the matrix’s cells are empty. However, the objective of a recommendation is not to complete all the fields; rather, it is to suggest a song to the users that they are interested to listen again or not.

3.4 Collaborative-Based Filtering System

In information retrieval, we frequently identify similar users and suggest things that they like. Instead of using the characteristics of the object to propose it in this kind of system, we group clients into clusters of individuals who have similar traits and then make suggestions for each one based on the preferences of its cluster.

From Table 3, we recognize that user-1 and user-2 given the audio album ratings are virtually identical; therefore, we can infer that user-1 and user-2 will find album-3 to be of ordinary interest, while user-2 and user-3 will find album-4 to be highly recommended. Some subscribers have different desires because user-1 and user-3 are in direct opposition to one another.

3.5 Demographic-Based Filtering System

By referring to the demographic information, it is also possible to understand the kind of individual who likes an object. User characteristics are taken into account by demographic optimization techniques, which then use the demographic information as a foundation to provide necessary suggestions. Data is collected both consciously and unconsciously as users interact with the service. The suggestions are based on how similarly rated items were by different reviewers.

3.6 Proposed Model

Our hybrid learning recommendation strategy has generated a lot of discussion. A combined collection of user-based quality, collaborative, and demographic models may be used to solve the challenge of determining the evaluation of users and items in a matrix using the suggestion technique. This approach gains from the demographic characteristic association between customer similarities. In comparison to the baseline, this increases prediction and can resolve this issue.

When it comes to making the appropriate recommendation for such users, the sparsity problem poses a significant barrier to optimization techniques. Because of the increase in features available and customers with scant reviews and customer data, this problem has grown even more. This makes it tough to compare two users who are similar. With the use of different classes, a characteristic combining hybrid technique resolves the sparsity issue and lowers the margin of error. As shown in Fig. 1, it includes the customer rating collaborative filtering technique with correlating individual demographic characteristics.

The proposed recommender system is divided into four main steps according to the user's profile information such as finding similarity, finding neighborhood selection, recommendation, and classification. By using filters used in this work, the user profiles can be modeled. It is considered content-based filtering if the account includes a group of characteristics gathered from the product specifications the customers found appealing.

A consumer's profile attributes serve as demographic representation. The collaborative filtering analyses the presence of a list of evaluated products in the profile. Due to its effectiveness and capacity for identifying comparable neighbor users, the collaborative approach is frequently utilized as a recommendation system. When this approach is employed alone, however, the rapid rise in the variety of users and objects may result in a sparsity problem.

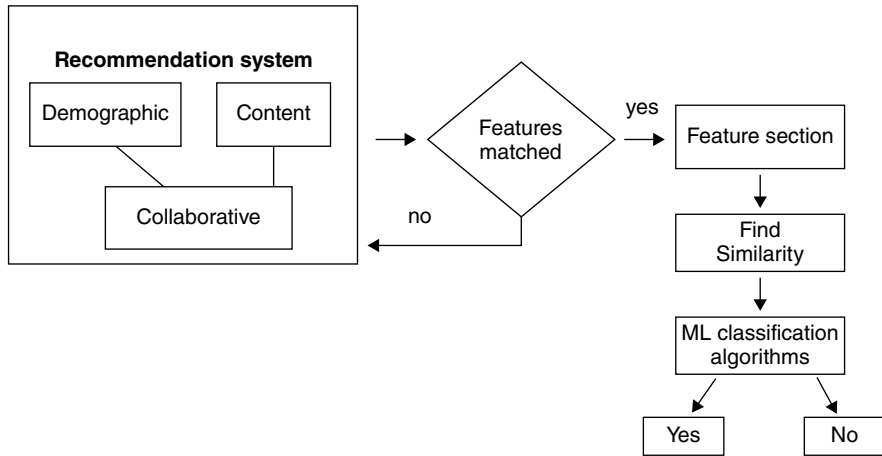


Fig. 1 Hybrid recommendation model

4 Experimental Results

4.1 Random Forest

This classifier uses ensemble techniques to create a series of decision trees. From a random subset of training examples, each tree is created. It is additionally noise-resistant and robust. The effectiveness of this approach as a classifier using machine learning has already been effectively validated. Eq. (1) defines the mathematical formula for calculating the Gini index (GI) for determining the feature importance score while employing RF:

$$GI = 1 - \left[\left(P_+^2 \right) + \left(P_-^2 \right) \right] \quad (1)$$

Here, P_+ and P_- determine the probability of positive and negative classes.

4.2 AdaBoost

It is another popular classification algorithm that may be applied in applications of machine learning. It is a widely used boosting classification that increases accuracy by combining several ineffective classifiers into one more accurate and satisfying classification. A training rate of 1 and 10 estimation methods has been utilized in this work. This model verifies if the classifier is good or not by using Eq. (2):

$$f(x_i) = \sum_{t=1}^T \alpha_t h_t(x_i) \quad (2)$$

4.3 Mean Absolute Error (MAE)

It computes the mean of the absolute deviation including all ratings as specified in Eq. (3) between both the expected score and the observed rating. p_i represents the predicted score in absolute numbers, and r_i represents the actual score in absolute terms. Smaller amounts must be taken into account as they produce better outcomes:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |p_i - r_i| \quad (3)$$

4.4 Root Mean Squared Error (RMSE)

Equation (4) defines it as the sampling standard deviation of the variances between expected values and the real values. The observed score is r_i , while p_i is the anticipated score. Lower scores should be taken into account as they produce better outcomes:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2} \quad (4)$$

4.5 Classification Results

Tables 4 and 5 demonstrate how random forest classification is used to create the MAE and RMSE accuracy scores. Similarly, Tables 6 and 7 show how AdaBoost is used to create MAE and RMSE accuracy scores. By doing the random forest classification experiments with a separate set of tree branches, it is obvious that efficiency was enhanced when demographic factors and content-based factors were merged. It is clear that the random forest algorithm outperforms AdaBoost in performance. The visualization of the proposed model outcomes is shown in Fig. 2.

Table 4 Performance of recommendation models by random forest using MAE filter

Album	MAE metric on filters			
	Content	Collaborative	Demographic	Hybrid
1	0.88	0.81	0.71	0.65
2	0.81	0.78	0.75	0.62
3	0.86	0.75	0.82	0.52
4	0.89	0.72	0.79	0.58
5	0.91	0.79	0.69	0.63

Table 5 Performance of recommendation models by random forest using RMSE filter

Album	RMSE metric on filters			
	Content	Collaborative	Demographic	Hybrid
1	0.83	0.74	0.84	0.54
2	0.79	0.72	0.62	0.47
3	0.74	0.79	0.74	0.62
4	0.81	0.75	0.7	0.41
5	0.83	0.69	0.63	0.47

Table 6 Performance of recommendation models by AdaBoost using MAE filter

Album	MAE metric on filters			
	Content	Collaborative	Demographic	Hybrid
1	0.93	0.83	0.76	0.72
2	0.86	0.8	0.8	0.69
3	0.91	0.77	0.87	0.59
4	0.94	0.74	0.84	0.65
5	0.96	0.81	0.74	0.70

Table 7 Performance of recommendation models by AdaBoost using RMSE filter

Album	RMSE metric on filters			
	Content	Collaborative	Demographic	Hybrid
1	0.89	0.71	0.87	0.62
2	0.81	0.81	0.74	0.54
3	0.79	0.84	0.79	0.69
4	0.72	0.89	0.81	0.53
5	0.89	0.73	0.68	0.65

5 Conclusion

In this work, we proposed a new hybrid approach for a recommendation system that simultaneously combines demographic information and user-to-user-based collaborative filtering. According to the findings, demographics can significantly enhance the total suggestion. It is clear that the random forest algorithm outperforms AdaBoost in performance. In the future, we hope to employ particular consumers who exclusively rate a select number of things as well as perhaps additional item-related criteria.

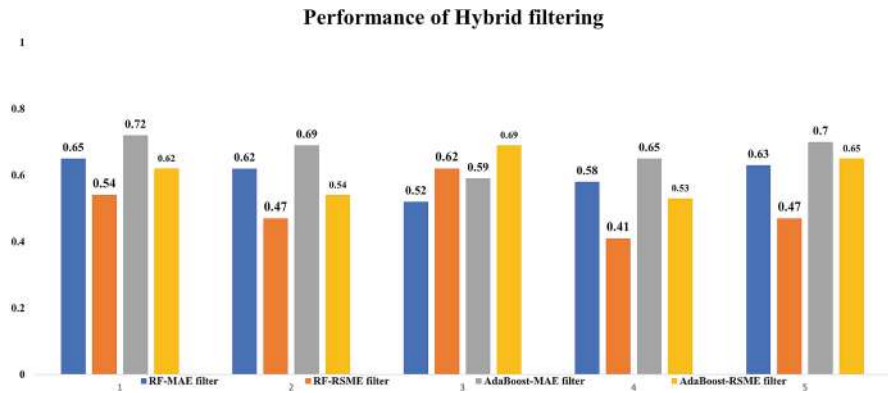


Fig. 2 Performance of hybrid filtering

References

1. Zhang, D., Liu, L., Wei, Q., Yang, Y., Yang, P., & Liu, Q. (2020). Neighborhood aggregation collaborative filtering based on knowledge graph. *Applied Sciences*, 10(11), 3818.
2. Yang, S., & Cai, X. (2022). Bilateral knowledge graph enhanced online course recommendation. *Information Systems*, 107, 102000.
3. Romadhony, A., Al Faraby, S., & Pudjoatmodjo, B. (2013, March). Online shopping recommender system using hybrid method. In *2013 International Conference of Information and Communication Technology (ICoICT)* (pp. 166–169). IEEE.
4. Yun, Y., Hooshyar, D., Jo, J., & Lim, H. (2018). Developing a hybrid collaborative filtering recommendation system with opinion mining on purchase review. *Journal of Information Science*, 44(3), 331–344.
5. Yao, G., & Cai, L. (2015). User-based and item-based collaborative filtering recommendation algorithms design. University of California, San Diego.
6. Kim, J. K., Suh, J. H., Ahn, D. H., & Cho, Y. H. (2002). A personalized recommendation methodology based on collaborative filtering. *Journal of Intelligence and Information Systems*, 8(2), 139–157.
7. Dhelim, S., Aung, N., Bouras, M. A., Ning, H., & Cambria, E. (2022). A survey on personality-aware recommendation systems. *Artificial Intelligence Review*, 55(3), 2409–2454.
8. Liu, G. (2022). An ecommerce recommendation algorithm based on link prediction. *Alexandria Engineering Journal*, 61(1), 905–910.
9. Valentino, V. H., Setiawan, H. S., Saputra, A., Haryanto, Y., & Putra, A. S. (2021). Decision support system for thesis session pass recommendation using AHP (analytic hierarchy process) method. *International Journal of Educational Research & Social Sciences*, 2(1), 215–221.
10. Onile, A. E., Machlev, R., Petlenkov, E., Levron, Y., & Belikov, J. (2021). Uses of the digital twins concept for energy services, intelligent recommendation systems, and demand side management: A review. *Energy Reports*, 7, 997–1015.
11. Hui, B., Zhang, L., Zhou, X., Wen, X., & Nian, Y. (2022). Personalized recommendation system based on knowledge embedding and historical behavior. *Applied Intelligence*, 52(1), 954–966.
12. Khosravani, M. R., Nasiri, S., & Reinicke, T. (2022). Intelligent knowledge-based system to improve injection molding process. *Journal of industrial information Integration*, 25, 100275.

13. Iwendi, C., Ibeke, E., Eggoni, H., Velagala, S., & Srivastava, G. (2022). Pointer-based item-to-item collaborative filtering recommendation system using a machine learning model. *International Journal of Information Technology & Decision Making*, 21(01), 463–484.
14. Zhang, H., Min, F., & Wang, S. (2014). A random forest approach to model-based recommendation. *Journal Of Information & Computational Science*, 11(15), 5341–5348.
15. Alshammari, G., Kapetanakis, S., Alshammari, A., Polatidis, N., & Petridis, M. (2018). A hybrid feature combination method that improves recommendations. In *Computational Collective Intelligence: 10th International Conference, ICCCI 2018, Bristol, UK, September 5–7, 2018, Proceedings, Part I* 10 (pp. 209–218). Springer International Publishing.
16. Alshammari, G., Jorro-Aragoneses, J. L., Kapetanakis, S., Petridis, M., Recio-García, J. A., & Díaz-Agudo, B. (2017). A hybrid CBR approach for the long tail problem in recommender systems. In *Case-Based Reasoning Research and Development: 25th International Conference, ICCBR 2017, Trondheim, Norway, June 26–28, 2017, Proceedings* 25 (pp. 35–45). Springer International Publishing.
17. F. Maxwell Harper and Joseph A. Konstan. (2015). The MovieLens Datasets: History and Context. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 5, 4, Article 19 (December 2015), 19 pages. DOI=<https://doi.org/10.1145/2827872>

An Exploratory Review of Machine Learning and Deep Learning Applications in Healthcare Management



Narasimha Rao Vajjhala  and Philip Eappen 

Abstract The growing interest in machine learning (ML) and deep learning (DL) applications for healthcare management has led to a plethora of research outputs. In this exploratory study, the state-of-the-art ML and DL applications in healthcare management are synthesized to provide an overview of the advancements, challenges, and future directions. A total of 65 studies were analyzed, focusing on four main categories: patient care, medical imaging, electronic health records, and healthcare operations. The review highlights the potential of ML and DL in improving healthcare outcomes and efficiency while emphasizing the need for further research to address concerns such as data privacy, model interpretability, and integration into clinical practice.

Keywords Healthcare management · Machine learning · Deep learning · Predictive analytics · Medical decision support · Electronic health records · Patient outcome prediction · Personalized medicine · Medical imaging analysis · Natural language processing

1 Introduction

Machine learning (ML) and deep learning (DL) have emerged as powerful tools to transform healthcare management, driven by increasing computational power, vast amounts of data, and innovative algorithmic advancements [1, 2]. The rapid advancements in artificial intelligence (AI), particularly ML and DL, have presented a unique opportunity to revolutionize the healthcare sector [3]. The massive volume of patient data, coupled with the growing computational power, has paved the way

N. R. Vajjhala
University of New York Tirana, Tirana, Albania

P. Eappen (✉)
Cape Breton University, Cape Breton, NS, Canada
e-mail: narasimharao@unyt.edu.al

for innovative applications of ML and DL to address various challenges in healthcare management [4]. These applications have the potential to not only enhance clinical decision-making and optimize resource allocation but also improve patient outcomes and overall healthcare efficiency [5, 6]. This exploratory study aims to provide a comprehensive understanding of the current state-of-the-art ML and DL applications in healthcare management, examining trends, challenges, and future research directions. The integration of ML and DL into healthcare management has gained significant momentum in recent years, with researchers and practitioners exploring a wide range of applications, such as disease diagnosis, medical imaging, electronic health record (EHR) analysis, and healthcare operations optimization [7–11]. By leveraging the power of these advanced algorithms, healthcare professionals can uncover hidden patterns and insights from complex data, leading to more accurate, personalized, and timely care delivery. Despite the promising advancements, several challenges remain in the implementation of ML and DL for healthcare management, including data privacy, model interpretability, and the integration of these technologies into clinical workflows [12, 13]. Addressing these challenges is crucial to ensure the widespread adoption and acceptance of ML and DL in healthcare settings.

2 Methodology

The primary research objectives of this exploratory review are to:

- Identify the main applications of ML and DL in healthcare management.
- Investigate the potential benefits and challenges associated with the implementation of ML and DL techniques in healthcare settings.
- Provide insights into the current state-of-the-art and future directions in this rapidly evolving field.

A systematic search of PubMed, IEEE Xplore, ScienceDirect, and Google Scholar databases was performed to identify relevant studies published between January 2015 and March 2023. The search was limited to English-language articles. Keywords used include “machine learning,” “deep learning,” “artificial intelligence,” “healthcare management,” “medical imaging,” “patient care,” “electronic health records,” “confidentiality,” “data protection,” “GDPR” (General Data Protection Regulation), “HIPAA” (Health Insurance Portability and Accountability Act), and “healthcare operations.” After screening, 65 studies were included in the review.

The methodological quality of the included studies was assessed using the Newcastle-Ottawa Scale for cohort and case-control studies, and the Cochrane Collaboration’s Risk of Bias Tool for randomized controlled trials. Due to the expected heterogeneity of the ML/DL applications and healthcare management settings, a narrative synthesis was conducted. The results were summarized in a descriptive manner, focusing on the ML/DL techniques used, the specific healthcare

management applications, and the reported improvements in efficiency, patient outcomes, or resource utilization. Additionally, we identified common challenges and future directions for research in this area.

3 Findings

In this exploratory systematic review, we investigated the applications of ML and DL techniques in healthcare management. The 65 studies were categorized into four main application areas, namely, patient care, medical imaging, EHRs, and healthcare operations. The most reported ML techniques included decision trees (DT), support vector machines (SVM), random forests (RF), and logistic regression (LR). Among DL techniques, artificial neural networks (ANNs) and convolutional neural networks (CNNs) were the most frequently employed. The use of DL has significantly increased over the past 5 years, indicating a growing trend toward more complex models in healthcare management research.

3.1 *Patient Care*

ML and DL techniques have demonstrated the potential to improve patient care by predicting disease outcomes, facilitating early diagnosis, and personalizing treatment plans. These advanced computational methods are expected to revolutionize various aspects of healthcare, including diagnostics, treatment planning, and personalized medicine. For instance, Malo [14] demonstrated the use of a CNN to detect skin cancer with accuracy comparable to dermatologists. The applications identified from the literature on patient care include disease prediction and prognosis, predicting patient readmissions, and personalized medicine and drug discovery [15–17].

3.1.1 Disease Prediction and Prognosis

ML has been employed in various aspects of cancer care, including diagnosis, treatment planning, and prognosis [14, 18–21]. For instance, Kourou et al. [21] reviewed ML applications in cancer diagnosis and prognosis, highlighting the potential of these techniques in improving patient outcomes. DL techniques have shown remarkable success in analyzing cancer imaging and genomic data [22, 23]. Yu [24] demonstrated the use of DL to predict lung cancer patient outcomes from histopathology images, indicating its potential in personalized treatment planning. Additionally, Di Noia et al. [25] applied DL to genomic data for predicting survival in patients with brain tumors.

ML techniques have been utilized for the early detection and management of diabetes [26–28]. Jaiswal et al. [28] conducted a review of ML applications in diabetes diagnosis and management, emphasizing their potential in improving patient outcomes and reducing healthcare costs. ML has been employed to predict cardiovascular risk and improve patient outcomes [29–31]. Shah et al. [29] elaborated on the superiority of ML algorithms over traditional risk prediction models in predicting the onset of cardiovascular diseases. DL techniques have been applied to cardiac imaging and electrocardiogram (ECG) analysis for improved diagnostics and risk assessment [32].

3.1.2 Predicting Patient Readmissions

Patient readmissions are a significant concern for healthcare systems due to their impact on patient outcomes and healthcare costs. Identifying patients at risk of readmission can help in early intervention and improving care management. ML and DL techniques have shown promise in predicting patient readmissions using EHR data, enabling healthcare providers to make data-driven decisions [33]. Several ML algorithms have been applied for predicting patient readmissions, including LR, SVM, DT, RF, and gradient boosting machines (GBM) [34]. These approaches have shown potential in identifying risk factors and predicting readmission probabilities.

Feature engineering plays a crucial role in improving prediction performance. Studies have found that incorporating demographic, clinical, and social factors can improve model accuracy. Model performance varies depending on the algorithm used, the dataset, and the specific patient population. In general, ensemble methods like RF and GBM tend to outperform single models like LR and SVM. Model interpretability is an essential aspect of ML applications in healthcare. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) have been employed to increase the interpretability of complex models [35, 36]. DL techniques, particularly ANNs and recurrent neural networks (RNNs), have been applied to predict patient readmissions [37, 38]. DL models can automatically learn representations from raw EHR data, reducing the need for manual feature engineering. CNNs and long short-term memory (LSTM) networks have been particularly effective in capturing temporal patterns in EHR data [39, 40].

3.2 Medical Imaging

ML and DL have revolutionized medical imaging by automating image analysis and interpretation, resulting in improved accuracy and efficiency [41–43]. DL techniques have shown remarkable success in medical imaging applications [44]. Litjens et al. [45] conducted a comprehensive review of DL techniques for medical image analysis, including segmentation, detection, and classification tasks.

Similarly, Sahiner et al. [46] highlighted the potential of DL in medical imaging, emphasizing its transformative impact on disease diagnosis and treatment.

3.3 *Electronic Health Records*

The analysis of EHRs using ML and DL has enabled the extraction of valuable insights for improved patient care and decision-making. The applications of ML and DL techniques in EHRs include predictive modeling as well as phenotyping and patient stratification [47].

3.3.1 Predictive Modeling

ML and DL techniques have been employed to build predictive models using EHR data to diagnose diseases, predict disease progression, and estimate prognosis [48]. These models often rely on a combination of structured and unstructured data from EHRs, including clinical notes, laboratory results, and imaging data. Supervised learning algorithms such as LR, SVM, and RF have been used for disease diagnosis and predicting patient outcomes. DL models, including CNNs and RNNs, have shown promise in processing complex and unstructured EHR data, such as medical images and time-series data [37, 49]. ML and DL techniques have been used to develop early warning systems and risk stratification models that can identify high-risk patients and predict adverse events, such as hospital readmissions, sepsis, or in-hospital mortality [33, 34, 38]. GBM, RF, and neural networks have been employed to build risk prediction models for various clinical outcomes. LSTM networks and other RNN architectures have been applied to model temporal dependencies in EHR data and predict time-sensitive clinical events. Leveraging EHR data, ML and DL models can generate personalized treatment recommendations by predicting patient-specific responses to interventions, drug therapies, or dosages. Reinforcement learning has been explored to optimize individualized treatment strategies by learning from patient data in EHRs. Collaborative filtering and matrix factorization techniques have been applied to EHRs to provide personalized treatment recommendations based on similarities between patients and their treatment histories.

3.3.2 Phenotyping and Patient Stratification

Phenotyping, the process of identifying and classifying individuals based on their observable traits, plays a crucial role in personalized medicine and patient stratification [50–52]. Recent studies have applied ML techniques, such as DT, RF, and SVM, to analyze EHR data and identify phenotypes associated with specific diseases, including diabetes, hypertension, and cancer [27, 53, 54]. DL techniques,

particularly RNN and LSTM networks, have been employed to model temporal relationships in EHR data for phenotyping [55, 56]. The integration of genomic data with ML and DL techniques has enabled the identification of novel disease-related genetic variants and gene expression patterns. These approaches have been particularly effective in complex diseases, such as Alzheimer's disease, autism, and various cancers [20, 57, 58]. Studies have used unsupervised ML techniques, such as clustering and dimensionality reduction, to identify disease subtypes based on gene expression patterns, while supervised techniques, including SVM and RF, have been employed for phenotype prediction [27, 59]. The integration of multi-omics data, including genomics, transcriptomics, proteomics, and metabolomics, has emerged as a promising approach to obtaining a comprehensive understanding of disease mechanisms and phenotypes. Several studies have employed ML and DL techniques, such as ensemble learning, deep autoencoders, and multitask learning, to integrate multi-omics data for improved patient stratification and personalized medicine [60, 61].

3.4 Healthcare Operations

ML and DL have been applied to optimize various aspects of healthcare operations, such as resource allocation and workflow management. Remote patient monitoring (RPM) is an essential aspect of modern healthcare. ML techniques have been applied to RPM to predict patient deterioration, enabling timely interventions [62]. For instance, Futoma et al. [49] developed an ML algorithm that successfully predicted patient deterioration using EHRs. DL techniques have been employed in telemedicine to facilitate remote diagnostics and consultations. For example, Gulshan et al. [63] utilized a DL algorithm to diagnose diabetic retinopathy in retinal images, a crucial step toward remote diagnosis and management of the condition. The applications of ML and DL in healthcare operations include patient scheduling and appointment optimization, hospital resource management (e.g., bed allocation and staff scheduling), and workflow optimization and process improvement.

4 Recommendations for Future Research

The exploratory systematic review of ML and DL applications in healthcare management has revealed promising advancements in patient care, resource optimization, and overall efficiency. Nevertheless, there are several areas that warrant further investigation to ensure the successful implementation and integration of ML and DL techniques in healthcare settings. Future research should focus on developing methodologies for ensuring data quality and standardization, as well as handling missing or incomplete data in healthcare datasets. Techniques for data preprocessing, imputation, and integration can help overcome these challenges,

leading to more accurate and reliable ML and DL models. The complexity of DL models often limits their interpretability, making it challenging for healthcare practitioners to understand and trust the results. Developing more transparent and explainable models is crucial for their adoption in clinical settings. Future research should explore explainable AI approaches to enhance the interpretability of ML and DL models in healthcare management. The performance of ML and DL models should be evaluated across diverse healthcare settings and patient populations to ensure their generalizability. Future studies should focus on cross-validation, external validation, and the development of standardized evaluation metrics to assess the performance of ML and DL models in various healthcare contexts.

The use of ML and DL in healthcare management raises ethical and legal concerns, such as patient privacy, data security, and algorithmic bias. Future research should explore ways to address these concerns by developing privacy-preserving ML and DL techniques, implementing robust data security measures, and ensuring fairness and transparency in the development and deployment of models. The successful implementation of ML and DL in healthcare management requires the collaboration of multiple stakeholders, including healthcare providers, data scientists, engineers, and policymakers. Future research should also promote interdisciplinary collaboration through the development of shared frameworks, terminologies, and best practices, as well as by fostering partnerships between academic institutions, healthcare organizations, and industry partners.

5 Conclusion

The exploratory systematic review of ML and DL applications in healthcare management has revealed significant progress and potential benefits in patient care, resource optimization, and overall healthcare efficiency. ML and DL techniques have demonstrated promising results in key application areas, such as patient risk stratification, healthcare resource optimization, disease diagnosis and prognosis, and personalized medicine. Despite the advancements, there are several challenges and limitations that need to be addressed to ensure the successful integration and adoption of ML and DL techniques in healthcare settings. These include data quality and standardization, model interpretability, generalizability, ethical concerns, and interdisciplinary collaboration. By focusing on the future research recommendations outlined in this review, the healthcare community can continue to develop innovative ML and DL solutions that address these challenges, ultimately improving patient care and healthcare management processes. In conclusion, the growing importance of ML and DL applications in healthcare management is evident from the existing literature. As we continue to make progress in this domain, it is crucial to address the identified challenges and limitations, as well as foster interdisciplinary collaboration to maximize the potential benefits of ML and DL in healthcare management. By doing so, we can transform the way healthcare is delivered and managed, enhancing patient outcomes, and optimizing resources across the healthcare ecosystem.

References

1. Zou, K.H. and J.Z. Li. *Enhanced Patient-Centricity: How the Biopharmaceutical Industry Is Optimizing Patient Care through AI/ML/DL in Healthcare*. 2022. MDPI.
2. Khan, S. and T. Yairi, *A review on the application of deep learning in system health management*. Mechanical Systems and Signal Processing, 2018. **107**: p. 241–265.
3. Angehrn, Z., et al., *Artificial intelligence and machine learning applied at the point of care*. Frontiers in Pharmacology, 2020. **11**: p. 759.
4. Wiljer, D. and Z. Hakim, *Developing an artificial intelligence-enabled health care practice: rewiring health care professions for better care*. Journal of medical imaging and radiation sciences, 2019. **50**(4): p. S8–S14.
5. Pallathadka, H., et al., *Impact of machine learning on management, healthcare and agriculture*. Materials Today: Proceedings, 2021.
6. Javaid, M., et al., *Significance of machine learning in healthcare: Features, pillars and applications*. International Journal of Intelligent Networks, 2022. **3**: p. 58–73.
7. Battineni, G., et al., *Applications of machine learning predictive models in the chronic disease diagnosis*. Journal of personalized medicine, 2020. **10**(2): p. 21.
8. Shickel, B., et al., *Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis*. IEEE journal of biomedical and health informatics, 2017. **22**(5): p. 1589–1604.
9. Currie, G., et al., *Machine learning and deep learning in medical imaging: intelligent imaging*. Journal of medical imaging and radiation sciences, 2019. **50**(4): p. 477–487.
10. Pianykh, O.S., et al., *Improving healthcare operations management with machine learning*. Nature Machine Intelligence, 2020. **2**(5): p. 266–273.
11. Fairley, M., D. Scheinker, and M.L. Brandeau, *Improving the efficiency of the operating room environment with an optimization and machine learning model*. Health care management science, 2019. **22**: p. 756–767.
12. Callahan, A. and N.H. Shah, *Machine learning in healthcare*, in *Key Advances in Clinical Informatics*. 2017, Elsevier. p. 279–291.
13. Shailaja, K., B. Seetharamulu, and M. Jabbar. *Machine learning in healthcare: A review*. in *2018 Second international conference on electronics, communication and aerospace technology (ICECA)*. 2018. IEEE.
14. Malo, D.C., et al. *Skin Cancer Detection using Convolutional Neural Network*. in *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*. 2022. IEEE.
15. Collins, A. and Y. Yao, *Machine learning approaches: data integration for disease prediction and prognosis*. Applied Computational Genomics, 2018: p. 137–141.
16. Huang, Y., et al., *Application of machine learning in predicting hospital readmissions: a scoping review of the literature*. BMC medical research methodology, 2021. **21**(1): p. 1–14.
17. Nayariseri, A., et al., *Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery*. Current drug targets, 2021. **22**(6): p. 631–655.
18. Doppalapudi, S., R.G. Qiu, and Y. Badr, *Lung cancer survival period prediction and understanding: Deep learning approaches*. International Journal of Medical Informatics, 2021. **148**: p. 104371.
19. Qiu, H., et al., *Applications of artificial intelligence in screening, diagnosis, treatment, and prognosis of colorectal cancer*. Current Oncology, 2022. **29**(3): p. 1773–1795.
20. Fatima, N., et al., *Prediction of breast cancer, comparative review of machine learning techniques, and their analysis*. IEEE Access, 2020. **8**: p. 150360–150376.
21. Kourou, K., et al., *Applied machine learning in cancer research: A systematic review for patient diagnosis, classification and prognosis*. Computational and Structural Biotechnology Journal, 2021. **19**: p. 5546–5555.
22. Giger, M.L., *Machine learning in medical imaging*. Journal of the American College of Radiology, 2018. **15**(3): p. 512–520.

23. Esteva, A., et al., *A guide to deep learning in healthcare*. Nature medicine, 2019. **25**(1): p. 24–29.
24. Yu, K.-H., et al., *Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features*. Nature communications, 2016. **7**(1): p. 12474.
25. Di Noia, C., et al., *Predicting Survival in Patients with Brain Tumors: Current State-of-the-Art of AI Methods Applied to MRI*. Diagnostics, 2022. **12**(9): p. 2125.
26. Kopitar, L., et al., *Early detection of type 2 diabetes mellitus using machine learning-based prediction models*. Scientific reports, 2020. **10**(1): p. 11981.
27. Anderson, A.E., et al., *Electronic health record phenotyping improves detection and screening of type 2 diabetes in the general United States population: A cross-sectional, unselected, retrospective study*. Journal of biomedical informatics, 2016. **60**: p. 162–168.
28. Jaiswal, V., A. Negi, and T. Pal, *A review on current advances in machine learning based diabetes prediction*. Primary Care Diabetes, 2021. **15**(3): p. 435–443.
29. Shah, D., S. Patel, and S.K. Bharti, *Heart disease prediction using machine learning techniques*. SN Computer Science, 2020. **1**: p. 1–6.
30. Doppala, B.P., et al., *A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques*. Journal of Healthcare Engineering, 2022. **2022**.
31. Hannun, A.Y., et al., *Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network*. Nature medicine, 2019. **25**(1): p. 65–69.
32. Laad, M., et al., *Cardiac Diagnosis with Machine Learning: A Paradigm Shift in Cardiac Care*. Applied Artificial Intelligence, 2022. **36**(1): p. 2031816.
33. Ashfaq, A., et al., *Readmission prediction using deep learning on electronic health records*. Journal of biomedical informatics, 2019. **97**: p. 103256.
34. Artetxe, A., A. Beristain, and M. Grana, *Predictive models for hospital readmission risk: A systematic review of methods*. Computer methods and programs in biomedicine, 2018. **164**: p. 49–64.
35. Shi, P., A. Gangopadhyay, and P. Yu. *LIVE: A Local Interpretable model-agnostic Visualizations and Explanations*. in 2022 IEEE 10th International Conference on Healthcare Informatics (ICHI). 2022. IEEE.
36. Hu, C., et al., *Explainable Machine-Learning Model for Prediction of In-Hospital Mortality in Septic Patients Requiring Intensive Care Unit Readmission*. Infectious Diseases and Therapy, 2022. **11**(4): p. 1695–1713.
37. Reddy, B.K. and D. Delen, *Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology*. Computers in biology and medicine, 2018. **101**: p. 199–209.
38. Chen, T., et al., *Machine learning methods for hospital readmission prediction: systematic analysis of literature*. Journal of Reliable Intelligent Environments, 2022. **8**(1): p. 49–66.
39. Han, S., et al., *Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing*. Journal of Biomedical Informatics, 2022. **127**: p. 103984.
40. Lin, C., et al. *Early diagnosis and prediction of sepsis shock by combining static and dynamic information using convolutional-LSTM*. in 2018 IEEE International Conference on Healthcare Informatics (ICHI). 2018. IEEE.
41. Ahamed, F. and F. Farid. *Applying internet of things and machine-learning for personalized healthcare: Issues and challenges*. in 2018 International Conference on Machine Learning and Data Engineering (iCMLDE). 2018. IEEE.
42. Wang, L., et al. *Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation*. in Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2018.
43. Razzak, M.I., S. Naz, and A. Zaib, *Deep learning for medical image processing: Overview, challenges and the future*. Classification in BioApps: Automation of Decision Making, 2018: p. 323–350.

44. Zhou, S.K., et al., *A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises*. Proceedings of the IEEE, 2021. **109**(5): p. 820–838.
45. Litjens, G., et al., *A survey on deep learning in medical image analysis*. Medical image analysis, 2017. **42**: p. 60–88.
46. Sahiner, B., et al., *Deep learning in medical imaging and radiation therapy*. Medical physics, 2019. **46**(1): p. e1–e36.
47. Lareyre, F., et al., *Applications of artificial intelligence for patients with peripheral artery disease*. Journal of Vascular Surgery, 2022.
48. Sandhya, K. and J.J. James, *Prognosis and Diagnosis of Disease Using AI/ML Techniques, in Disruptive Developments in Biomedical Applications*. 2022, CRC Press. p. 37–51.
49. Futoma, J., S. Hariharan, and K. Heller. *Learning to detect sepsis with a multitask Gaussian process RNN classifier*. in *International conference on machine learning*. 2017. PMLR.
50. Lee, T.C., et al. *Clinical implementation of predictive models embedded within electronic health record systems: a systematic review*. in *Informatics*. 2020. MDPI.
51. Aledhari, M., et al., *Federated learning: A survey on enabling technologies, protocols, and applications*. IEEE Access, 2020. **8**: p. 140699–140725.
52. Cesario, A., et al., *Personalized clinical phenotyping through systems medicine and artificial intelligence*. Journal of personalized medicine, 2021. **11**(4): p. 265.
53. Hong, N., et al., *Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries*. Journal of biomedical informatics, 2019. **99**: p. 103310.
54. Zeng, Z., et al., *Natural language processing for EHR-based computational phenotyping*. IEEE/ACM transactions on computational biology and bioinformatics, 2018. **16**(1): p. 139–153.
55. Lee, C., J. Rashbass, and M. Van der Schaar, *Outcome-oriented deep temporal phenotyping of disease progression*. IEEE Transactions on Biomedical Engineering, 2020. **68**(8): p. 2423–2434.
56. Choi, E., et al., *Using recurrent neural network models for early detection of heart failure onset*. Journal of the American Medical Informatics Association, 2017. **24**(2): p. 361–370.
57. Valliani, A.A.-A., D. Ranti, and E.K. Oermann, *Deep learning and neurology: a systematic review*. Neurology and therapy, 2019. **8**: p. 351–365.
58. Kong, Y., et al., *Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier*. Neurocomputing, 2019. **324**: p. 63–68.
59. Castaldi, P.J., et al., *Machine learning characterization of COPD subtypes: insights from the COPD Gene study*. Chest, 2020. **157**(5): p. 1147–1157.
60. Zhang, X., et al., *OmiEmbed: a unified multi-task deep learning framework for multi-omics data*. Cancers, 2021. **13**(12): p. 3047.
61. Zaghlool, S.B. and O. Attallah. *A Review of Deep Learning Methods for Multi-omics Integration in Precision Medicine*. in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2022. IEEE.
62. Futoma, J., et al., *The myth of generalisability in clinical research and machine learning in health care*. The Lancet Digital Health, 2020. **2**(9): p. e489–e492.
63. Gulshan, V., et al., *Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs*. Jama, 2016. **316**(22): p. 2402–2410.

Bone Fracture Prediction Using Machine Learning and Deep Learning Techniques



Satya Vamsi Kumar Appala , S. V. V. D. Jagadeesh , M. Durga Satish , and B. Sridevi

Abstract Bones play a significant role in human body support, allowing us to use our muscles to walk, ride a bike, and hold a child. A bone fracture is a medical condition that occurs when a bone is cracked or broken due to an accident or trauma. Fractures can range in severity from a small crack in the bone to a complete break, where the bone is broken into two or more pieces. Accidents are one of the most common causes of bone fractures, and they can occur in a variety of ways, such as falls, sports injuries, car accidents, and workplace accidents. The severity of the fracture depends on the amount of force applied to the bone, the location of the fracture, and the age and health of the person.

To deal with this problem, it requires an essential need of prediction system for the detection of bone fractures in the human body. The artificial intelligence subfield of machine learning offers distinguished assistance in event prediction and trains its models using data from real-world occurrences.

In this chapter, ML and DL algorithms for predicting bone fractures using support vector machine, decision tree, random forest, and convolutional neural network (CNN) algorithms are discussed. Out of these, SVM outperformed with 86% accuracy.

Keywords Bone fracture · Support vector machines · Decision tree · Random forest · Convolutional neural network · X-ray images

S. Vamsi Kumar Appala (✉)

Department of Computer Science, B V Raju College, Vishnupur, Bhimavaram, AP, India

S. V. V. D. Jagadeesh · M. Durga Satish · B. Sridevi

Department of Computer Science & Engineering, Vishnu Institute of Technology, Vishnupur, Bhimavaram, India

e-mail: s.jagadeesh@vishnu.edu.in; durgasatish.m@vishnu.edu.in; sridevi.b@vishnu.edu.in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

325

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_24

1 Introduction

Bone fractures refer to a break or a crack in one or more bones in the body. This is a common medical problem that affects people of all ages and genders. The prevalence of bone fractures varies depending on various factors, including age, gender, lifestyle, and underlying medical conditions.

According to the World Health Organization, approximately one in three women and one in five men over the age of 50 years will experience a bone fracture due to osteoporosis, a condition that weakens bones and makes them more prone to breaking [8]. In addition to this, fractures can occur due to trauma or injuries, sports activities, falls, and other accidents.

The impact of bone fractures on an individual's quality of life can be significant, as it can cause pain, discomfort, and immobility. It can also result in a loss of productivity, decreased ability to perform daily activities, and increased healthcare costs. Moreover, some individuals may experience long-term complications such as chronic pain, disability, or even a decreased lifespan.

A bone fracture is a medical condition that occurs when a bone is cracked or broken due to an accident or trauma. Fractures can range in severity from a small crack in the bone to a complete break, where the bone is broken into two or more pieces. Some common symptoms of a bone fracture include pain, swelling, bruising, and difficulty moving the affected limb or body part. In some cases, there may be an audible cracking or popping sound at the time of the injury.

Bone fracture detection using machine learning algorithms is a rapidly developing field that has the potential to improve the accuracy and efficiency of fracture diagnosis. There are several approaches to using machine learning for bone fracture detection, but most methods rely on the analysis of medical imaging data such as X-rays, CT scans, and MRI scans.

2 Literature Review

D P Yadav (2021) [2] used deep neural network, and datasets from TCIA and the Indian Institute of Engineering Science and Technology were used. Using Softmax and Adam optimizer, he assesses the model's performance. Using fivefold cross-validation, the suggested model's classification accuracy for the healthy and broken bones is 92.44%.

Alhat Asmitha (2021) [3] employed a convolutional neural network (CNN) method to carry out edge detection on aggregate scales. The proposed system has been demonstrated to be significantly more noise-tolerant than the currently available edge detectors, and sufficiently resilient to retrieve the essential features and perform the required analysis on crucial regions of the images.

Tabassum Nahid Sultana (2022) used machine learning algorithms such as k-nearest neighbor, support vector machine (SVM), artificial neural network, and CNN classifiers for leg and hand bone fracture detection [4].

Yangling Ma (2021) introduces the crack-sensitive CNN (CrackNet), a new classification network that is sensitive to fracture lines. To determine if each bone region is broken, we first identify the 20 various types of bone regions included in X-ray images using faster region with convolutional neutral network (Faster R-CNN). A total of 1052 photos, 526 of which are broken and the remaining 882 are not, are utilized to test our system [1].

3 Dataset Description

A bone fracture X-ray image dataset typically contains a collection of medical images of human bones, specifically X-ray images of bones that have been fractured or broken. The dataset may also include X-ray images of non-fractured bones for comparison.

Table 1 describes the dataset collection which includes X-ray images of several joints that are fractured and unfractured. This dataset consists of 8863 X-ray images of both fractured and unfractured for training the model and 600 X-ray images of both fractured and unfractured for evaluating the model [9].

The training dataset consists of 4480 fractured samples and 4383 unfractured samples, while the test dataset contains 360 fractured samples and 240 unfractured samples. Fig. 1 shows an example of fractured and unfractured images in the dataset.

Having this information helps in understanding the class distribution within the dataset. It is important to have a balanced dataset, or at least have a representative number of samples for each class, to avoid biases and ensure that the model can learn from both fractured and unfractured cases effectively.

In this case, the training dataset appears to be relatively balanced, with a slight difference of 97 samples between the fractured and unfractured classes. This balance is essential for the model to learn the characteristics and features of both classes adequately.

For the test dataset, there are 360 fractured samples and 240 unfractured samples. However, it is worth mentioning that the dataset size alone does not determine the performance of the model. The quality and representativeness of the samples, as well as the complexity of the underlying patterns, also play a crucial role. It is always beneficial to gather a diverse and comprehensive dataset that covers various

Table 1 Train test dataset representation

Dataset	Train	Fractured	4480
		Unfractured	4383
	Test	Fractured	360
		Unfractured	240



Fig. 1 Fractured and unfractured images in the dataset

scenarios and variations of fractures to improve the model's ability to generalize to real-world cases.

Additionally, it is important to consider other factors such as data augmentation techniques and model architecture choices that can help address class imbalances and improve the model's performance on both fractured and unfractured samples.

4 Proposed Methodology

In the proposed methodology (Fig. 2), the set of trained fractured and unfractured images is preprocessed to avoid any noise existing in the images. After removing noise from the images, the images are applied to the models to perform feature extraction. Once the features are extracted, then the model is used for the prediction of identifying the fracture in the bones.

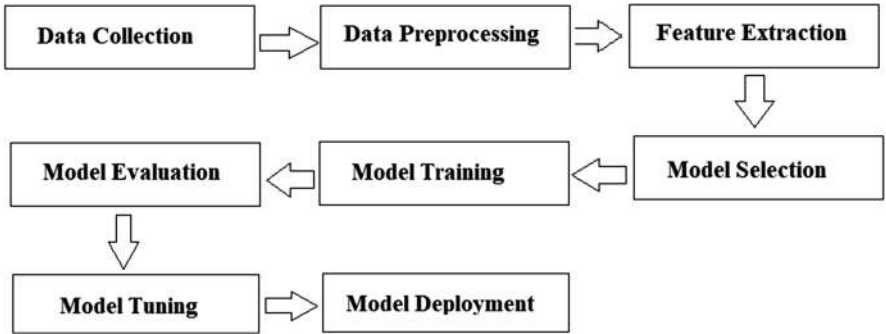


Fig. 2 Classification framework

5 Dataset Preprocessing and Feature Extraction.

(a) Data Preprocessing:

Data preprocessing is a common technique used in machine learning to transform raw data into a more suitable format for training a model. The goal of data preprocessing is to make the data easier to understand, visualize, and analyze and to reduce noise and variability in the data. In the context of image datasets, data preprocessing can include the following steps:

- 1. *Data cleaning*: This involves removing any corrupted or irrelevant images from the dataset.
- 2. *Resizing and cropping*: The images in the dataset may have different sizes and aspect ratios. Resizing and cropping the images to a uniform size can make it easier to train a model.
- 3. *Normalization*: Normalizing the pixel values of the images can make the training process more stable and efficient. This can involve subtracting the mean pixel value and dividing by the standard deviation.
- 4. *Data augmentation*: Data augmentation is a technique used to increase the size of the dataset by applying transformations to the images such as rotation, flipping, and zooming. This can help to improve the generalization of the model.
- 5. *Label encoding*: If the dataset has categorical labels, label encoding can be used to convert them into numerical values.

(b) Feature Extraction:

Here are the steps to apply feature extraction on an image dataset using a pretrained CNN:

- 1. *Load the pretrained CNN*: There are many pretrained CNNs available, such as VGG, ResNet, and Inception. We can load the pretrained CNN using a library such as Keras or PyTorch.

2. *Freeze the layers:* We can freeze the layers of the pretrained CNN so that they are not updated during training. This allows us to use the pretrained CNN as a feature extractor without modifying its weights.
3. *Extract features:* We can pass the images in our dataset through the pretrained CNN and extract the features from one of its layers. This will give us a set of features for each image in our dataset.
4. *Train a model:* We can use the extracted features as input to a new model such as an SVM or a fully connected neural network and train it to perform the specific task.

6 Implementation.

The specified machine learning and deep learning classifiers are used to test the model in this step. There have been a few models developed, and their accuracy has been verified. The classifiers listed below are utilized in this project.

6.1 Machine Learning Classifiers

After the training process, the training model was analyzed, and feature sets were introduced to the untrained model. Based on these feature sets, the model generated a set of predictions, which were subsequently compared with the valid labels to assess the testing accuracy of the model.

(a) SVM

For classification and regression tasks, the widely used machine learning algorithm SVM is utilized. It is a supervised learning method that excels at handling challenging classification issues in high-dimensional domains.

The SVM algorithm separates the various classes in the training data by creating a hyperplane or set of hyperplanes in a high-dimensional space. The distance between the hyperplane and the nearest data points from each class is determined in a way that maximizes the margin between the two classes.

It can handle data that can be split linearly and nonlinearly by utilizing kernel functions to translate the original input space to a higher-dimensional space. As a result, data that may not be linearly separable in the initial input space can be effectively classified using SVM [5].

(b) Classification

(i) Decision Tree Classifier:

A well-liked machine learning technique called decision tree (DT) is utilized for both classification and regression problems. It builds a DT model that is associated with the model in the form of a tree structure that

divides a dataset into more manageable subgroups. The outcome is a tree containing leaf nodes and decision nodes.

Recursively dividing the data into the features that best distinguish the classes, such as Gini impurity or entropy, is how the method operates. The algorithm first chooses the feature at each node that best distinguishes the data, and then it generates a new branch for each conceivable value of that feature. This procedure is repeated until all leaf nodes have pure data, or until a stopping requirement is satisfied.

By moving up the tree from the root node to a leaf node based on the values of the features for the new data, the DT model can be used to make predictions about it. The projected class for the input data is the class label connected to the leaf node [10].

(ii) *Random Forest Classifier:*

A well-liked machine learning method called random forest (RF) is utilized for both classification and regression problems. It is an ensemble learning technique that brings together various DTs to produce a reliable and precise model.

The approach works by building a forest of DTs, where each tree is trained using a randomly chosen set of characteristics on a part of the data. Each tree divides the data during training according to a certain criterion, like Gini impurity or entropy, to produce a set of rules that identify the class of a new instance.

By casting a vote on the predictions of each individual tree in the forest, the RF model's final prediction is decided. This lessens the effect of overfitting and lowers the variance, which enhances the model's accuracy and stability [11].

6.2 Deep Learning Classifiers

Deep learning algorithms for bone fracture prediction, CNNs, are commonly preferred due to their effectiveness in image analysis tasks. CNNs have shown remarkable success in various computer vision applications, including medical imaging.

Here are some popular CNN architectures that are often used as a starting point for bone fracture prediction:

1. *VGGNet*: VGGNet is a deep CNN architecture that achieved excellent performance on the ImageNet dataset. It has a simple and uniform structure with small 3×3 filters and max pooling layers. VGGNet is relatively easy to understand and implement.
2. *ResNet*: ResNet (short for Residual Network) introduced the concept of residual blocks, which enable the training of much deeper networks without suffering from vanishing gradients. ResNet architectures, such as ResNet-50 or ResNet-101, have been widely used in various image recognition tasks.

3. *Inception*: Inception, or the Inception-v3 model, introduced the idea of using multilevel feature extraction with filters of different sizes within the same convolutional layer. This architecture is known for its efficiency and has been employed in numerous applications, including medical imaging [7].

While these pretrained CNN architectures provide a good starting point, fine-tuning is usually necessary to adapt the models to the specific task of bone fracture prediction. This involves training the model on a custom dataset of bone fracture images to improve its accuracy and performance on this specific task.

(a) *CNN Classifiers*

CNN is a type of neural network architecture that is commonly used for image and video recognition and classification tasks. CNNs are inspired by the structure and function of the visual cortex in animals and use multiple layers of filters to extract hierarchical features from the input data.

The input to a CNN is typically an image, which is represented as a 3D array of pixel values (height \times width \times channels). The first layer of the network is a convolutional layer, which applies a set of filters to the input image to produce a set of feature maps. Each filter extracts a particular type of feature, such as edges or corners, from the input image [12].

The feature maps produced by the convolutional layer are then passed through a non-linear activation function, such as rectified linear unit, to introduce nonlinearity into the model. The output of the activation function is then passed through a pooling layer, which reduces the spatial dimensionality of the feature maps and helps to make the model more robust to small variations in the input.

The process of applying convolutional layers and pooling layers is repeated multiple times, forming a hierarchy of features that becomes increasingly abstract and high level as the network gets deeper. Finally, the output of the last convolutional layer is passed through one or more fully connected layers, which perform the classification or regression task [6].

(b) *Recurrent Neural Networks*:

A recurrent neural network (RNN) is a type of neural network that is designed to process sequential data. Unlike feed-forward neural networks, which process input data in a single pass, RNNs have connections that allow them to persist information across previous time steps. This makes them well-suited for tasks involving sequential data, such as time series analysis, natural language processing, speech recognition, and handwriting recognition.

The key feature of RNNs is their ability to maintain a hidden state that captures information from previous time steps. At each time step, the RNN takes an input and combines it with the hidden state from the previous time step to generate an output and update the current hidden state. This recurrent connection allows the RNN to capture dependencies and patterns across time.

The most common type of RNN is the long short-term memory (LSTM) network. LSTMs address the vanishing gradient problem that occurs in traditional RNNs by introducing a gating mechanism. This mechanism enables LSTMs to

selectively remember and forget information from the past, making them more effective at capturing long-term dependencies.

7 Results and Discussion

The images after performing preprocessing and feature extraction are applied with the above-mentioned models to test the classification accuracy of the training and test dataset images.

Table 2 shows the results for different sets of models.

1. The SVM model is trained for bone fracture detection using X-ray images; it learns the optimal hyperplane that separates the two classes (fracture and non-fracture) in the feature space. During prediction, the trained SVM model applies this learned hyperplane to new, unseen X-ray images to generate results. The SVM model generates results by analyzing the input X-ray image's features and determining its predicted class based on the learned hyperplane. It uses the concepts of separating hyperplanes and support vectors to make the classification decision. The result generated by the SVM model indicates whether the input X-ray image is likely to be a fracture or non-fracture case. SVM achieved the highest accuracy (0.86) among all the algorithms, indicating that it correctly classified 86% of the instances in the dataset. It also demonstrated high precision, recall, and F1 score, all of which were 0.86. Overall, SVM performed well in terms of accuracy and the ability to correctly predict positive instances.
2. The CNN algorithm generates results by learning hierarchical representations of features from the input X-ray images. By designing a deep network architecture with multiple convolutional and pooling layers, the model can capture complex patterns and relationships in the images. The training process optimizes the internal parameters of the model to improve its ability to discriminate between fracture and non-fracture cases. During prediction, the trained CNN model applies its learned features to new X-ray images and makes predictions based on the patterns it has learned during training. The output prediction represents the model's estimation of the probability or class label associated with the input image. CNN had the second-highest accuracy (0.80) and demonstrated a balanced performance with high precision (0.78), recall (0.80), and F1 score (0.76). This indicates that the CNN algorithm performed well in terms of

Table 2 Results of different models

Algorithm	Accuracy	Precision	Recall	F1 Score
<i>SVM</i>	0.86	0.88	0.86	0.86
<i>RF</i>	0.65	0.74	0.65	0.64
<i>DT</i>	0.75	0.74	0.75	0.74
<i>CNN</i>	0.80	0.78	0.80	0.76
<i>RNN</i>	0.63	0.66	0.63	0.64

accuracy and achieved a good balance between identifying positive instances and minimizing false positives.

3. The DT algorithm generates results by following the decision rules encoded in the tree structure. Each internal node of the tree represents a decision based on a specific feature, while the leaf nodes represent the final class labels. By comparing the feature values of the input X-ray image with the splits in the tree, the algorithm determines the appropriate path to take and reaches a leaf node that assigns the class label. The decision rules in the DT model are learned during the training process. The algorithm determines the optimal feature splits by considering measures of impurity or information gain to create the most discriminative decision boundaries. By using the trained DT model to traverse the tree structure, the algorithm can generate results by assigning the predicted class label based on the reached leaf node. DT performed moderately well with an accuracy of 0.75, a precision of 0.74, a recall of 0.75, and an F1 score of 0.74. This suggests that the DT algorithm achieved reasonable correctness in classification and showed a balanced performance between precision and recall.
4. The RF algorithm generates results by combining the predictions of multiple DTs. Each DT independently classifies the input X-ray image based on the selected features and its internal structure. The voting mechanism ensures that the collective decision of the ensemble of DTs is taken into account to determine the final prediction. By leveraging the diversity and collective wisdom of multiple DTs, the RF algorithm can provide robust and accurate predictions for bone fracture detection in X-ray images. RF had a lower accuracy (0.65) compared with the other algorithms. It showed moderate precision (0.74) but lower recall (0.65) and F1 score (0.64). This indicates that the RF algorithm may have struggled with correctly identifying positive instances.
5. The hybrid CNN-RNN model generates results by combining spatial and temporal information from X-ray images. The CNN component captures spatial features, while the RNN component models temporal dependencies. By integrating both aspects, the model can make predictions based on the joint understanding of spatial patterns and temporal context. The hybrid CNN-RNN model generates results for new, unseen X-ray images, and the preprocessed images are fed into the trained hybrid CNN-RNN model. The model applies its learned parameters and features to the input images, passes them through the CNN and RNN components, and produces predictions based on the learned patterns. The output of the final layer represents the model's prediction for each input image, indicating the likelihood of it being a fracture or non-fracture case. RNN had the lowest accuracy (0.63) among all the algorithms. It demonstrated moderate precision (0.66), recall (0.63), and F1 score (0.64). The RNN algorithm may have faced challenges in correctly identifying positive instances and achieving higher accuracy.

SVM and CNN performed the best in terms of accuracy and achieved a good balance between precision and recall. DT showed moderate performance, while RF and RNN had lower accuracy and faced challenges in correctly identifying positive instances.

8 Comparing the Models

When compared with other machine learning and deep learning models applied to the dataset, SVM is best suited for this system, with a mean accuracy of 86%.

Comparison of Algorithms Using Evaluation Metrics:

By observing Fig. 3, the SVM classifier gives the best accuracy 86% among all the remaining classifiers.

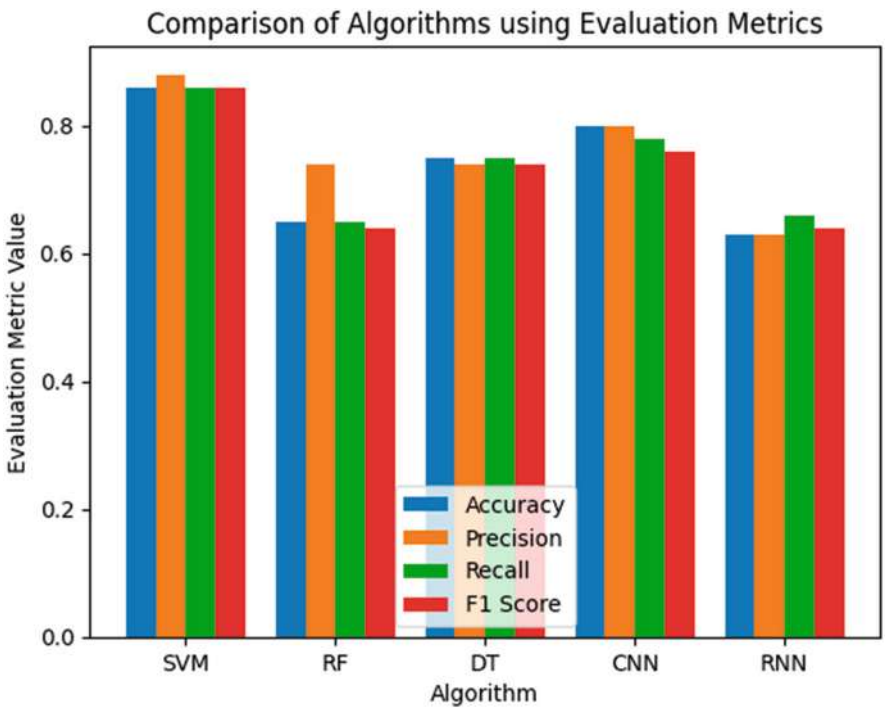


Fig. 3 Comparison of algorithms using evaluation metrics

9 Conclusion

Our study compared the performance of deep learning and traditional machine learning techniques for bone fracture prediction. Our results showed that the SVM model achieved the highest accuracy of 86%, followed by CNN with 80%, DT with 75%, RF with 65%, and RNNs with 63%. These findings suggest that deep learning techniques have the potential to improve the accuracy of bone fracture prediction models. However, traditional machine learning techniques can still provide useful results in certain cases. The choice of the appropriate technique depends on the specific characteristics of the dataset and the objectives of the study. Our study contributes to the growing body of literature on the application of machine learning techniques in healthcare and provides insights for future research in this area.

References

1. Yangli Ma, Yixin Luo Bone fracture detection through the two-stage system of Crack-Sensitive Convolutional Neural Network November 17, 2020, from <https://doi.org/10.1016/j.imu.2020.100452>.
2. D P Yadav, Sandeep Rathor, Bone Fracture Detection and Classification using Deep Learning Approach, July 17 2020
3. Alhat Asmita, Shinde Shunhangi, Bhor Sakshi, Pawale Kirti, P.R .Patil BONE FRACTURE DETECTION USING MACHINE LEARNING , IRJETS, Volume 3, Issue 12, December 2021
4. Tabassum Nahid Sultana, Asma Parveen, Efficient Bone Fracture Detection And Classification Using Machine Learning Approaches Volume 19, issue 2, June 2022
5. Vaishnav Kalbhor , Siddhant Mishra , Yash Walke , Bone Fracture Detection using CNN and SVM IJIRT Volume 8, Issue 3, August 2021
6. Kallimpudi Bhaskara sai kiran, B Satyasaivani BONE FRACTURE DETECTION USING CONVOLUTIONAL NEURAL NETWORKS, IJCRT, Volume 10, Issue 6, June 2022
7. Meena, T., & Roy, S. (2022). Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift. *Diagnostics (Basel, Switzerland)*, 12(10), 2420. <https://doi.org/10.3390/diagnostics12102420>
8. Sozen T., Ozisik L., & Basaran N. C. (2017). An overview and management of osteoporosis. *European Journal of Rheumatology*, 4(1), 46–56. <https://doi.org/10.5152/eurjrheum.2016.048>
9. Vuppala Adithya Sairam. (2023). Bone Fracture Detection using X-rays. Kaggle. <https://www.kaggle.com/datasets/vuppalaadithyasairam/bone-fracture-detection-using-xrays>
10. Sijia Chu, Aijun Chu, Lyuzhou Chen, Xi Zhang, Xiurong Shen, Wan Zhou, Shandong Ye, Chao Chen, Shilu Zhang, Li Zhang, Yang Chen, Ya Miao, Wei Wang, Machine learning algorithms for predicting the risk of fracture in patients with diabetes in China, *Heliyon*, Volume 9, Issue 7, 2023, e18186, ISSN 2405-8440. <https://doi.org/10.1016/j.heliyon.2023.e18186>.
11. Mohanty, Sareeta & Senapati, Manas. (2023). Fracture detection from X-ray images using different Machine Learning Techniques. 1–6. 10.1109/CCPIS59145.2023.10291652.
12. Sultan Mamun, Md Al Amin, Ajlan Jamal Ali and Jingwei Li. (2024). Bone Fracture Detection from X-ray Images using a Convolutional Neural Network (CNN). 10.20944/preprints202410.1320.v1.

Plant Disease Detection Using Modern Deep Learning Approach: YOLOv7



Ayan Banerjee, Arkaprava Mazumder, Ayush Kumar Shaw,
Udit Narayana Kar, Sovan Bhattacharya, and Chandan Bandyopadhyay

Abstract Agriculture is an essential part of every country's economy. However, it is widely affected by diseases and natural disasters. Natural disasters are inevitable, but in the case of diseases, if they can be detected in their early stages, their growth can be stopped with appropriate measures and action, boosting agricultural yields and reducing crop losses. In this chapter, we attempt to deploy the YOLOv7 model for the automated disease detection of plants. YOLOv7 uses deep neural networks and the Pytorch model to identify and classify items quickly and accurately. The model we have used is already available online, and we have further optimized it to our needs to obtain the highest accuracy. Finally, we trained the model with the perfect batch size and epochs that give an accuracy score of 94%, which is a good score for multiclass real-time detection of diseased plants.

A. Banerjee (✉) · A. Mazumder · A. K. Shaw
Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal,
India
e-mail: Ayan.B@labvantage.com

U. N. Kar
Vellore Institute of Technology, Vellore, Andhra Pradesh, India
e-mail: uditnarayana.k@vitap.ac.in

S. Bhattacharya
Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal,
India

Department of CSE, National Institute of Technology, Durgapur, West Bengal, India

C. Bandyopadhyay
Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal,
India

Department of CSE, University of Bremen (Former Post Doctoral Fellow), Bremen, Germany

Keywords Deep learning · Digital image processing · TensorFlow · YOLOv7 · Real-time image processing · Plant diseases detection

1 Introduction

Plant diseases are a severe hazard to world agriculture because they result in significant crop losses, decreased food production, and adverse economic effects. Early detection and prompt action are essential to halting the spread of illnesses and reducing their harmful effects on crop production. Traditional disease detection techniques, which rely on expert manual inspection, take a lot of time, are subjective, and are frequently constrained by the knowledge and accessibility of experienced employees. A rising number of people are interested in using computer vision and deep learning techniques to automate identifying plant diseases. In object detection, deep learning has shown exceptional performance in various computer vision applications. YOLOv7, an advancement of the well-known YOLO (You Only Look Once) framework, provides an effective and precise object detection solution, making it a good choice for applications that detect plant diseases. The framework of YOLOv7 is demonstrated clearly in Fig. 1.

1.1 Motivation and Objectives

The research aims to investigate how the YOLOv7 algorithm can be used in plant disease identification. The urgent problem of plant diseases, which represent a severe risk to the security of the world’s food supply, is the subject of this essay. First, the study aims to produce a precise and effective solution for real-time plant disease identification using YOLOv7, a cutting-edge object detection algorithm. The YOLOv7 model will be trained using a sizable dataset of annotated photos of various plant illnesses and healthy plants. Second, by contrasting it with previously

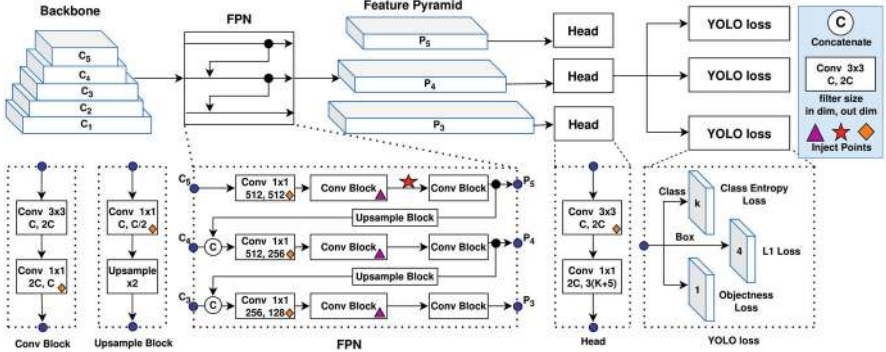


Fig. 1 YOLOv7 architecture

used techniques, the research assesses the model's performance regarding detection accuracy, speed, and robustness. The goal is to develop a reliable and valuable tool to help farmers, agricultural specialists, and researchers identify and control plant diseases early on, improving crop yields and reducing financial losses.

This chapter mainly aims to investigate the use of YOLOv7 in this field to create a reliable and effective tool that can help farmers and researchers identify diseases quickly so that they can put effective disease management plans into action to protect crop health. In an overall sense, it can be objectified that this chapter intends to advance agricultural practices by enabling early disease diagnosis and intervention, thereby enhancing crop yields and maintaining food security while also contributing to the area of plant pathology.

1.2 Challenges and Contribution

There are many difficulties that the study on plant disease detection using YOLOv7 must carefully analyze and overcome. First, gathering a comprehensive and varied dataset with precisely annotated photos depicting different plant illnesses and healthy plants is a challenging process that requires much work and subject-matter expertise. Second, it can be difficult to effectively identify and categorize plant illnesses due to their intrinsic complexity and unpredictability, which can appear in various ways and vary depending on the plant species and environmental factors. Finally, assessing the model's success necessitates developing relevant metrics, contrasting it with other approaches, and confirming its generalization abilities on hypothetical data. These issues must be resolved to assure the dependability, sturdiness, and practical application of the plant disease detection system employing YOLOv7. This chapter is arranged into the following sections: Introduction, followed by Research Review, then Methodology, Experimental Setup, Results, and Discussion, and finally Conclusion and Future Work, closing with Acknowledgments. In the end, the references to the Research Review are mentioned.

2 Research Review

In the field of research, several existing works aim at detecting plant diseases using different approaches. Some of these works are discussed in this section.

To assist the farmer in identifying the papaya illness, Md. Sagar Hossen et al. [1] suggested a method that uses a single algorithm, similar to convolutional neural networks (CNNs), to forecast papaya illness. With 99% accuracy, Payal Bose et al. [2] created an SVM-based system for diagnosing leaf diseases in medicinal plants. A target detection system based on YOLOv5 has been put into place by Hosting Zhang et al. [3] on forward-looking sonar pictures. Shuun Luo et al. [4] proposed a method that primarily defines an enhanced YOLOv5-aircraft model for multiple detections

at once. Bhattacharya et al. [5] proposed an advanced approach to detecting plant disease from the images of the leaves of individual plants. They optimized the Keras deep learning model and accomplished the detection task. Zhao Zheng et al. [6] proposed a YOLOv3 model with mAP-based 69.74% accuracy to detect Bearing-Cover. With a target picture size of 200×200 , a deep learning technique has been put out by Imdadul Haque et al. [7] to predict Malabar Nightshade sickness. Lakshma-narao et al. [8] have suggested a classification and prediction method for plant diseases using ConvNets. A YOLOv5 model was proposed by Zhaoyi Chen et al. [9] to identify plant diseases with an accuracy of 86.8%. Samhan et al. [10] developed a CNN-based Alzheimer's disease classification system with 100% training accuracy, 0.0012 training loss, and 97% validation accuracy. Bhattacharya et al. [11] proposed different ML techniques' capabilities to predict the future of budding researchers. A bowl of rice leaf disease detection technique has been proposed by Nguyen Thai-Nghe et al. [12]. Although this model for detecting rice illness is based on an Android application, no testing outcome can characterize the suggested model's level of perfection. Based on the CNN model and a refined model known as MobileNetv2, Murat Koklu et al. [13] created a classification of grapevine leaf diseases. A CNN-based BLeafNet model has been created by Shreyan Ganguly et al. [14] to detect plants from leaf pictures. This method uses the ResNet-50 architecture and RGB values to extract features. Ashwinkumar et al. [15] developed an automated plant leaf disease classification using MobileNet based on CNN to lessen the farmer's loss. Imdadul Haque et al. [16] developed a deep learning method for identifying pepper diseases utilizing pretrained models from Keras, including NASNetMobile, InceptionResNet-V2, ResNet152-V2, Xception, MobileNet-V2, and VGG-19. Sagar Hossen et al. [17] introduced a machine learning method to examine the risk factor for liver disease that has a satisfactory accuracy of 86.14% and is based on Random Forest algorithms. Using information from PlantVillage AlexNet and GoogleLeNet, Mohanty et al. [18] found 14 crop species in their model. In those 14 species, 26 plant diseases were discovered. Fuentes et al. [19] used photos manually obtained with a camera and a small number of samples to identify plants with disease and insect infestations. Mishra et al. [20] applied deep CNN on the Plant Village dataset and a few real-time photos.

So from the above study, we have observed several pre-existing studies related to plant disease detection. Different technologies and models, such as CNN, Keras, YOLOv5, and so on, are used. In our research, we have tried the YOLOv7 model, which is more advanced and gives more accurate results even in real-time detection. Moreover, this model is never used by researchers for plant disease detection activities.

3 Methodology

There are numerous stages to the suggested YOLOv7-based plant disease detection process, which are discussed in this section in three categories. The first step is to develop a vast and varied library of annotated photos depicting plant illnesses and

healthy plants. The trained model is then adjusted to deploy devices with limited resources frequently utilized in agricultural contexts. The model's performance is then assessed using various criteria compared with other approaches, and its ability to generalize is verified using hypothetical data. The suggested methodology provides a workable response to the problems associated with plant disease detection, enabling prompt and precise disease identification for better crop management.

3.1 Data Collection

The *PlantDoc* dataset is downloaded as the primary dataset for our experiment. The dataset includes more than 54,000 high-resolution photos of 15 distinct plant species, tagged with 26 disease classes and one healthy class. The pictures were gathered from various places, such as open-access plant pathology collections, field research, and agricultural research facilities. The dataset is a helpful resource for the research community, fostering improvements in plant disease management and detection.

3.2 Data Preprocessing

The dataset now had to be preprocessed and fine-tuned to best fit different image processing models. The steps involved in the preprocessing stage are discussed below:

- *Image Augmentation*: The images of the datasets are in their raw state, i.e., different images are of different sizes, pixels, and qualities. Hence, augmentation is done on the whole dataset to normalize all the images before training and proceeding with further steps.
- *Remove Erroneous Data*: In the downloaded dataset, there are several images that are corrupted or only black photos. Those image files have to be deleted manually as irrelevant data will decrease the accuracy of the deep learning models.
- *Annotate the Whole Dataset Manually*: The YOLOv7 model requires an annotated or labeled image dataset, as discussed before. We have to manually prepare the annotated data carefully to get accurate results.

4 Experimental Setup

In this section, we have discussed the two instances that we have considered for the experimental purpose of obtaining the best possible model for our objective to be achieved.

- *Variation of Epoch with Fixed Batch Size:* Here, we have implemented our model with different values of epoch and a fixed batch size. Through this experiment, we obtained accuracy concerning the various epochs and studied the change in accuracy of our model with the change in the value of the epochs.
- *Variation of Batch Size with Fixed Epoch:* Just like we discussed in the previous section, we will execute our proposed model with varying batch size values that keep the epoch value fixed. Following this, we will record the accuracy obtained for each of the considered batch sizes and inspect the fluctuation in accuracy with the alternation in batch sizes.
- *Comparative study:* Performed a comparative study of our model with previous existing work by Chen et al. [9] that performed plant disease detection using the YOLOv5 framework.

5 Results and Discussion

The following section consists of the results obtained from the experiments performed based on the instances mentioned in the previous section and a comparison study of the model we have used with models that have been used by researchers before. For the comparative study, we have considered previous research results on the YOLOv5 framework used for plant disease detection.

5.1 Results at Different Epochs for Fixed Batch Size

In this condition, we kept the batch size fixed at 15 but varied the epochs randomly until we got a maximum accuracy score. The results obtained in this case are presented with the help of Table 1. It can be observed that at 290 epochs, the models give a maximum accuracy score of 94%. After 290 epochs, the accuracy scores start decreasing gradually. Notably, 290 epochs are the best for training our dataset on the YOLOv7 model. Figure 2 illustrates the fluctuation in accuracy with change in epochs by comparing the accuracy.

Table 1 Accuracy of varying epochs at fixed batch size (15)

Accuracy at varying epoch at fixed batch size (15)		
Batch	Epoch	Accuracy score
15	50	0.56
15	90	0.68
15	150	0.55
15	200	0.85
15	290	0.94
15	300	0.88

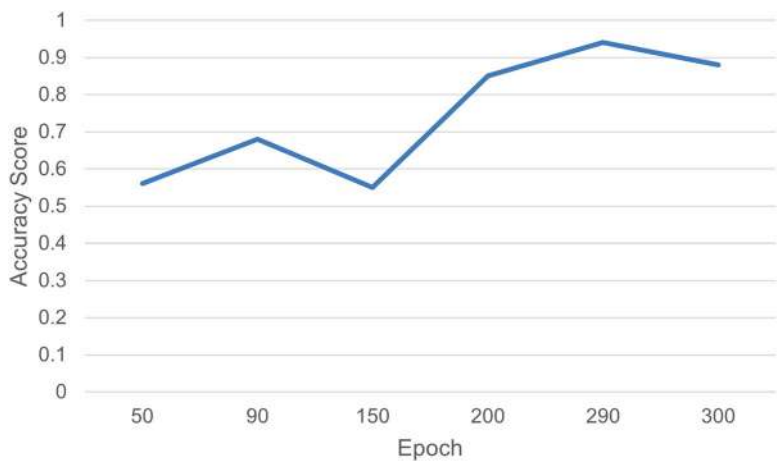


Fig. 2 Varying epochs at fixed batch size

Table 2 Accuracy scores of varying batch size at fixed epoch (290)

Accuracy at varying batch size at fixed epoch (290)		
Batch	Epoch	Accuracy score
5	290	0.65
10	290	0.75
15	290	0.94
20	290	0.81
25	290	0.77

5.2 Results at Different Batch Sizes for Fixed Epoch

In this case, we kept the epochs fixed at 290 and varied the batch size to 5, 10, 15, 20, and 25. The outcomes of this experiment are shown in Table 2. From the table, it is concluded that with a batch size of 15, the YOLOv7 model shows greater accuracy than other batch sizes. Figure 3 represents the distortion in the values of accuracy with change in the batch size, illustrating the formation of the highest peak at batch size 15.

5.3 Results Obtained from Comparative Study

In the article of Chen et al. [9] they performed the plant disease detection task using the YOLOv5 framework and obtained a maximum accuracy of 86.8% testing accuracy, whereas our YOLOv7 framework gives a maximum accuracy of 94%. We represented the accuracy variation with the help of the bar plot in Fig. 4. Hence, we can say that YOLOv7 is a more accurate framework than YOLOv5.

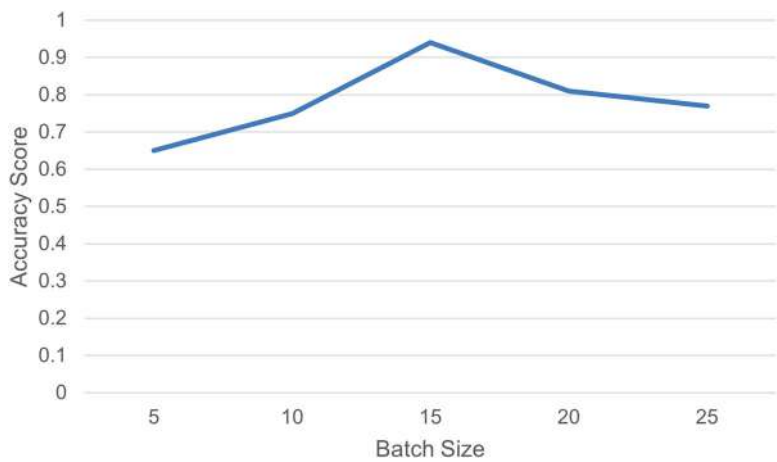


Fig. 3 Varying epochs at fixed batch size

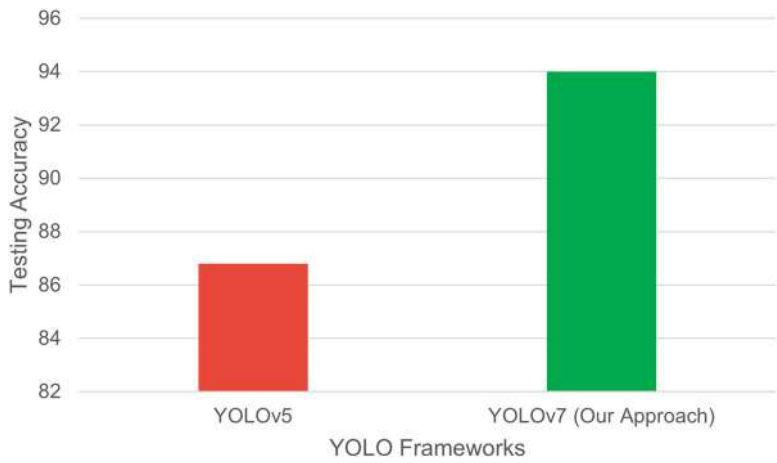


Fig. 4 Accuracy variation of YOLOv5 and YOLOv7 (comparative study)

From the above cases, we can conclude that considering batch size as 15 with 290 epochs will be a perfect choice for training the YOLOv7 model with the annotated PlantDoc dataset to get the highest accuracy score and precise detection task.

6 Conclusion and Future Scope

After the successful execution of the experiments, we derived that when 290 epochs were run on a batch size of 15, an accuracy of 94% was acquired. This gives us our

model, which is ready to detect plant diseases. Performing the comparative study, we can also conclude that YOLOv7 can give more accurate and precise results than other previous versions of the YOLO framework. For future work, we can deploy this model on a Raspberry Pi for detection through the camera, which will be attached to a drone for mobility over large agricultural fields to capture images for our model to detect diseases. The data obtained from the inspection will be recorded and uploaded in the app designed for this purpose, along with the GPS coordinates, for the farmers to acknowledge which part of their land is disease-affected so that they can use appropriate measures to eradicate the problems and gain maximum yield from their fields. Overall, this will uplift both the personal and global economies, which are affected due to the loss of crop yields because of the lack of detection or late detection of the diseases.

Acknowledgments

- Ethics Approval
The authors confirm that the work described has not been published before, and it is not under consideration for publication elsewhere.
- Consent to Participate
The authors agree to participate in the conference.
- Conflict of Interest
We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Hossen, M.S., Haque, I., Islam, M.S., Ahmed, M.T., Nime, M.J., Islam, M.A.: Deep learning based classification of papaya disease recognition. In: 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), pp. 945–951 (2020). IEEE
2. Bose, P., Dutta, S., Goyal, V., Bandyopadhyay, S.K.: Leaf diseases detection of medicinal plants based on support vector machine classification algorithm. *J. Pharm. Res. Int* **33**, 111–119 (2021)
3. Zhang, H., Tian, M., Shao, G., Cheng, J., Liu, J.: Target detection of forward-looking sonar image based on improved yolov5. *IEEE Access* **10**, 18023–18034 (2022)
4. Luo, S., Yu, J., Xi, Y., Liao, X.: Aircraft target detection in remote sensing images based on improved yolov5. *IEEE Access* **10**, 5184–5192 (2022)
5. Bhattacharya, S., Banerjee, A., Ray, S., Mandal, S., Chakraborty, D.: An advanced approach to detect plant diseases by the use of cnn based image processing. In: *Innovations in Computer Science and Engineering: Proceedings of the Tenth ICICSE, 2022*, pp. 467–478. Springer, Singapore (2023)
6. Zheng, Z., Zhao, J., Li, Y.: Research on detecting bearing-cover defects based on improved yolov3. *IEEE Access* **9**, 10304–10315 (2021)
7. Haque, I., Mojumdar, M.U., Chakraborty, N.R., Rana, M.S., Siddiquee, S.M.T., Ashik, M.M.H.: Malabar nightshade disease detection using deep learning technique. In: *Sustainable Advanced Computing: Select Proceedings of ICSAC 2021*, pp. 173–183. Springer, Singapore (2022)

8. Lakshmanarao, A., Babu, M.R., Kiran, T.S.R.: Plant disease prediction and classification using deep learning convnets. In: 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), pp. 1–6 (2021). IEEE
9. Chen, Z., Wu, R., Lin, Y., Li, C., Chen, S., Yuan, Z., Chen, S., Zou, X.: Plant disease recognition model based on improved yolov5. *Agronomy* **12**(2), 365 (2022)
10. Samhan, L.F., Alfarrar, A.H., Abu-Naser, S.S.: Classification of alzheimer's disease using convolutional neural networks (2022)
11. Bhattacharya, S., Banerjee, A., Goswami, A., Nandi, S., Pradhan, D.K.: Machine learning based approach for future prediction of authors in research academics. *SN Computer Science* **4**(3), 1–11 (2023)
12. Thai-Nghe, N., Tri, N.T., Hoa, N.H.: Deep learning for rice leaf disease detection in smart agriculture. In: Artificial Intelligence in Data and Big Data Processing: Proceedings of ICABDE 2021, pp. 659–670. Springer, Cham (2022)
13. Koklu, M., Unlersen, M.F., Ozkan, I.A., Aslan, M.F., Sabanci, K.: A cnn-svm study based on selected deep features for grapevine leaves classification. *Measurement* **188**, 110425 (2022)
14. Ganguly, S., Bhowal, P., Oliva, D., Sarkar, R.: Bleafnet: a bonferroni mean operator based fusion of cnn models for plant identification using leaf image classification. *Ecological Informatics* **69**, 101585 (2022)
15. Ashwinkumar, S., Rajagopal, S., Manimaran, V., Jegajothi, B.: Auto- mated plant leaf disease detection and classification using optimal mobilenet based convolutional neural networks. *Materials Today: Pro- ceedings* **51**, 480–487 (2022)
16. Haque, I., Islam, M.A., Roy, K., Rahaman, M.M., Shohan, A.A., Islam, M.S.: Classifying pepper disease based on transfer learning: A deep learn- ing approach. In: 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), pp. 620–629 (2022). IEEE
17. Hossen, M.S., Haque, I., Sarkar, P.R., Islam, M.A., Fahim, W.A., Khatun, T.: Examining the risk factors of liver disease: A machine learning approach. In: 2022 7th International Conference on Communication and Electronics Systems (ICCES), pp. 1249–1257 (2022). IEEE
18. Mohanty, S.P., Hughes, D.P., Salath/e, M.: Using deep learning for image- based plant disease detection. *Frontiers in plant science* **7**, 1419 (2016)
19. Fuentes, A., Yoon, S., Kim, S.C., Park, D.S.: A robust deep-learning- based detector for real-time tomato plant diseases and pests recognition. *Sensors in Agriculture: Volume 1* **17**, 153 (2019)
20. Mishra, S., Sachan, R., Rajpal, D.: Deep convolutional neural network based detection system for real-time corn plant disease recognition. *Procedia Computer Science* **167**, 2003–2010 (2020)

Analysis of the Life Insurance Business Performance Based on COVID by Using Machine Learning Algorithms



P. Nithya, C. D. Nandakumar, and S. Srinivasan

Abstract COVID-19 has had a worldwide impact. Economically, it has a ripple effect on the entire financial market. COVID has caused significant problems for the majority of insurance companies. COVID had a substantial impact on the life insurance industry. Most people nowadays have both life and health insurance based on how well that insurance company is performing. Because the majority of them purchased insurance, the company's share capital and risk are both increasing (Nti et al. *Artif Intell Rev* 53(4):3007–3057, 2020). As the share price of the insurance company fluctuates, insurance uses a hedging concept that minimizes risk while maximizing profit. Because the share price is volatile, it is difficult to forecast. Stock prediction is one of the most difficult tasks in artificial intelligence (Nti et al. *J Big Data* 7(1), 2020). Consequently, using machine learning techniques, it is possible to forecast how the insurance industry will perform on the day of its closing price. We collected data from Yahoo Finance for HDFC Life Insurance, Star Health, and Allied Insurance, SBI Life Insurance, ICICI Prudential Life Insurance, and Life Insurance Corporation of India from January 1, 2018, to December 31, 2022. The methodology of work is carried out using various models of machine-learning algorithms. The algorithms used to analyze company performance are Bayesian ridge regression, support vector machine, random forest regression, multinomial Naive Bayes, XGBoost regression, and Bayesian ridge regression. By employing these algorithms to analyze future share prices and boost the accuracy of stock price predictions (Nti et al. *Open Comput Sci* 10:153–163, 2020), investors can use the analysis to make investments in the company as well as buy insurance policies. They can also determine the most suitable forecasting models to produce the best estimations with the least amount of error (Nti et al. *Appl Comput Syst* 25(1):33–42, 2020).

P. Nithya (✉) · C. D. Nandakumar · S. Srinivasan
Department of Mathematics & Actuarial Science, B.S. Abdur Rahman Crescent Institute of Science & Technology, Chennai, Tamil Nadu, India
e-mail: cdnandakumar@crescent.education; srinivasan@crescent.education

Keywords Hedging · Machine-learning algorithms · Forecasting · COVID-19 · Artificial intelligence

1 Introduction

Today's economies rely heavily on the stock market. The rupee rate will either appreciate or depreciate based on market performance. Prices for commodities such as gold, curd oil, stocks, and commodities, among others, may rise or fall. However, due to COVID, the entire world will suffer a significant loss in 2019. COVID has had an impact all over the world [5]. Because of COVID, both developed and developing countries face numerous challenges. The majority of the insurance providers were impacted by this circumstance. Many life insurance companies are doing extremely well following the impact of COVID. The primary reason for this is that prior to COVID, only 40–50% of people might have insurance policies, whereas, after COVID, 80–90% of people were to have insurance plans. Therefore, it is clear that the company has a high risk while also receiving a sizable amount of premium from the policyholder. Hedging is the process of reducing risk while increasing profit [6]. The stock market's future trend is predicted by using linear regression, exponential smoothing, and time series forecasting. Identify the predicted stock market trends for the next month and also identify the accuracy of the share price of the company [7].

It is extremely difficult and dynamic to predict stock market volatility. The long-term stock price trend is chosen using a linear multi-factor model. The price trend is examined using a variety of feature algorithms [8]. By analyzing various algorithms, random forest is excellent for building a long-term portfolio [9]. There are several causes of stock market volatility, including macroeconomic factors, inflation, and changes in currency. This study makes predictions about stock price changes based on technical indicators, macroeconomic data, fundamental market information, and so on [10]. A life insurance company performs exceptionally well when using the concept of hedging. An examination of how the company performs well by implementing machine learning algorithms. Any problem can be solved more easily by humans. Machines can incorporate human creativity and expertise through the use of various algorithms [11]. Umer Ghania [12] especially focuses on investment in the stock market. Stocks are the most popular financial instrument for constructing a portfolio. Constructing a long short-term memory (LSTM)-based model to forecast closing index prices. Humans' needs were met and machines made their lives much easier. Numerous data issues can be solved by machine learning methods. It aids in foreseeing and predicting the future [13]. The outcomes demonstrate that the single-layer LSTM model offers a superior fit and excellent prediction accuracy [14].

2 Description of Data

Due to market fluctuations, volatility, and non-stationary data, analyzing stock market prices is a difficult task nowadays. It is possible to predict stock prices with artificial intelligence. In this study, various life insurance company types are analyzed in terms of how they handle the COVID situation and how well they are currently doing. To assess the performance of the company using various types of machine learning algorithms, financial information was gathered from Yahoo Finance.

The historical data of the five life insurance companies have been compiled from Yahoo Finance—HDFC Life Insurance Company Limited, Star Health and Allied Insurance Company Limited, SBI Life Insurance Company Limited, ICICI Prudential Life Insurance Company Limited, and Life Insurance Corporation of India from January 1, 2018, to December 31, 2022. The data include stock price information such as Open, High, Low, Close, Adjusted Close, and Volume. The stock's closing price for the day has been the only information derived. The stock price is predicted based on these variables. These parameters are used to train and test the model. Only ICICI Prudential Life Insurance Company Limited, SBI Life Insurance Company Limited, and HDFC Life Insurance Company Limited were listed prior to January 1, 2018. However, Star Health and Allied Insurance Company Limited was listed on December 10, 2021, and Life Insurance Corporation of India was listed on May 17, 2022.

3 Algorithm Selection

The present work has to execute the machine learning algorithms in our overall stock market closing price of the insurance company. As mentioned, we gathered data from Yahoo Finance which contains Open, High, Low, and Closing prices and 52-week low and high share prices. There are many machine learning algorithms available. Machine learning means classifying the data and making decisions based on the data. It is used to train machines and perform actions. The computer is to act like a human being by giving training based on past experience and predicted data. There are mainly four types, namely, supervised, semi-supervised, unsupervised, and reinforcement learning. Supervised learning is known as labeled data which means that the analysis part is based on the input and output of the dataset. Unsupervised learning algorithms are unlabeled data—to train the machine with some inputs, but the output is not known. Semi-supervised is a combination of both supervised and unsupervised learning. It helps to increase the accuracy of the model. Reinforcement learning algorithms are based on feedback. Based on the outcome, the next step of the plan is decided. On the analysis of the past records, receive feedback after every step and decide the next step to go.

The algorithms used to analyze company performance are Bayesian ridge regression, support vector machine (SVM), random forest regression, multinomial Naive Bayes, and XGBoost regression.

3.1 Bayesian Ridge Regression

One of the consistently attractive compressive methods to lessen the effects of multicollinearity in both linear and nonlinear regression models is ridge regression. When there are solid or nearly solid direct relationships between the indicator factors, this is referred to as multicollinearity [3].

3.2 Support Vector Machine

One of the more efficient machine learning algorithms for predicting the future dataset is the SVM algorithm. Using the available data, the main objective is to draw the best line or decision boundary.

It allows you to plot raw information on an n-dimensional graph. SVM is actually mainly used for tasks involving classification and regression. Share market performance measures are employed to plot multidimensional coordinate planes in this case. This algorithm represents the most robust and effective financial tool for predicting the stock market [4].

3.3 Random Forest Regression

An algorithm for supervised learning called random forest regression employs the outfit learning approach for regression. The supervised learning algorithm includes forecasting from various algorithms to produce a more precise forecast than a single product. KNN and decision tree are related to this technique. To evaluate the effectiveness of the model, it enables us to solve the classification and regression tasks. It combines numerous decision trees with numerous subsets of the dataset. The model is more effective when the input is based on the average prediction of all the trees. Higher accuracy is obtained and overfitting is avoided because of the larger number of trees in the forest.

3.4 Naive Bayes

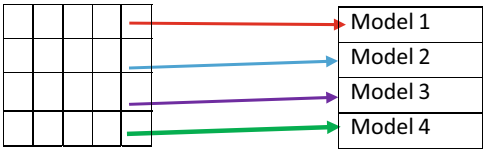
This is one of the simplest and most successful supervised learning algorithms. The Bayes theorem served as the foundation for this idea. This model provides rapid predictions with good accuracy. It has multidimensional training datasets. This model forecasts results based on the likelihood that an object will occur.

3.5 XGBoost Regression

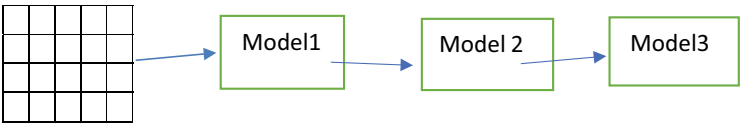
Regression predictive modeling can be done with the help of XGBoost, an appropriate gradient-boosting development—how to implement the best practice methodology of frequent k-fold cross-validation to analyze an XGBoost regression model and how and where to fit the accuracy of the model that can be used to forecast new data [5].

3.6 Bagging and Boosting

The characteristics of bagging and boosting are similar but not the same. When bagging, they stored all of the data and divided it according to machine learning models, line by line.



Boosting means the entire data process through all the models.



4 Methodology

To examine the performance of five insurance companies using machine learning models, the volatility of each company's closing stock price is determined. The majority of the time, share price data may not be stationary. Therefore, first, smooth the data for future prediction. Time series is a useful technique for seeing and predicting future trends based on data collected over a period of time. A specific amount of time is used to collect the data for the time series. When a time series is referred to as stationary, it signifies that the mean, variance, and autocorrelation do not change over time. Over time, it does not do anything. It maintains its stability over time. The data might be transformed into stationery using a variety of well-liked methods. Augmented data are transformed into stationery using the Dickey-Fuller and Kwiatkowski-Phillips-Schmidt-Shin tests. One reason to stabilize the data before analysis is stationary; otherwise, the results and findings will be inaccurate. It produces false information. Non-stationery data are not reliable. It is unpredictable and cannot be predicted or modeled. Over time, the mean, variance, and autocorrelation are not constant. The conclusion can be false given the non-stationary data. From 2018 to 2022, the following data were obtained from Yahoo Finance. High fluctuations can be seen in the data. To begin, convert the data into stationery for smoothing reasons.

Figure 1 depicts the stationary data points of HDFC Life Insurance Company. There were extremely large fluctuations in the data from 2019 to 2022 as a result of COVID. The closing price of the shares abruptly dropped after 2020, and the company only gradually recovered. The difference in closing price is currently not very high.

The stationary data points for SBI Life Insurance Company are shown in Figure 2. The data fluctuated greatly between 2019 and 2022 as a result of COVID. The stock's closing price abruptly dropped after 2020, and the corporation only gradually recovered. Now the variation in closing price is not too much high. From 2020 on, the closing stock price progressively rose; before and after COVID, the price fluctuated.

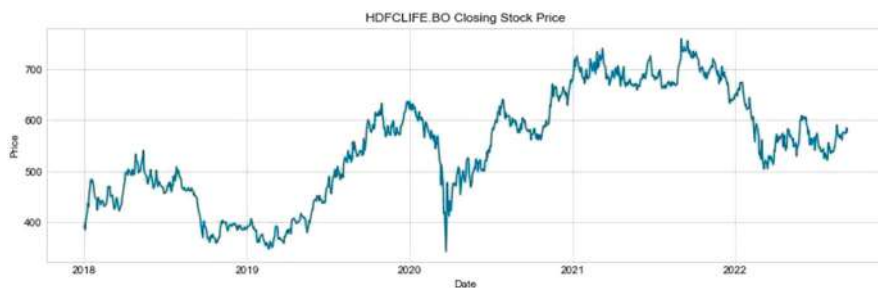


Fig. 1 HDFC Life Insurance Company Limited



Fig. 2 SBI Life Insurance Company Limited



Fig. 3 ICICI Prudential Life Insurance Company Limited

The stationary data points for ICICI Prudential Life Insurance Company are shown in Fig. 3.

Due to COVID, the data had incredibly high variations from 2019 to 2022. After 2020, the share closing price sharply decreased, and the company’s performance only gradually improved. Currently, the closing price difference is not that large.

The presented f1 score is a development of two simpler evaluation criteria. Before delving into the specifics of the F1 score, let us take a moment to reflect and review the metrics that underpin it. The F1 score is indeed the average of both recall and precision. They both seem to be rated, so using the harmonic mean is a rational choice. Each metric has benefits and drawbacks.

$$F1 = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{Precision} = \text{True Positive} / (\text{True Positive} + \text{False Positive})$$

$$\text{Recall} = \text{True Positive} / (\text{True Positive} + \text{False Negative})$$

True Positive means the number of samples correctly predicted as Positive; True Negative means the number of samples correctly predicted as Negative; False Positive means the number of samples correctly predicted as Positive; False Negative means the number of samples correctly predicted as Negative.

Table 1 F_1 score, RMSE, and MSE score for the insurance companies

No	Company	F_1 Score	Root mean square error (RMSE)	Mean square error (MSE)
1	ICICIPRULI.NS	0.787	3.5204	4.45857
2	SBILIFE.NS	0.836	5.14853	6.89245
3	HDFCLIFE.BO	0.749	3.48140	4.53347

Table 2 Algorithms score

No	Company	Bayesian ridge regression	Support vector machine	Random forest regression	Multinomial Naïve Bayes	XGBoost regression
1	ICICIPRULI.NS	0.99897	0.48788	0.99790	0.52548	0.99587
2	SBILIFE.NS	0.99934	0.48788	0.99883	0.47231	0.99763
3	HDFCLIFE.BO	0.99891	0.47058	0.99794	0.47231	0.99759
4	STARHEALTH.Ns	0.98706	0.42857	0.97687	0.53571	0.95691
5	LICI.NS	0.94419	0.47058	0.97017	0.50000	0.84574

The F_1 score value is more than 0.9 means that the company has a very good reputation. If the value lies between 0.8 and 0.9, it means it is good. If the value lies between 0.5 and 0.8, it means ok and the value less than 0.5 means not good. Here the F_1 score value lies between 0.5 and 0.8. We selected these three firms out of the five because, whereas LIC was just listed in 2022, the other three companies were listed before 2019. For the analysis part, the closing price data value is insufficient. So, when determining the F Score value, only consider ICICI, SBI, and HDFC. According to Table 1, the SBI F score ratio was quite high as compared with all other companies.

Even though there are several machine learning algorithms, just a few were selected for this study. Based on these algorithms and the closing prices of the insurance companies, random forest regression outperforms all others (Table 2). The second one is Bayesian ridge regression, which outperforms all others significantly. SBI Life Insurance Company does exceptionally well overall from the random forest algorithm. As a result, the user has the option of investing in shares as well as buying a policy from SBI Life Insurance.

5 Conclusion

Overfitting occurs when a model performs well on training data but fails to perform well on test data. It has a low basis and a high variance. Underfitting occurs when the model performs poorly on both training and testing data. It has a high basis and a high variance. A generalized model has a low basis and a low variance.

In this case, the variance value for the entire company is less than one. As a result, the risk to the company is very low. These types of company shares are growing slowly if the investor who is afraid of taking risks can purchase these

stocks. Because the risk is low, the return is also low. Risk-seeking investors will not immediately purchase these types of shares because the return is so low and it will take a while for them to get a better return.

Notwithstanding, if the root mean square error and mean square error values are both greater than 1, our model cannot provide a better solution. Naturally, a lower value indicates a lower model error. Moreover, the error term will be high in this case. Random forest performs quite well based on algorithms. The number, i.e., 0.99, is nearer to one. Random forest algorithms, therefore, perform the best matched to the data when compared with the others. SBI Life Insurance Company has good results using the random forest algorithm. Additionally, it is getting closer to one in order for them to invest and buy the policy from that firm. However, due to the stock market's high level of volatility, a person cannot make a judgment solely based on an algorithm. The performance of the organization is impacted by numerous varying aspects. A future trend can be predicted based on the closing price.

Finally, we draw the conclusion that the share price is merely a numerical value, and we use this to accurately predict the future using machine learning algorithms. Using natural language processing, other factors, besides the share price, are taken into account, such as the company's financial health, the state of the market, its credibility, its relationships with customers, and so on. Everything is easily accessible to us. From an investor's perspective, they are prepared to invest and buy shares right away, but they must wait at least a year.

References

1. Nti IK, Adekoya AF, Weyori BA. A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*. 2020;53(4):3007–3057.
2. Nti IK, Adekoya AF, Weyori BA. A comprehensive evaluation of ensemble learning for stock-market prediction. *Journal of Big Data*. 2020;7(1)
3. Nti IK, Adekoya AF, Weyori BA. Efficient Stock-Market Prediction Using Ensemble Support Vector Machine. *Open Comput Sci*. 2020;10:153–163.
4. Nti IK, Adekoya AF, Weyori BA. Predicting Stock Market Price Movement Using Sentiment Analysis: Evidence From Ghana. *Applied Computer Systems*. 2020;25(1):33–42. Available from: <https://doi.org/10.2478/acss-2020-0004>.
5. Masoud, Najeb MH. (2017) "The impact of stock market performance upon economic growth." *International Journal of Economics and Financial Issues* 3 (4): 788–798.
6. M Umer Ghania , M Awaisa and Muhammad Muzammula. "Stock Market Prediction Using Machine Learning(ML)Algorithms", *Computing and Artificial Intelligence Journal Regular Issue*, Vol. 8 N. 4 (2019), 97–116 eISSN: 2255-2863
7. Thomas Fischera, Christopher Kraussb. "Deep learning with long short-term memory networks for financial market predictions", *Standard-Nutzungsbedingungen*, ISSN 1867-6707
8. Xianghui Yuan1, Jin Yuan 1, Tianzhao Jiang2, and Qurat ul Ain, "Integrated Long-Term Stock Selection Models Based on Feature Selection and Machine Learning Algorithms for China Stock Market", *ACCESS*.2020.2969293.
9. Edgar P. Torres P, Myriam Hernández-Álvarez1, Edgar A. Torres Hernández2, and Sang Guun Yoo, "Stock Market Data Prediction Using Machine Learning Techniques", *Springer Nature Switzerland*, https://doi.org/10.1007/978-303011890-7_52

10. Md. Mobin Akhtar, Abu Sarwar Zamani, Shakir Khan, Abdallah Saleh Ali Shatat, Sara Dilshad, Faizan Samdani. "Stock market prediction based on statistical data using machine learning algorithms", Journal of King Saud University-Science Volume 34, Issue 4, June 2022, 101940
11. Nath Bhandari, Binod Rimal, Nawa Raj Pokhrel, Ramchandra Rimal, Keshab R. Dahal, Rajendra K.C. Khatri. "Predicting stock market index using LSTM", [Machine Learning with Applications](#) Volume 9, 15 September 2022, 100320
12. PanelMalti Bansal, Apoorva Goyal, Apoorva Choudhary. "Stock Market Prediction with High Accuracy using Machine Learning Techniques", Procedia Computer Science Volume 215, 2022, <https://doi.org/10.1016/j.procs.2022.12.02>
13. Batta Mahesh." Machine Learning Algorithms - A Review", International Journal of Science and Research (IJSR) ISSN: 2319-7064
14. Osisanwo F.Y.*1, Akinsola J.E.T.*2, Awodele O.*3, Hinmikaiye J. O.*4, Olakanmi O.*5, Akinjobi J. "Supervised Machine Learning Algorithms: Classification and Comparison", International Journal of Computer Trends and Technology (IJCTT), ISSN: 2231-2803

An Ensemble Model of Skin Disease Detection Using CNN and Transfer Learning



Bhagyalaxmi K. , Vemulapally Vennela, N. Tirumal Reddy,
and Shaik Saba Maheen

Abstract Skin diseases are more prevalent than other diseases. Skin diseases can be challenging to diagnose. They could spread fast to other parts of the body if left untreated, leading to serious infection. Skin diseases refer to a range of conditions that affect the skin. These diseases can affect people of every age group and can range from mild, temporary conditions to chronic, life-long conditions. Skin diseases can affect various parts of the skin and can manifest in various forms such as rashes, itching, redness, scaling, blisters, or discoloration. Some skin diseases can be caused by genetic factors, infections, allergies, autoimmune disorders, environmental factors, or lifestyle habits. This paper proposes an ensemble model for various skin diseases detection using convolutional neural networks (CNNs). The proposed model combines three different CNN architectures, namely, Sequential CNN, ResNet-101v2, and DenseNet-121, to improve the accuracy and robustness of skin diseases classification. The proposed model takes skin lesion images as input and produces an average distribution over various skin diseases as output. The model is trained and evaluated on a large-scale skin disease dataset with 10 different skin diseases, achieving state-of-the-art performance in terms of accuracy. The results show that the proposed ensemble model outperforms individual CNN architectures and other state-of-the-art skin disease detection models. The main objective of the project is to create an application that helps in identifying the kind of skin disease easily.

Keywords Ensemble model · CNN · ResNet · DenseNet · Skin disease · Transfer learning

Bhagyalaxmi K. (✉) · V. Vennela · N. Tirumal Reddy · S. S. Maheen
Matrusri Engineering College, Saidabad, Hyderabad, India
e-mail: bhagyalaxmi@matrusri.edu.in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,
Springer Proceedings in Mathematics & Statistics 441,
https://doi.org/10.1007/978-3-031-51338-1_27

357

1 Introduction

Skin diseases are more prevalent than other diseases [1]. Skin diseases can be challenging to diagnose. They could spread fast to other parts of the body if left untreated, leading to serious infection. Skin diseases refer to a variety of conditions that affect the skin. These diseases can affect people of every age group and can range from mild, temporary conditions to chronic, life-long conditions. Skin diseases can affect different parts of the skin and can manifest in various forms such as rashes, itching, redness, scaling, blisters, or discoloration [2]. Some skin diseases can be caused by genetic factors, infections, allergies, autoimmune disorders, environmental factors, or lifestyle habits. Skin diseases can be challenging to diagnose. They could spread fast to other parts of the body if left untreated, leading to serious infection. Skin diseases refer to a range of conditions that affect the skin. Some ignorance among people might lead to skin cancer. To stop the growth and spread of skin diseases, early diagnosis is necessary. Skin disease diagnosis and treatment are more time-consuming and costly. Some common skin diseases include acne, eczema, psoriasis, rosacea, dermatitis, hives, skin cancer, warts, and shingles. Treatment for skin diseases depends on the specific condition and may involve medications, lifestyle changes, topical creams or ointments, light therapy, or surgery. It is important to seek medical attention if you experience any unusual symptoms or changes in your skin. Fungal infections, bacteria, allergies, viruses, and so on can lead to various skin diseases. A skin disease could alter the texture or color of the skin. In the existing approach, the increased skin disease is identified at the later stage using biopsy only [3]. This chapter proposes three stages of learning methods in order to construct an ensemble model with the properties of convolutional neural network (CNN) and transfer learning. The main portion of this study is in extracting features and estimating the distribution of data automatically in a fast and accurate way [4, 5]. According to the experimental findings, the ensemble model successfully classifies images of skin diseases with an accuracy of 96%.

Some of the research works R. Yasir, A. Rahman, and N. Ahmed proposed a method by which they detect various skin diseases by using computer vision techniques. Jainesh Rathod, Vishal wazhrode, and Anirudh sodha used CNN for skin lesions with 90% [6] accuracy. P.B. Manoorkar and Prof D.K. Kamat used the bioimpedance measurement method for the analysis and classification of different skin diseases with 75% accuracy [7]. This chapter is organized as follows: theoretical background is described in Sect. 2, the dataset and the proposed system are given in Sect. 3, the experimental results are given in Sect. 4, and finally, a conclusion is given in Sect. 5.

2 Background

2.1 Convolutional Neural Network

A particular sort of artificial neural network called a CNN is made with the express purpose of processing and analyzing high dimensional input and also including photographs. In computer vision applications including segmentation, object identification, and picture classification, it is frequently employed [8]. The ability of CNNs to automatically recognize spatial hierarchies of features from incoming data is its primary characteristic [9]. Here is a formula for calculating the width/height of one output channel: $n_{out} = \frac{n_{in} + 2p - k}{s} + 1$; n_{in} : width/height of input map; p : convolution padding size; k : convolution kernel width/heights; s : convolution stride size. Typically, CNNs are made up of a various number of convolutional layers that apply a various number of filters to the input image in order to extract local features and then pooling layers that downsample the feature maps. The spatial connection between nearby pixels in a picture is intended to be taken advantage of by CNNs, making them far more effective [10]. In a CNN, there are several layers that perform different operations on the input data [11]. The architecture of the CNN is represented in Fig. 1.

Here are some of the most common layers in a CNN: The first layer is the input layer followed by the convolutional layer: This layer performs the main operation of a CNN, which is convolution. It applies a set of filters to the input image and produces a set of output features that capture various aspects of the input image. Pooling Layer [12]. This helps to reduce the parameters in the model and also makes the model more robust to small variations in the input. Fully Connected Layer: This layer performs a matrix multiplication operation between the input vector and a set of learnable weights, followed by a bias term and an activation function. This produces the final output of the model, which can be used for classification, regression, or other tasks. The classification layer (output layer) is the final layer of the network that predicts the final result. And we can add several others called the Activation Layer, the Batch Normalization Layer, and the Dropout Layer for further feature extraction [13].

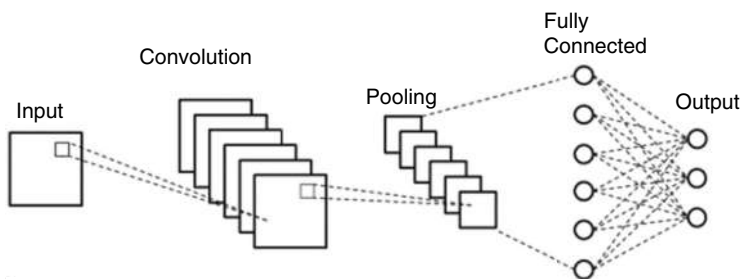


Fig. 1 CNN architecture

2.2 DenseNet

Each layer of the network in DenseNet is connected to every other layer in a feed-forward fashion [14]. Concatenating the features of all earlier layers and using them as input for the subsequent layer is how this is accomplished. Due to the dense connectivity pattern produced, “DenseNet” was coined. As a result, accuracy is increased and overfitting is decreased because of Dense Net’s improved feature reuse and gradient flow.

2.3 Resnet101

He et al. introduced ResNet-101 in 2016; however, ResNet-101v2 is an updated CNN architecture. The ResNet-101v2 design has a number of advantages over the original ResNet-101 architecture, such as batch normalization before activation functions that decrease internal covariate shift and expedite training [15]. ResNet uses the network layers to fit a residual mapping $F(x) + x$, instead of trying to learn the desired underlying mapping $H(x)$ directly with stacked layers.

2.4 Ensemble Model

An ensemble model is a kind of integrated machine learning model that forecasts from various separate models to increase overall forecast accuracy. With ensemble models, you may use the advantages of several different models to balance out the shortcomings of any single model. Building ensemble models can be done in a number of different ways [9]. A popular technique is termed bagging, in which many models are trained individually using various subsets of the training data, and their forecasts are then aggregated by average or voting. Another method is called boosting, in which different models are trained one after the other while each model aims to fix the flaws of the preceding model [16].

As shown in Fig. 2, an ensemble model that combines the sequential CNN model, Densenet121, and Resnet101 is used and the final prediction is done based on the average of the above models.

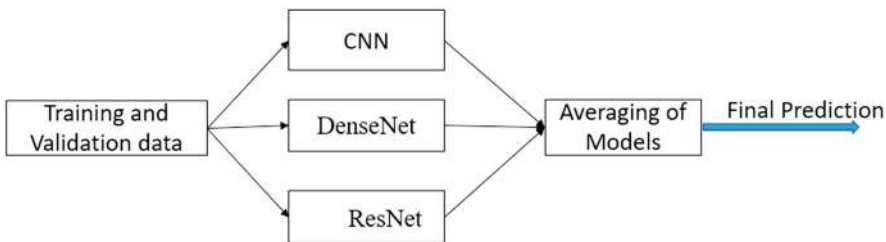


Fig. 2 Ensemble model

3 Materials and Methods

3.1 Skin Disease Dataset

This dataset consists of approximately 27,153 images of 10 different skin disease images. This entire dataset is further split into the training data of 19,358 images, validation data of 5042 images, and testing data of 2753 images. Ten various skin diseases that are used are eczema, warts molluscum and viral infections, melanoma, atopic dermatitis, basal cell carcinoma, melanocytic nevi, benign keratosis-like lesions, psoriasis pictures lichen planus and related diseases, seborrheic keratosis and other benign tumors, tinea ringworm candidiasis, and other fungal infections [2]. The images are converted into 256×256 pixels (Fig. 3).

3.2 Experiments and Setup

In this section, the proposed experimental setup for classification is described. An ensemble model has been developed in Python 3.6.9, due to the availability of the most common machine learning and neural network libraries. In this work, Tensorflow [17] and Keras libraries are used, and they can be run on both the central processing unit (CPU) and graphical processing unit (GPU) [6]. Google Collaboratory (a.k.a. Colab) is a cloud service based on Jupyter Notebooks for demonstrating machine learning education and research. It provides 12.7 GB of RAM, 33 GB of hard disk, and a runtime for deep learning and free-of-charge access to a robust GPU. The GPUs available in Colab often include Nvidia K80s, T4s, P4s, and P100s. In this work, learning models are implemented using Colab services.

3.3 Methodology

This research proposes an ensemble learning method for the classification of skin diseases. As shown in Fig. 2, there are four different models CNN, densenet121,



Fig. 3 Images of skin diseases

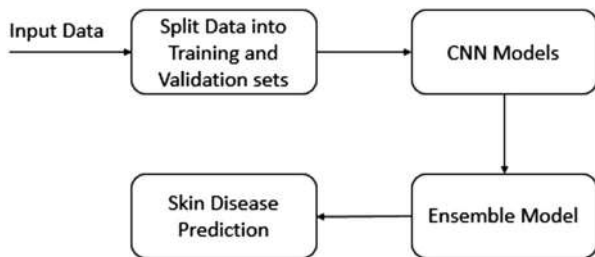


Fig. 4 Technical architecture

resnet101, and an ensemble model with the average of the above three models. CNN networks are highly affected in front of noise and rotation. To prevent this, CNN separately learns features from the raw image. The complementary behaviors of input images exploit better features for classification [18]. Finally, the results of CNNs are shared in a fully connected network to demonstrate the ability to classify skin diseases. The technical architecture of the model is given in Fig. 4.

This network consists of three layers including the input layer which is extracted features of CNN, one hidden layer, and an output layer with the number of classes in the dataset. The activation functions for the hidden layer and output layer are ReLU and SoftMax, respectively. All layers, i.e., those in the same dense block and transition layer, share their weight among multiple inputs, allowing deeper layers to use early extracted features. ResNet-101 is a CNN with 101 layers deep. Input data is split into training, validation, and testing datasets, then these data split with a ratio 70:20:10 split is trained with various models, then the models are combined forming an ensemble model which is a combination average of all the three models named CNN, DenseNet, and ResNet, and then final prediction of the disease is done by using ensemble model through its image.

4 Evaluation Results

All input skin disease images were resized into 256×256 pixels before feeding them into the ensemble model. Some of the skin disease images are shown in Fig. 3 which are being used in all the models used in this chapter. CNN contains four convolutional layers, two max pool layers, one flatten layer, and one dropout layer of 0.5 [5], with a kernel size of 3 and 10 filters by using padding and ReLU activation function, whereas the final output layer uses SoftMax as activation function. The ensemble model is described in Fig. 5.

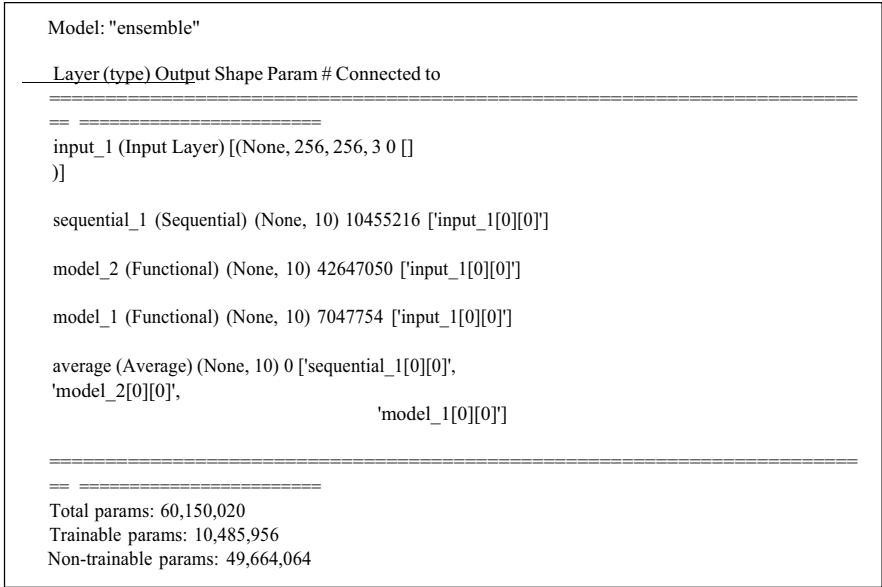


Fig. 5 Ensemble model

4.1 Comparison with Related Work in Terms of Accuracy

This CNN model predicts with 95% accuracy for 20 epochs. Densenet121 has 120 convolutions and 4 average pools and is a pre-trained model predicted with 75% accuracy for 15 epochs. Resnet101 is a pre-trained model with 101 deep layers predicted with 85% accuracy for 10 epochs. Finally, an ensemble model that is created using the above three models CNN, DensNet121, and ResNet101 is a pre-trained model with 101 deep layers predicted with 85% accuracy for 10 epochs. Finally, an ensemble model is created using the above three models, namely, CNN, DensNet121, and ResNet101. The final prediction of an ensemble model is made using an average of these three models, and this model predicts with an accuracy of 96% for 5 epochs as given in Fig. 6. In Table 1, the comparison of various models used in this chapter with their accuracies is shown.

5 Conclusion

In this chapter, a general model of skin disease images and a CNN were made, which not only removes unwanted features and soothes skin disease boundaries but also extracts useful features for image classification [5]. The evaluation results demonstrate that most configurations outperformed individual models, but there

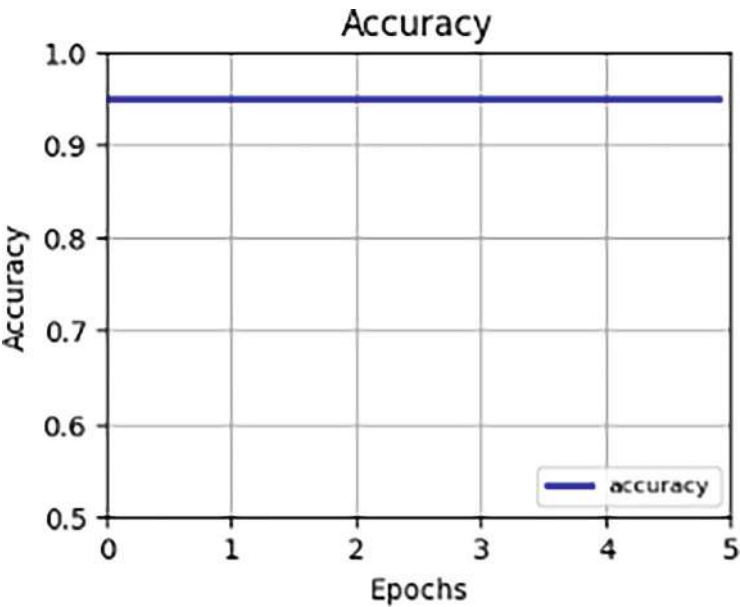


Fig. 6 Accuracy graph of ensemble model. Accuracy = (TP+TN)/(TP +TN + FP + FN), where *TP* = True Positives, *TN* = True Negatives, *FP* = False Positives, and *FN* = False Negatives

Table 1 Comparison of different models

Model	Accuracy
Ensemble	96%
CNN	95%
DenseNet	75%
ResNet	85%

were also cases where individual models outperformed some configurations. The CNN model predicts with 95% accuracy for 20 epochs. DenseNet121 has 120 convolutions and 4 Avgpool and is a pre-trained model predicted with 75% accuracy for 15 epochs. ResNet101 is a pre-trained model with 101 deep layers predicted with 85% accuracy for 10 epochs. This emphasizes that combining different architectures is also an important factor, along with the ensemble model of buildings with other factors such as network depth, size, and total number of parameters. In the evaluation results, the best model was an ensemble model (DenseNet121, CNN, and ResNet) with 96% accuracy for 5 epochs. This research results in an entity model that can be used in medical imaging applications. This project can be further improved in the future by recommending precautions and medical treatments for the skin disease that is being diagnosed.

References

1. Santy, A., & Joseph, R. (2015) "Segmentation Methods for Computer Aided Melanoma Detection." Global Conference of Communication Technologies
2. R. Yasir, M. A. Rahman and N. Ahmed, Dermatological disease detection using image processing and artificial neural network.
3. Flow T. Nidhal K, Abbadi A, Saadi N.D, Muhsin A, AL-Dhalimi and Restom H. Psoriasis Detection Using Skin Color and Texture Features. Journal of Computer Science.
4. R S Goud, Priyanka S Gadre, Jyoti B Gaikwad and Priyanka K Wagh, "Skin Disease Diagnosis System using Image Processing and Data Mining,
5. Krizhevsky, A., ILYA, S., & Geoffrey, E. (2012) "ImageNet Classification With the Deep Convolutional Neural Networks." Advances in Neural Information Processing Systems
6. Suneel Kumar and Ajit Singh. Image processing for recognition of skin diseases. International Journal of Computer Applications, 149(3):37–40, 2016
7. Alexander Wang, Jacob Scharcanski, and Paul Fieguth. Automatic skin lesion segmentation via iterative stochastic region merging.
8. Sigurdur Sigurdsson, Peter Alshede Philipsen, Lars Kai Hansen, Monika Gniadecka, and Hans-Christian Wulf. Detection of skin cancer by classification of raman spectrum
9. YuanYo, Lo YC. Improving Dermoscopic Image Segmentation with Enhanced Convolutional-Deconvolutional Networks.
10. Hyun Chung and Guillermo Sapiro. Segmenting skin lesions with partial-differential equations- based image processing algorithms. IEEE transactions on Medical Imaging.
11. Andrew G.H, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobile Nets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Google Inc., arXiv:1704.04861v1 [cs.CV]. April 2017.
12. Anitha G.K, Deepak M.C. Machine Learning Techniques for learning features of any kind of data: A Case Study. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 2014 December;3(12).
13. Li Y, Shen L. Skin Lesion Analysis towards Melanoma Detection Using Deep Learning.
14. Kassem MA, Hosny KM, Fouad MM. Skin Lesions Classification into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning, in IEEE Access. 2020;8:114822–114832. DOI:<https://doi.org/10.1109/ACCESS.2020.3003890>.
15. Patil R, Bellary S. Transfer learning based system for melanoma type detection. Revue d'Intelligence Artificielle. 2021;35:(2):123–130. <https://doi.org/10.18280/ria.350203>
16. Patil R, Bellary S. Transfer learning based system for melanoma type detection. Revue d'Intelligence Artificielle. 2021;35:(2):123–130. <https://doi.org/10.18280/ria.350203>
17. M Emre Celebi, Quan Wen, Sae Hwang, Hitoshi Iyatomi, and Gerald Schaefer. Lesion border detection in dermoscopy images using ensembles of thresholding methods. Skin Research and Technology, 19(1):e252–e258, 2013.
18. ImageRecognitionTensorFlow, Available: https://www.tensorflow.org/tutorials/image_recognition. Accessed 2018 March; 28.
19. R Sumithra, Mahamad Suhil, and DS Guru. Segmentation and classification of skin lesions for disease diagnosis. Proceeds Computer Science, 45:76–85, 2015.

Session-Based News Recommendation System



V. Vemani, Vaibhav Chemboli, and Pusarla Sindhu

Abstract In today's digital age, the vast amount of available content makes it increasingly challenging for users to find relevant news articles that match their interests. News recommendation algorithms play a pivotal role in alleviating this problem by providing personalized recommendations based on user preferences. Session-based recommendation systems have gained significant attention in recent years due to their ability to capture user preferences and behavior in a more granular and dynamic manner. Unlike conventional recommendation systems that rely on historical data or long-term preferences, session-based recommendation systems prioritize the user's most recent interactions within a session. This enables instant identification of user preferences and facilitates the delivery of relevant recommendations. In this research paper, the authors explore the challenges faced by news recommendation algorithms, which play a crucial role in helping users discover relevant news articles in the era of information overload, and discuss the temporal structure of news content, the diversity of user tastes, the cold start problem faced by new users, and the significance of transformer models in the natural sciences.

To address these challenges, the authors propose the utilization of session-based recommendation systems and investigate various approaches for constituting session-based recommendation systems, including the use of deep learning models, collaborative filtering, and hybrid techniques. Additionally, the authors highlighted the significance of the XLNet architecture combined with causal language modeling for accurate predictions within the session-based recommender system implementation.

Keywords Session-based recommendation · Sequential recommendation · Causal language modeling · Transformers · XLNet · Next item prediction

V. Vemani · V. Chemboli · P. Sindhu (✉)

Department of CSE, GITAM School of Technology, GITAM Deemed to be University,
Visakhapatnam, India

e-mail: spusarla@gitam.edu

1 Introduction

With the abundance of news articles being released every day, it can be difficult for users to sort through them all to find the ones that are applicable to their interests. When a user's past behavior and preferences are considered, news-based recommendation systems can make suggestions for items that are fair to be of interest to them.

The transitory nature of news content makes developing recommendation systems based on it one of the main challenges. The relevancy of a news article can swiftly wane over time because they are continually being released. This means that, to continue to be effective, recommendation systems must be able to quickly adapt to shifting trends and user preferences. The wide range of user preferences is another difficulty. While some individuals might be more excited about sports or entertainment, others might be more fascinated by politics and current affairs. To give individualized recommendations that are pertinent to each user, recommendation systems must be able to consider these individual characteristics. The cold-start issue is also a difficulty for news recommendation algorithms, to sum up. When a new user registers for the service and has never interacted with the system before, this happens. In this situation, the system must rely on additional information sources, such as human input or demographic information, in order to deliver pertinent recommendations.

A session-based recommendation system focuses on logging user preferences and activity during a specific session. In contrast to traditional recommendation systems that rely on historical data or long-term preferences, session-based systems are designed to respond to rapidly changing user interests and preferences. Session-based recommendation systems have gained popularity recently because of their ability to provide real-time, personalized recommendations based on a user's current interests. Therefore, they are particularly well-suited for applications such as news portals and e-commerce.

Users will receive appropriate news articles via session-based news recommendation algorithms based on their recent activity and personal preferences. These programs stand out from traditional news recommendation services, which usually rely on detailed customer preferences and histories.

One of the important advantages of session-based news recommender systems is their ability to recognize the individual's current interests and provide real-time recommendations. In the realm of news, this is essential because an article's relevance can quickly dwindle with time. These systems focus on short-term user behavior within a certain session, making them more able to adapt to changing trends and provide useful recommendations.

2 Literature Review

Hidasi and Karatzoglou introduced a session-based recommendation approach using recurrent neural networks (RNNs) with top-k gains and a novel loss function that optimizes the model to predict the next item in a session. Experimental results show improved performance compared with traditional RNN-based methods [1]. Song et al. proposed a cluster-based meta-learning (CBML) model for session-based recommendation. CBML leverages user clustering to transfer knowledge across similar sessions and improves recommendation accuracy. Experimental results demonstrate the effectiveness of the proposed approach compared with existing methods [2]. Mikolov et al. introduced the concept of word embeddings and their compositionality using distributed representations, the Word2Vec model, which learns continuous representations of words based on their co-occurrence patterns. This work has had a significant impact on natural language processing and recommendation systems [3]. Grbovic et al. focused on e-commerce recommendations and proposed a scalable approach for generating product recommendations. Their method utilizes item co-purchasing information to build a recommendation model that is effective at scale. Experimental results demonstrate the practicality and efficiency of the proposed approach [4]. Vasile et al. introduced Metaprod2vec, a model that incorporates side information about products to learn embeddings for recommendation. By considering product attributes and metadata, Metaprod2vec improves the quality of product representations. The paper described the effectiveness of the approach through experiments on real-world datasets [5]. Li et al. proposed a neural attentive session-based recommendation model. Their approach employs attention mechanisms to capture the relevance between items in a session. Experimental results demonstrate the effectiveness of the proposed model in predicting the next item accurately [6]. Vaswani et al. introduced the transformer model, which revolutionized natural language processing tasks by utilizing self-attention mechanisms. The transformer architecture has been widely adopted in various domains and has significantly improved the performance of recommendation systems [7].

Yang et al. presented XLNet, a generalized autoregressive pretraining approach for language understanding. XLNet overcomes the limitations of previous models by considering all possible permutations of the input sequence during training, enabling bidirectional context modeling. The model achieves state-of-the-art results on several benchmark tasks [8]. De Souza Pereira Moreira et al. proposed Transformers4rec, a framework that combines natural language processing techniques with sequential/session-based recommendation. The model incorporates pre-trained transformer-based language models to enhance the representation learning for recommendation tasks. Experimental results demonstrate improved performance compared with traditional approaches [9]. Guo et al. proposed a joint neural network model for session-aware recommendation. Their approach combines a session representation module with an item representation module to capture the dynamics of user sessions. Experimental results show the effectiveness of the proposed

model in capturing user preferences and generating accurate recommendations [10]. Chenglong Wang et al. proposed an improved XLNet text classification model based on the problems of long-term dependence and insufficient contextual semantic expression in previous pre-trained language models. The authors used XLNet pre-trained language modeling to represent text as low-dimensional word vectors to obtain sequences [11].

For session-based recommendation systems, cutting-edge pre-training methods for NLP tasks like XLNet have shown promising results. XLNet's pre-training challenge is training the model to anticipate a randomly masked word in a phrase, unlike earlier models like BERT. To avoid pushing the model to anticipate masked words based on their position in the input sequence, XLNet, unlike those earlier models, considers all potential permutations of the input sequence. As a result, XLNet [12, 13] can handle long-term dependencies and preserve more intricate interactions between input tokens.

XLNet employs causal language modeling (CLM), a state-of-the-art approach to language modeling. CLM teaches a model to recommend the subsequent word in a sequence given all preceding words, like classical language modeling. Yet unlike traditional language models that use left-to-right or right-to-left prediction, XLNet uses a bidirectional approach that enables the model to take into consideration all probable contexts for each word in the sequence.

The key innovation of XLNet is the adoption of a permutation-based training approach, which enables the model to learn from all feasible permutations of the input sequence. The model is trained to predict each word in the sequence based on all possible combinations of the other words in the sequence.

3 Architecture

The author used Transformers4Rec's architecture [14] to accomplish the task of session-based new recommendations. Transformers4Rec's fundamental concept is to leverage the transformer's self-attention mechanism to record both the item-level and sequence-level dependencies of user interactions. To accomplish this, the session is represented as a series of objects, and each item is encoded using an embedding layer. The attention weights for each item in the sequence in relation to others are analyzed by the multi-head self-attention layer.

Tokenizer, transformer architecture, and a head are the three main parts of models for diverse NLP tasks. However, in the presented work, the authors used transformer architecture [15] and its configuration classes, combining unique heads created especially for the recommendation problem depicted in Fig. 1.

To create the interaction embedding, the Feature Processing module normalizes and adds input embeddings that can be either sparse category or continuous in nature. Before being given to the Sequence Processing module, this sequence of embeddings is first masked using the CLM-based Sequence Masking module. This module is made up of stacked transformer blocks with an alterable block count and

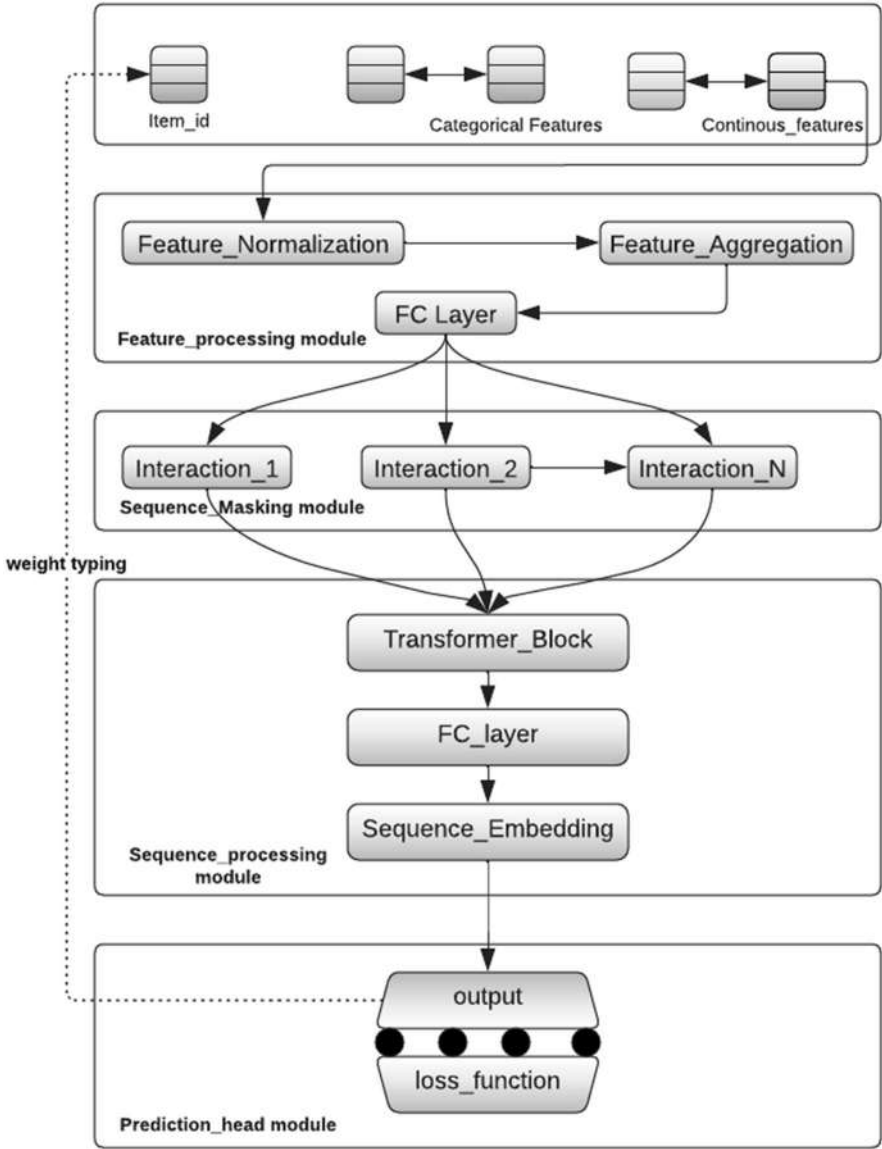


Fig. 1 Transformers4Rec meta-architecture

an XLNet architecture type. Each place in the sequence is given a vector that is projected to represent an embedding of the sequence. Following that, the prediction module is set up for various functions, such as item prediction or sequence-level predictions. The authors utilized the item’s prediction head with an output layer that uses the weight-tying embedding technique to tie the resultant layer to the item

matrix weights. The resultant embedding is passed through the softmax layer to generate scores for each item including the loss function called cross-entropy loss.

3.1 Feature Processing

The technique of integrating various input features into a single vector representation is known as feature aggregation. The design allows for the normalization and aggregation of numerous interaction-level features in a variety of ways utilizing the two types of aggregation functions concatenation and element-wise. In contrast to the element-wise merge, which combines the additional features by the item embedding after first summing up the additional features element-wise, the item id feature is simply concatenated with the other input features in the concatenation merge. The system also allows for the representation of continuous numerical features using a Soft One-Hot Encoding. Concatenation merging was employed in this work to aggregate the input.

3.2 Masking

Masking is a method used in machine learning models to deliberately hide or mask some portions of the input data while they are being trained. Masking's goal is to have the model develop more robust and broad representations by preventing it from depending too much on particular input components. One kind of masking used in language models, when the models are employed for language-generating tasks, is called CLM masking. With CLM masking, part of the words in the input sentence are randomly masked off, so the model must guess them based on the context. The architecture is trained to recommend the next word in a sentence given the preceding words.

3.3 Sequence Processing

The transformer is employed by the XLNet model for sequence processing. In contrast to traditional left-to-right or right-to-left masking, XLNet represents dependencies between all locations in the sequence using a permutation-based method. To do this, auto-regressive permutation, which trains the model to predict the tokens depending on the input sequence's permutation selected at random, is utilized.

Although not to the same extent as other transformer models, XLNet employs masking. Typical transformer models mask the input in a forward or backward direction, in contrast to XLNet, which is trained to predict the token at a particular

place based on all other positions in the sequence, regardless of their position in relation to the present position. A segment-level recurrence mechanism that allows the model to maintain an explicit recollection of the previous segments that it has processed makes this possible.

Relative positional encoding, which is one of the methods employed by XLNet to improve sequence processing in addition to the permutation-based method and segment-level recurrence mechanism, encodes the relative location of tokens as a vector and adds it to the usual positional encoding of each token. As a result, the model can more precisely depict the connections among tokens at different places in the sequence.

3.4 Prediction Head

To generate recommendations, the Transformers4Rec system uses a prediction head, which is a neural network layer put on top of the transformer block. This head receives the sequence of interaction embeddings generated by the transformer block and generates a prediction for an exclusive item in the recommendation list.

The next item prediction head is used for task-based recommendation tasks that are session-based. Using the transformer block's output as input, it applies the softmax function to generate scores for the entire list of suggestions. After that, the user is advised to go on to the item with the highest chance.

4 Methodology

The experiment's training and evaluation methodologies, metrics, and hyperparameter tuning process are all described in this section. Additionally, it gives an overview of the preprocessing procedures employed on the datasets used in the study.

4.1 Dataset

Due to the anonymous browsing habits of news consumers and the availability of useful data regarding recent user interactions, the G1 news dataset [16] was chosen for session-based recommendation studies. The dataset comprises textual representations and details about the news stories read, as well as sampling individual's sessions and interactions from the news portal over a 16-day time period. Pre-processing included adding data frames to hourly indices, clustering interactions into sessions, removing sessions with fewer than 100 interactions, and adding temporal characteristics to uniquely identify timestamps. The final dataset was made up of 2 days' worth of data, split 80/20 between train and test.

Table 1 Hyperparameters

Hyperparameters	Values
Learning rate (lr)	0.000666
Train batch size	384
Maximum sequence length	124
Epochs	10
Gradient sequence length	1
Aggregation	concat
Masking	CLM

All the features’ metadata, including their minimum and maximum values (for continuous features), data types, and whether they are continuous, are contained in the schema that is utilized for feature aggregation. As previously mentioned, the authors employed the concatenation and merge aggregation types and selected the masking type as causal language modeling.

4.2 Hyperparameters

First, the model architecture and hyperparameters were carefully selected, including, among other things, the batch size, learning rate, size of the embedding and hidden layers, and the number of layers in the transformer block presented in Table 1. The model was trained on the training set once the model architecture and hyperparameters were chosen. The loss function used is called cross entropy.

The model’s performance on the validation set was assessed after each training session in order to spot any overfitting issues and adjust the hyperparameters. According to the recommendation task, various evaluation metrics, including recall and normalized discounted cumulative gain, were employed to measure the model’s effectiveness.

5 Results and Discussion

NDCG, which stands for normalized discounted cumulative gain, measures how good the ordering of the recommended result is, and it is a valid measure for recommendation systems [17]. It considers both the recommended items’ relevance and where they fall in the list of recommendations. The NDCG formula is as follows:

$$NDCG@k = DCG@k / IDC@k$$

where $DCG@k$ is the discounted cumulative gain that weighs each relevancy score for its respective position at position k , $IDC@k$ is the ideal discounted cumulative

gain refers to the expected cumulative gain at position k , and k refers to the length of the recommendation list. By ranking the ground truth items according to their relevance score and computing the DCG at position k , the optimal DCG may be determined.

On the contrary, recall quantifies the fraction of relevant items that are suggested in the top- k list.

$$\text{recall}@k = \frac{\text{(number of relevant items that are recommended)}}{\text{(total number of applicable items)}}$$

where the relevant items are those that the user has already dealt with, and those are well-regarded by the recommendation job.

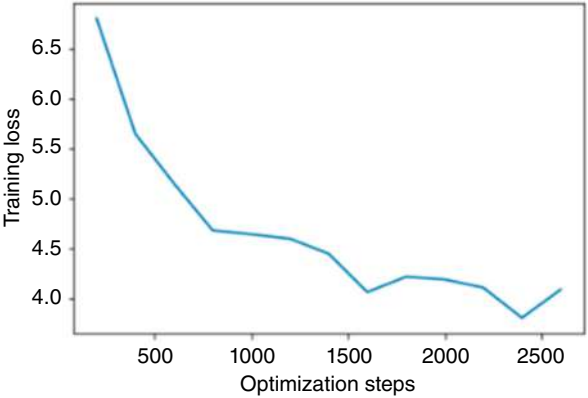
The model’s performance represented in Table 2 has shown significant improvement with an NDCG@10 score of 0.14 and an NDCG@20 score of 0.16. In addition, the recall score has also improved considerably, with a score of 0.26 for recall@10 and 0.34 for recall@20.

From Fig. 2, it is observed that the loss of the model is decreasing when compared with the optimization steps. Even though news recommendation typically involves a lower relevancy of recommended items, these scores and learning rates demonstrate that the model employed can still recommend a substantial proportion of relevant items, as indicated by the recall score.

Table 2 NDCG and recall values at the 10th and 20th epochs

Metrics	Value
NDCG@10	0.14
NDCG@20	0.16
Recall@10	0.26
Recall@20	0.34

Fig. 2 Training loss versus optimization steps



6 Conclusion

The proposed XLNet-based session-based recommendation system demonstrated strong predictive capabilities for determining subsequent user interactions. This study highlights the advantages of using XLNet in session-based recommendation systems and identifies promising areas for future research. Further investigations can explore alternative language models or architectural modifications to optimize system performance. Integrating additional data, such as user demographics or item attributes, could enhance the accuracy of the proposed model.

References

1. Hidasi, B., & Karatzoglou, A. (2018, October). Recurrent neural networks with top-k gains for session-based recommendations. In *Proceedings of the 27th ACM international conference on information and knowledge management* (pp. 843–852).
2. Song, J., Xu, J., Zhou, R., Chen, L., Li, J., & Liu, C. (2021, October). CBML: A cluster-based meta-learning model for session-based recommendation. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 1713–1722).
3. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.
4. Grbovic, M., Radosavljevic, V., Djuric, N., Bhamidipati, N., Savla, J., Bhagwan, V., & Sharp, D. (2015, August). E-commerce in your inbox: Product recommendations at scale. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1809–1818).
5. Vasile, F., Smirnova, E., & Conneau, A. (2016, September). Metaprod2vec: Product embeddings using side-information for recommendation. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 225–232).
6. Li, J., Ren, P., Chen, Z., Ren, Z., Lian, T., & Ma, J. (2017, November). Neural attentive session-based recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 1419–1428).
7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
8. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32.
9. de Souza Pereira Moreira, G., Rabhi, S., Lee, J. M., Ak, R., & Oldridge, E. (2021, September). Transformers4rec: Bridging the gap between nlp and sequential/session-based recommendation. In *Proceedings of the 15th ACM Conference on Recommender Systems* (pp. 143–153).
10. Guo, Y., Zhang, D., Ling, Y., & Chen, H. (2020). A joint neural network for session-aware recommendation. *IEEE Access*, 8, 74205–74215.
11. Chenglong Wang, Fenglei Zhang, “The performance of improved XLNet on text classification”. In *Third International Conference on Artificial Intelligence and Electromechanical Automation (AIEA 2022)*, 1232900 (7 September 2022); <https://doi.org/10.1117/12.2646785>.
12. Suleiman Khan, 2019, Available: <https://towardsdatascience.com/bert-roberta-distilbert-xl-net-which-one-to-use-3d5ab82ba5f8>.
13. Xu Liang, 2019, Available: <https://towardsdatascience.com/what-is-xl-net-and-why-it-outperforms-bert-8d8fce710335>.

14. Nvidia, GitHub, Available: https://nvidia-merlin.github.io/Transformers4Rec/main/model_definition.html
15. Transformers4Rec, Model Architecture, 2021, Available: https://nvidia-merlin.github.io/Transformers4Rec/stable/model_definition.html
16. Gabriel Moreira, Kaggle, 2019, Available: <https://www.kaggle.com/datasets/gspmoreira/news-portal-user-interactions-by-globocom>
17. Pranay Chandekar, Medium, 2020, Available: <https://towardsdatascience.com/evaluate-your-recommendation-engine-using-ndcg-759a851452d1>

A Fusion-Based Approach for Generating Image Captions



Samatha J.  and G. Madhavi

Abstract Image captioning is a demanding task that involves generating textual descriptions of images. In recent years, deep learning approaches, such as convolution neural networks (CNNs) and long short-term memory (LSTM) networks, have been extensively used to address this task. In this approach, the CNN is used to extract features from the input image, while the LSTM is used to produce a succession of words that describe the image. The CNN-LSTM architecture has been revealed to achieve modern performance on various image captioning datasets. In this chapter, we present a review of the current progress in image captioning using CNN and LSTM, including the different architectures and techniques used in the literature. Adding the capability to read aloud the image subtitle has increased the sophistication of image captioning in comparison to other works in the field.

Keywords Image captioning · Deep learning · CNN · LSTM · Feature extraction

1 Introduction

Image captioning refers to the procedure of generating a textual description of an image by considering the properties of objects, actions, and interactions depicted within the image. This process involves utilizing an image caption generator, which takes an image as input and produces a descriptive caption as output. To accomplish image captioning, the computer performs two primary tasks [1, 2]. The first task involves comprehending and interpreting the image using computer vision techniques. The second task entails generating captions in natural language based on

Samatha J. (✉)

Research Scholar JNTUK, Assistant Professor, Matrusri Engineering College, Saidabad, Hyderabad, Telangana, India

G. Madhavi

Assistant Professor, University College of Engineering, JNTUK, Kakinada, Andhra Pradesh, India

the features identified in the image. With the progression in deep learning techniques for computer vision, the availability of large-scale datasets, and powerful computing resources, it is now feasible to generate image captions with increasingly accurate results.

The process of generating textual descriptions for images is important and necessary for the following reasons [4, 6]:

- **Accessibility:** Image captioning enhances the accessibility of visual content to people with visual impairments or other disabilities. By providing a textual description of an image, individuals who are unable to see the image can still understand the content [9].
- **Enhanced Communication:** Image captioning can facilitate communication among individuals who have diverse linguistic backgrounds [13]. A textual description can serve as a common ground for individuals to discuss and understand the content of an image.
- **Information Retrieval:** Image captioning can improve the precision and effectiveness of image-based search engines by enabling users to search for images based on their textual descriptions [10]. This can be particularly helpful for finding particular images among large collections of images.
- **Automation:** Image captioning can aid in the development of automated systems that need to understand and interpret visual content, such as autonomous vehicles or robots. By providing a textual description of an image, these systems can enhance their understanding of the environment and carry out tasks with improved effectiveness [17].
- **Creativity:** Image captioning can also serve creative endeavors, such as generating captions for images in art or photography. This can enhance the overall aesthetic and artistic value of an image.

Image captioning finds extensive applications across diverse fields [8, 15]. Here are some examples:

- **Visual Impairment:** Image captioning has the potential to be utilized for textual descriptions of images to individuals who are visually impaired. This can aid them in gaining a deeper comprehension of visual content.
- **Social Media:** Image captioning can be utilized to enhance the user experience on social media platforms by providing automated descriptions of images uploaded by users. This can improve accessibility and engagement on social media.
- **Image Search:** Image captioning can be used to enhance the precision of image search engines by allowing users to search for images based on their textual descriptions. It has the potential for finding specific images or identifying similar images.
- **Advertising:** Image captioning can be utilized for automatically generating captions for product images in advertising. This can save time and resources while also improving the precision of the captions.

- **Robotics:** Image captioning can support the development of autonomous robots that need to understand and interpret visual content so as to perform tasks. Robots involved in tasks such as cleaning or caregiving can leverage the advantages of image captioning.
- **Education:** Image captioning can be utilized in educational materials to provide textual descriptions of visual content. This can aid students with learning disabilities or individuals who have limited proficiency in the language of instruction in comprehending the content more effectively [14].
- **Entertainment:** Image captioning can be employed to produce humorous or creative captions for images in entertainment media such as comics, memes, and cartoons.

2 Related Work

Many researchers started working on object recognition in images, but only providing the names of objects recognized does not make such a good impression as a full, human-like description. Natural language descriptions will remain a challenge to be solved if the machines do not think, talk, and behave like humans. Different techniques were proposed by the researchers to make the computer understand the image and generate sentences regarding it. Below are several examples of existing systems in this domain:

In their work, Hoxha et al. [3] presented a framework for captioning remote sensing images utilizing beam search. Their approach generates multiple captions for the sensed data and selects the most prominent caption by considering the lexical similarity with the most analogous images.

For image captioning, Srushti Shinde et al. [5] explored various deep learning algorithms, including convolution neural network (CNN), VGG16, RNN, and ResNet. For the analysis, the Flickr_8K dataset was used. In comparison to other models, ResNet 101 among them fared well.

Gupta et al. [7] employed deep learning and natural language processing techniques to develop an image caption generator specifically designed for visually impaired individuals. The approach involved training the model using transfer learning on the Flickr_8K dataset to generate captions for the test data.

In their study of several image captioning techniques, Mohammad Shahnawaz Alam et al. [12] compared their findings with more current research. Several system applications and use cases were also explored.

Vaibhav Pandit et al. [16] developed a DeepCap model for image captioning on black-and-white images using deep learning algorithms, which achieved an accuracy of 45.77% on validation data.

3 Methodology

The image caption generator utilizes a combination of CNN and long short-term memory (LSTM) in order to generate captions for images [10, 18]. Specifically, the generator employs CNN to extract features from the images and LSTM to generate descriptive sentences. By incorporating LSTM, the generator overcomes the limitations of the recurrent neural network [11], enabling it to handle long sentences effectively. The hybrid architecture of CNN and LSTM, as depicted in Fig. 1, is employed for the purpose of caption generation in this system.

The fusion-based architecture of CNN-LSTM is depicted in Fig. 1, showcasing the overall design of the system. The system takes an image of dimensions $224 \times 224 \times 3$ as input. The CNN component extracts the image’s features, resulting in a 2048-dimensional feature vector that serves as the input for the LSTM. The LSTM processes the input received from the previous layer across its various layers, progressively computing the output. The number of LSTM layers depends on the maximum length of the sentence to be generated. To distinguish the beginning and end of a sentence, the LSTM utilizes special tokens such as “<start>” and “<end>.”

In this chapter, the Flickr_30K dataset is used for caption generation. Among the 30,000 images, 22,000 are trained, and the remaining are used for testing and validation purposes. When comparing the accuracy between the Flickr_30K dataset and the Flickr_8K dataset, the results indicate that the Flickr_30K dataset provides higher levels of accuracy. Figure 2 exhibits a sample Flickr_30K text dataset.

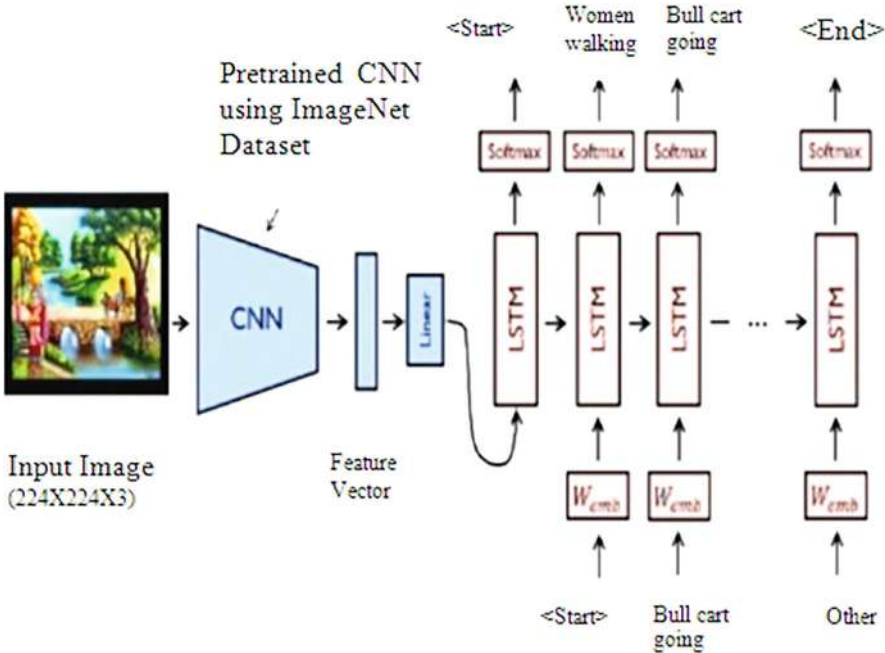


Fig. 1 Fusion-based architecture of CNN-LSTM


```

1003163366_44323f5815.jpg,a man sleeping on a bench outside with a white and black dog sitting next to him .
1003163366_44323f5815.jpg,A shirtless man lies on a park bench with his dog .
1003163366_44323f5815.jpg,man laying on bench holding leash of dog sitting on ground
1007129816_e794419615.jpg,A man in an orange hat starring at something .
1007129816_e794419615.jpg,A man wears an orange hat and glasses .
1007129816_e794419615.jpg,A man with gauges and glasses is wearing a Blitz hat .
1007129816_e794419615.jpg,A man with glasses is wearing a beer can crocheted hat .
1007129816_e794419615.jpg,The man with pierced ears is wearing glasses and an orange hat .
1007320043_627395c3d8.jpg,A child playing on a rope net .
1007320043_627395c3d8.jpg,A little girl climbing on red roping .
1007320043_627395c3d8.jpg,A little girl in pink climbs a rope bridge at the park .
1007320043_627395c3d8.jpg,A small child grips onto the red ropes at the playground .
1007320043_627395c3d8.jpg,The small child climbs on a red ropes on a playground .
1009434119_febe49276a.jpg,A black and white dog is running in a grassy garden surrounded by a white fence .
1009434119_febe49276a.jpg,A black and white dog is running through the grass .
1009434119_febe49276a.jpg,A Boston Terrier is running in the grass .
1009434119_febe49276a.jpg,A Boston Terrier is running on lush green grass in front of a white fence .
1009434119_febe49276a.jpg,A dog runs on the green grass near a wooden fence .
1012212859_01547e3f17.jpg,"A dog shakes its head near the shore , a red ball next to it ."
1012212859_01547e3f17.jpg,A white dog shakes on the edge of a beach with an orange ball .
1012212859_01547e3f17.jpg,"Dog with orange ball at feet , stands on shore shaking off water"
1012212859_01547e3f17.jpg,white dog playing with a red ball on the shore near the water .

```

Fig. 2 Sample text dataset

The steps involved in the image caption generator can be described as follows:
 Step 1: Import all the required packages.

Step 2: The following steps are undertaken for data cleaning on the Flickr_30K dataset:

- Loading the flickr_30k file and extracting the contents into a string.
- Creating a description dictionary that associates images with a list of five captions.
- Cleaning the descriptions by removing punctuation, converting the text to lowercase, and eliminating words containing numbers.
- Extracting all unique words and constructing the vocabulary based on the descriptions.
- Generating a list of preprocessed descriptions and saving them in a file.

Step 3: Feature vector extraction from images.

This step involves utilizing transfer learning, a technique that leverages pre-trained models trained on extensive datasets, to extract features for our task. Specifically, we employ the Xception model, which has undergone training on the ImageNet dataset containing 1000 distinct classes for classification purposes. The Xception model can be readily imported from the Keras applications library. However, as the Xception model was originally designed for ImageNet, some modifications are necessary for seamless integration with our specific model. We remove the final classification layer to obtain a 2048-dimensional feature vector. It is important to note that the Xception model expects images of size $299 \times 299 \times 3$ as input.

Step 4: Dataset loading for model training.

In this step, we perform the following tasks to load the dataset for training the model:

Table 1 Caption generator

	Xi		Yi
i	Image feature vector	Partial caption	Target word
1	Image_1	startseq	the
2	Image_1	startseq the	black
3	Image_1	startseq the black	cat
4	Image_1	startseq the black cat	sat
5	Image_1	startseq the black cat sat	on
6	Image_1	startseq the black cat sat on	grass
7	Image_1	startseq the black cat sat on grass	endseq

- Developing a function responsible for loading the text file and returning the list of image names as a string.
- Constructing a dictionary that associates each photo in the list with its corresponding captions. To facilitate the LSTM model in identifying the beginning and end of each caption, we append the <start> and <end> identifiers.
- Creating a dictionary that pairs image names with their respective feature vectors. These feature vectors have already been extracted using the Xception model in the previous step.

Step 5: Vocabulary tokenization.
To enable the computer to comprehend English words, we need to represent them using numerical values. This process involves the following steps:

- Associating each word in the vocabulary with a distinct index value. To accomplish this, we can employ the tokenizer function available in the Keras library, which allows us to create tokens based on the vocabulary.
- Saving the tokens corresponding to the vocabulary in a separate file to preserve this mapping for future use.

Step 6: Data generator creation.
In this step, we create a data generator to facilitate the supervised learning task. The following tasks are performed:

To train the model effectively, we need to provide input–output pairs. For training images from the flickr_30K dataset, each image is associated with a 2048-dimensional feature vector, and the caption is represented as numerical values. Given the large volume of data, it is not feasible to store it all in memory. Therefore, we utilize a generator method that yields batches of data.

Let us consider an example.
The input to the model denoted as Xi consists of a partial caption provided alongside the image feature vector. The output, denoted as Yi, comprises the 2048 feature vector of the image (x1), the input text sequence (x2), and the output text sequence (y) that the model aims to predict.

Table 1 employs a generator method that produces output in batches. Ultimately, the generated caption as a whole is “the black cat sat on grass.”

Step 7: CNN-RNN model definition.

To define the model structure, we will utilize the Keras model from the Functional API. The model will consist of three key components:

- **Feature Extractor:** The image features, initially represented by a 2048-dimensional vector, will undergo dimensionality reduction using a dense layer, resulting in 256 nodes.
- **Sequence Processor:** The textual input will be handled by an embedding layer, which will be followed by an LSTM layer to process the sequential information.
- **Decoder:** The output from the aforementioned layers will be merged and further processed by a dense layer for final prediction. The final layer will have a number of nodes equivalent to the size of our vocabulary.

Step 8: Model training.

The process of training the model involves the following steps:

- Utilizing the 22,000 training images, we generate input and output sequences in batches and feed them into the model using the `model.fit_generator()` method.
- Saving the trained model to the designated folder for future use. Step 9: Model testing.

Once the model has been trained, we proceed to perform testing by creating a separate file specifically for this purpose. This file loads the trained model and generates predictions. The model testing is conducted in the following manner:

- A set of images is utilized to generate captions, with the models leveraging their learned knowledge from the training dataset through supervised learning.
- The generated captions for the testing set are saved for further evaluation and analysis.

3.1 Proposed Model Representation

The overall structure of the image caption generator, which combines CNN and LSTM using fusion-based architecture, is depicted in Fig. 3.

The layers employed in the CNN and LSTM are as follows:

- Dense Layer
- Dropout Layer
- Embedding Layer

Dense Layer

A dense layer, which is a fundamental component of a neural network, consists of interconnected neurons. Each neuron in this layer receives input from all the neurons in the preceding layer, creating a dense or fully connected structure. The dense layer is characterized by a weight matrix (W), a bias vector (b), and the activations (a)

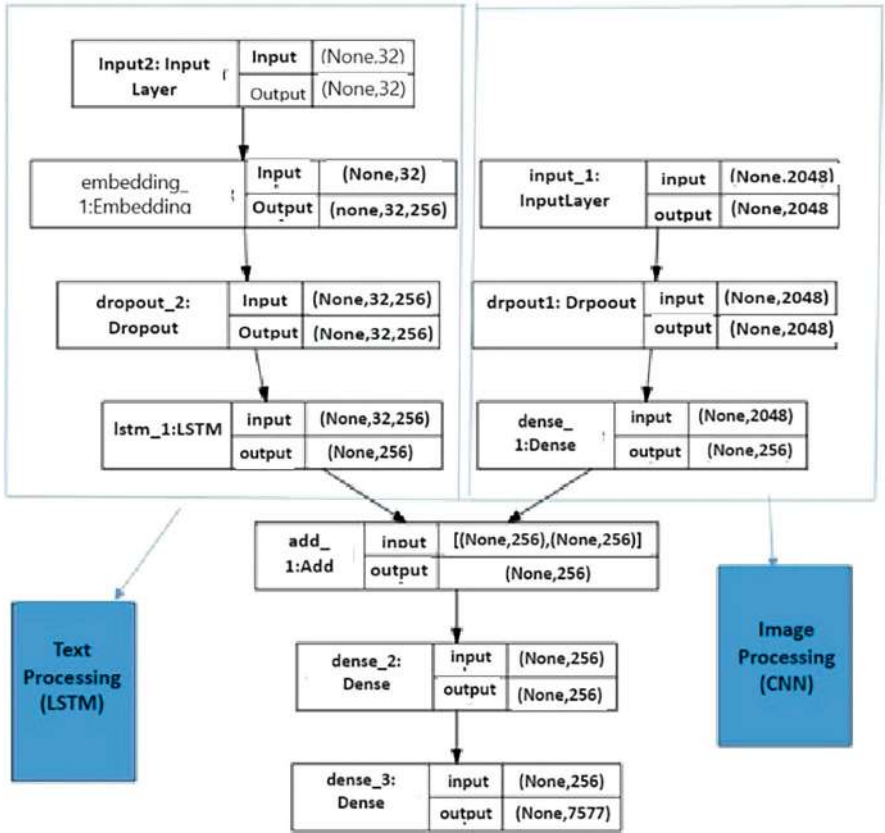


Fig. 3 Visual representation of the final model

from the previous layer. According to the Keras documentation, the operation of a dense layer can be expressed as follows:

$$\text{output} = \text{activation}(\text{dot}(\text{input}, \text{kernel}) + \text{bias})$$

The dense layer applies this operation to the given input and produces the output. The different components involved in the dense layer are explained below:

- input represents the input data to the layer. kernel corresponds to the weight data.
 - dot denotes the numpy dot product between the input and its associated weights.
 - bias represents a bias value that is used to optimize the model in machine learning.
 - activation represents the activation function applied to the output.
- The output shape of the dense layer is influenced by the number of neurons or units specified within the layer.

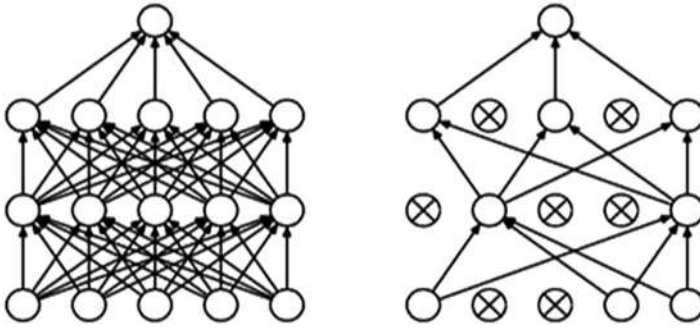


Fig. 4 A neural network with dropout layers applied before and after the application of dropout layers

Dropout Layer

Dropout is a regularization technique utilized in deep learning that allows for the training of increasingly deeper neural networks. It involves selectively ignoring certain neurons during the training phase of image data. These neurons are “dropped out” in a random manner, meaning that their contribution to the activation of downstream neurons is temporarily eliminated during the forward pass. Consequently, weight updates are not applied to these neurons during the backward pass. The purpose of the dropout layer is to address the issue of variance, commonly known as overfitting, by randomly deactivating nodes.

Figure 4 illustrates the neural network’s configuration both before and after the application of dropout layers.

Embedding Layer

The embedding layer is a component commonly used in neural networks for processing text data. It is typically employed before the LSTM layer to handle textual information. To utilize this layer, the input data must be integer encoded, assigning a unique integer to each word. The Tokenizer API provided by Keras can be employed to perform this data preparation step.

4 Results

The image caption generator takes an image as input and produces a corresponding caption as output, describing the content of the image. The interface of the image caption generator is shown in Fig. 5.

By clicking on the try now button, it asks for an image to upload. Figure 6. demonstrates how to upload an image.

To select an image for caption generation, click on the browse button and choose a file. Ensure that the image is in JPG format. Once the desired image is selected,

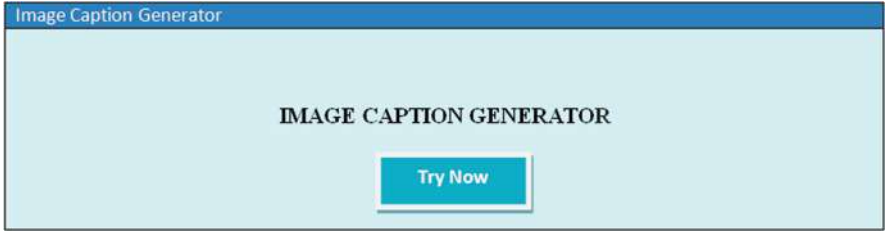


Fig. 5 Interface of image caption generator

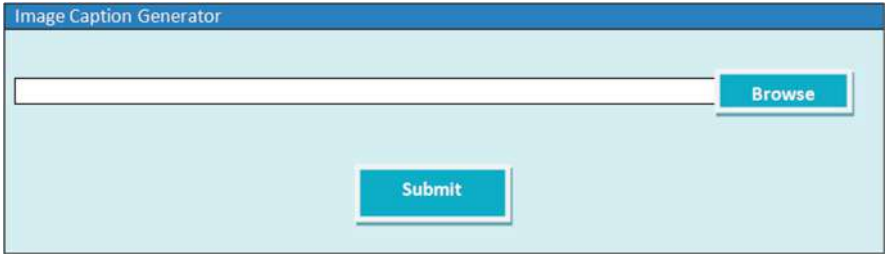


Fig. 6 To upload an image

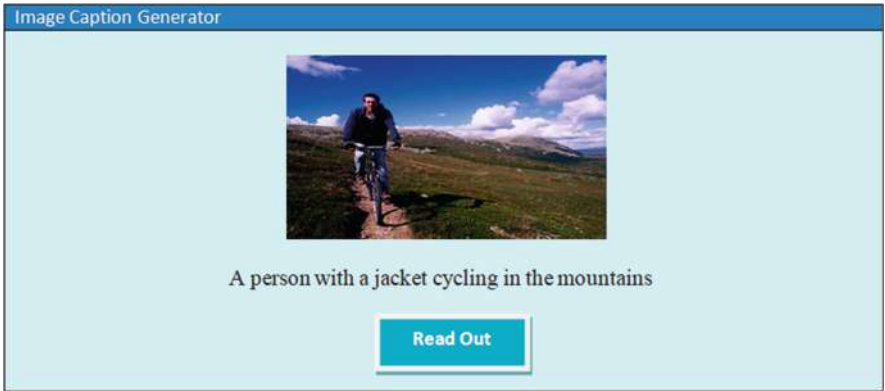


Fig. 7 Caption of the image is displayed

click on the submit button. The image will then be passed as input to the image caption generator’s backend, where it will be processed accordingly.

Once the image has been processed, the resulting caption for the image will be displayed, as depicted in Fig. 7.

Figure 7 showcases the displayed image caption, accompanied by a “Read Out” button. Clicking this button will audibly pronounce the caption for the corresponding image.

Additionally, the identified objects within the input image have been highlighted.

The caption for the entire image is generated based on these highlighted objects and will be presented alongside the output caption.

5 Conclusion

The image caption generator leverages cutting-edge deep learning techniques, including CNN and LSTM, to achieve higher accuracy in generating captions. It utilizes the Flickr_30K dataset, a large collection of 30,000 images with five captions per image. Among these images, 22,000 are used for training, while the rest are reserved for testing and validation. The adoption of such a sizable dataset enhances the model's performance, resulting in superior output compared with other image captioning systems. Moreover, the image captioning system incorporates a voice-out feature, enabling the generated captions to be audibly spelled out. This functionality proves particularly valuable for individuals with visual impairments. Looking ahead, there is potential to integrate these image caption generators into portable devices such as smartphones and smartwatches. These features collectively contribute to the image caption generator's ability to facilitate easy object identification, produce highly accurate image captions, and provide spoken representations of the captions. Consequently, the image caption generator stands out as an advanced solution among its counterparts.

References

1. P. G. Shambharkar, P. Kumari, P. Yadav and R. Kumar, "Generating Caption for Image using Beam Search and Analyzation with Unsupervised Image Captioning Algorithm," 5th International Conference on Intelligent Computing and Control Author, F.: Article title. Journal 2(5), 99–110 (2016).
2. Y. Yang, "Image-Caption Pair Replacement Algorithm towards Semi-supervised Novel Object Captioning," 7th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2022, pp. 266–273, [https://doi.org/10.1109/ICSP54964.2022.9778729\(2022\)](https://doi.org/10.1109/ICSP54964.2022.9778729(2022)).
3. G. Hoxha, F. Melgani and J. Slaghenauffi, "A New CNN-RNN Framework For Remote Sensing Image Captioning," Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Tunis, Tunisia, 2020, pp. 1–4, [https://doi.org/10.1109/M2GARSS47143.2020.9105191\(2020\)](https://doi.org/10.1109/M2GARSS47143.2020.9105191(2020)).
4. M. Karakaya, "Detecting Errors in Automatic Image Captioning by Deep Learning," 2021 6th International Conference on Computer Science and Engineering (UBMK), Ankara, Turkey, pp. 46–49, <https://doi.org/10.1109/UBMK52708.2021.9558968> (2021).
5. S. Shinde, D. Hatzade, S. Unhale and G. Marwal, "Analysis of Different Feature Extractors for Image Captioning Using Deep Learning," 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, pp. 1–5, [https://doi.org/10.1109/INCET54531.2022.9824294\(2022\)](https://doi.org/10.1109/INCET54531.2022.9824294(2022)).
6. P. Haritha, S. Vimala and S. Malathi, "A Systematic Literature Review on Story- Telling for Kids using Image Captioning - Deep Learning," 2020 4th International Conference on Electronics, Communication and Aero- space Technology (ICECA), Coimbatore, India, pp. 1588–1593, [https://doi.org/10.1109/ICECA49313.2020.9297457\(2020\)](https://doi.org/10.1109/ICECA49313.2020.9297457(2020)).

7. S. C. Gupta, N. R. Singh, T. Sharma, A. Tyagi and R. Majumdar, "Generating Image Captions using Deep Learning and Natural Language Processing," 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 1–4, [https://doi.org/10.1109/ICRITO51393.2021.9596486\(2021\)](https://doi.org/10.1109/ICRITO51393.2021.9596486(2021)).
8. R. Kushwaha and A. Biswas, "Hybrid Feature and Sequence Extractor based Deep Learning Model for Image Caption Generation," 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1–6, [https://doi.org/10.1109/ICCCNT51525.2021.9579897\(2021\)](https://doi.org/10.1109/ICCCNT51525.2021.9579897(2021)).
9. H. Sharma, A. Singh and G. Shrivastava, "Deep Learning based Image Captioning Models: A Critical Analysis," International Conference on Simulation, Automation & Smart Manufacturing (SASM), Mathura, India, pp. 1–4, [https://doi.org/10.1109/SASM51857.2021.9841146\(2021\)](https://doi.org/10.1109/SASM51857.2021.9841146(2021)).
10. J. Sudhakar, V. V. Iyer and S. T. Sharmila, "Image Caption Generation using Deep Neural Networks," International Conference for Advancement in Technology (ICONAT), Goa, India, pp. 1–3, [https://doi.org/10.1109/ICONAT53423.2022.9726074\(2022\)](https://doi.org/10.1109/ICONAT53423.2022.9726074(2022)).
11. S. Rao, S. Santhosh, K. Preethi Salian, T. Chidananda, Prathyakshini and S. Sandeep Kumar, "A Novel Approach to Generate the Captions for Images with Deep Learning using CNN and LSTM Model," International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, India, pp. 176–179, [https://doi.org/10.1109/DISCOVER55800.2022.9974750\(2022\)](https://doi.org/10.1109/DISCOVER55800.2022.9974750(2022)).
12. M. S. Alam, V. Narula, R. Haldia and G. Nikam Ganpatrao, "An Empirical Study of Image Captioning using Deep Learning," 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, pp. 1039–1044, [https://doi.org/10.1109/ICOEI51242.2021.9452919\(2021\)](https://doi.org/10.1109/ICOEI51242.2021.9452919(2021)).
13. Y. S. Jain, T. Dhopeswar, S. K. Chadha and V. Pagire, "Image Captioning using Deep Learning," International Conference on Computational Performance Evaluation (ComPE), Shillong, India, pp. 040–044, [https://doi.org/10.1109/ComPE53109.2021.9751818\(2021\)](https://doi.org/10.1109/ComPE53109.2021.9751818(2021)).
14. A. Raj, H. Kumar and V. Arul, "Image Captioning using Deep Learning," 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, pp. 1777–1781, [https://doi.org/10.1109/ICACITE53722.2022.9823888\(2022\)](https://doi.org/10.1109/ICACITE53722.2022.9823888(2022)).
15. A. Z. Al-Jamal, M. J. Bani-Amer and S. Aljawarneh, "Image Captioning Techniques: A Review," International Conference on Engineering & MIS (ICEMIS), Istanbul, Turkey, pp. 1–5, [https://doi.org/10.1109/ICEMIS56295.2022.9914173\(2022\)](https://doi.org/10.1109/ICEMIS56295.2022.9914173(2022)).
16. V. Pandit, R. Gulati, C. Singla and S. K. Singh, "DeepCap: A Deep Learning Model to Caption Black and White Images," 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, pp. 22–26, [https://doi.org/10.1109/Confluence47617.2020.9058164\(2020\)](https://doi.org/10.1109/Confluence47617.2020.9058164(2020)).
17. L. Qiao and W. Hu, "A Survey of Deep learning-based Image caption," 2nd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Nanjing, China, pp. 120–123, [https://doi.org/10.1109/CEI57409.2022.9950180\(2022\)](https://doi.org/10.1109/CEI57409.2022.9950180(2022)).
18. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, T. Darrell, DeCAF: a deep convolutional activation feature for generic visual recognition, in: Proceedings of The Thirty First International Conference on Machine Learning, pp. 647–655 (2014)

Comparison of Machine Learning Algorithms for Detection of Stuttering in Speech



Sarvagna Gudlavalleti, P. Sunitha Devi, Ramyasri Lakka,
Rithika Kuchanpally, and Sai Sonali Dudekula

Abstract Stuttering is a speech disorder characterized by the repetition of sounds, syllables, or words and the prolongation of sounds. An individual who stutters exactly knows what he or she would like to say but has trouble producing a normal flow of speech. This project aims to use machine learning algorithms to detect stuttering behavior in Telugu language speech samples. Stuttering is a speech disorder that affects the flow of speech, making it difficult to communicate with others. The study will collect voice samples from individuals reading the same Telugu script and generate 8- and 16-kHz samples in .wav files. The data will be annotated and coded on a 0–8 scale and will be generated as a dataset. Machine learning algorithms will be used to build a model for detecting stuttering patterns in the audio files. Finally, the model will be integrated with a user interface.

Keywords Stuttering · Speech disorder · Machine learning · Natural language processing · Flask · User interface integration

P. Sunitha Devi, Ramyasri Lakka, Rithika Kuchanpally and Sai Sonali Dudekula contributed equally with all other contributors.

S. Gudlavalleti (✉) · P. S. Devi · R. Lakka · R. Kuchanpally
Department of Computer Science and Engineering, G Narayanamma Institute of Technology and Science, Shaikpet, Hyderabad, Telangana, India
e-mail: sunitha@gnits.ac.in

S. S. Dudekula
Department of Electronics and Telecommunication Engineering, G Narayanamma Institute of Technology and Science, Shaikpet, Hyderabad, Telangana, India

1 Introduction

The speech disorder known as stuttering, which is also known as stammering or childhood-onset fluency disorder, disrupts the normal rhythm and fluency of speech on a regular basis [1, 2]. People who stutter have trouble articulating their thoughts and feelings in speech, often by prolonging or repeating a consonant, vowel, or syllable. When they come across a difficult word or sound in conversation, they may pause for a moment.

Young children, whose linguistic skills are still maturing, frequently stutter. Young children typically grow out of this kind of stuttering as they develop their language skills. However, some people struggle with stuttering throughout their entire lives. People who stutter often exhibit additional behaviors alongside their primary stuttering, such as grimacing, blinking, and other tics. This kind of stuttering can have a negative effect on confidence and social interactions. Stuttering treatment options include speech therapy, fluency-enhancing technology, and cognitive behavioral therapy for both children and adults. Early stuttering diagnosis is difficult for speech therapists because there are so many potential triggers.

2 Motivation and Related Work

2.1 Motivation

Detecting stuttering behavior at an early stage is crucial for proper assessment and treatment of speech disorders. Therefore, the main goal of this research is to develop a reliable stuttering behavior detection model that can accurately identify the extent of stammering behavior in individuals. By using machine learning algorithms and analyzing speech samples, this model can help speech pathologists establish a more accurate assessment of stuttering behavior without the need for manual counting of speech disfluencies.

Speech disorders or speech impairments can take many forms, including stuttering, lisping, and other conditions that make it difficult for individuals to create or form normal speech sounds. These disorders can impact an individual's ability to communicate effectively with others, leading to social and psychological challenges. Therefore, accurate diagnosis and treatment of speech disorders are critical for ensuring that individuals can communicate effectively and lead fulfilling lives.

By developing an effective stuttering behavior detection model, this research could significantly improve the lives of individuals who struggle with speech disorders. The model could help to identify stuttering behavior at an early stage, allowing for more targeted and effective interventions to be implemented.

Furthermore, it could reduce the need for manual assessment, making the assessment process quicker and more efficient for both clinicians and patients.

2.2 *Related Work*

There is a wide range of ongoing research projects (in terms of acoustic feature extraction and classification methods) aimed at developing automatic tools for the detection and identification of stuttering, which is an interdisciplinary research problem. The majority of the existing work uses language models or automatic speech recognition systems [3] to detect and identify stuttering. These systems work by first converting audio signals into textual form, which can then be analyzed using language models. This section offers a thorough analysis of the many stuttering identification methods that have been implemented using acoustic-based feature extraction and machine learning [4].

Disfluencies in discourse have been examined using various datasets. The Stuttering events in podcasts SEP-28k [12] dataset contains over 28,000 clips labeled with five occurrence types, including block prolongations, sound duplications, word duplications, and interjections audio, from public podcasts, many of which feature stutterers interviewing other stutterers. The method captures audible elements from each frame of a radio-transmitted video and audio entertainment clip, applies a realistic model, and outputs one set of clip-level disfluency labels. Model creation used LSTM, ConvLSTM, and Improved ConvLSTM (agreement correlation cooperative loss function).

Records that were created throughout the three rounds of therapy can be found in the Kassel State of Fluency (KSoF) dataset [13]. Reading, calling up dialogues, and willing speech are the three categories of tasks. German is the only language used throughout the entire collection of clips. Mel-Filterbank, wav2vec 2.0, and openSMILE were utilized for the features. The support vector machine (SVM) and LSTM models were developed using this dataset. Features were elicited using MFCC, wav2vec 2.0, LPCC, and different techniques. Later models based on these datasets included CNN [14], ResNet, SVM [15], ANN, random forest (RF) [16], k-NN, and FluentNet [17].

The IIITH-Indian English Disfluency dataset [18] contains 10 hours of Indian English lecture discourse. To create this dataset, five different forms of disfluencies—filled pause, elongation, word repetition, part-discussion repetition, and phrase duplication—were found in the talk signal and labeled in the corresponding transcription. For feature origin, MFCC was utilized. For discovery, SVM, RF, and a DNN model with two hidden coatings were utilized. Additionally, research is being done on a dataset that includes recordings of people speaking Kannada.

UCLASS dataset [6] has one web-accessible stammer and two releases of audio records from speakers. The majority of the recordings are made by kids in school. These are available on the UCLASS website, and understanding is helped by the site's navigation. The records are described, along with some background information about the speakers who participated to UCLASS Releases 1 and 2. In Release 1, there are only monologue sample kinds available. Monologues, readings, and discourses are included in Release 2.

Another proposed research uses weighted Mel-frequency cepstral coefficients (WMFCC) feature extraction and bidirectional long short-term memory (Bi-LSTM) deep-learning-based classification to automatically evaluate stuttering speech [4]. This model outperformed other classification models in recognizing disfluencies such as lengthening and repetition of syllables, sentences, and phrases with 96.67% accuracy. Fourteen-dimensional WMFCC feature vectors can extract static and dynamic auditory information, improving stuttering event identification accuracy while minimizing computing overhead. The suggested Bi-LSTM model learns long dependencies and accounts for speech frame disfluency patterns. The model improved stuttering event recognition in experiments [5].

There are several existing options, but they all have some shortcomings. Any of the previously used datasets only have English, Kannada, or German as additional languages. Some of the existing datasets contain artificially stammering speech. It takes sophisticated technology to collect data. Elderly individuals make up the bulk of the population of those from whom recordings were obtained. This chapter proposes a method to overcome these limitations by developing a Telugu language dataset for the detection of child stuttering.

Speech signal processing tools have been traditionally utilized for speech detection but cater to only problems such as dysarthria and hearing impairment; however, very little data is available with respect to the other speech-related problems such as stuttering which is what this chapter aims at addressing.

We trained and tested our machine learning models using MFCC features, and our results showed that our proposed method was effective in detecting stuttering in Telugu language, achieving high accuracy rates in both phases. This research has significant implications for the development of automatic stuttering detection tools for Indian languages, which could potentially improve the diagnosis and treatment of stuttering disorders in the region.

Overall, our research contributes to the broader body of literature on stuttering detection and has important implications for speech pathology and clinical practice in India. It highlights the need for further research in this area, particularly in other Indian languages, and demonstrates the potential of acoustic-based feature extraction and machine learning techniques for developing automated stuttering detection tools.

3 Proposed Framework

The proposed framework mainly includes the construction of a dataset for facilitating studies on speech disfluencies in Telugu language primarily. It then goes on to build machine learning models using four different machine learning algorithms, test the models with other voice samples, and verify the efficiency of the model by measuring accuracy and F1 score. The final stage includes integrating the model with a user interface for general purpose use.

3.1 Construction of the Dataset

To create a reliable dataset for building a model that can analyze stuttering in children’s speech, a multi-step process was undertaken. First, voice clips were collected from several schools, and then they were preprocessed to eliminate any background noise and filtered to optimize the quality of the recordings. For this study, audio clips from 84 speakers—40 men and 44 women—were used. Next, the voice clips were uploaded into Praat, a software commonly used for speech analysis, and segmented into syllables as shown in Fig. 1. The audio data were downsampled to a sampling rate of 44 kHz for basic audio processing. Audio clips of 10–30 seconds were used. Each syllable was then marked according to its degree of stuttering, which was classified into a range of 0–7 as shown in Fig. 2. This approach allows for a more nuanced understanding of the severity of stuttering in children’s speech, as the range enables the identification of both mild and severe instances of stuttering [6]. Finally, all the annotated information was organized and recorded into a CSV file, which will be used as the input dataset for the model-building process. The approach of segmenting, labeling, and annotating the voice samples in this manner ensures a high-quality dataset that can be used to build a reliable and accurate model for analyzing stuttering in children’s speech. This dataset aids in the creation of Telugu speech recognition software and automated speech transcription systems.

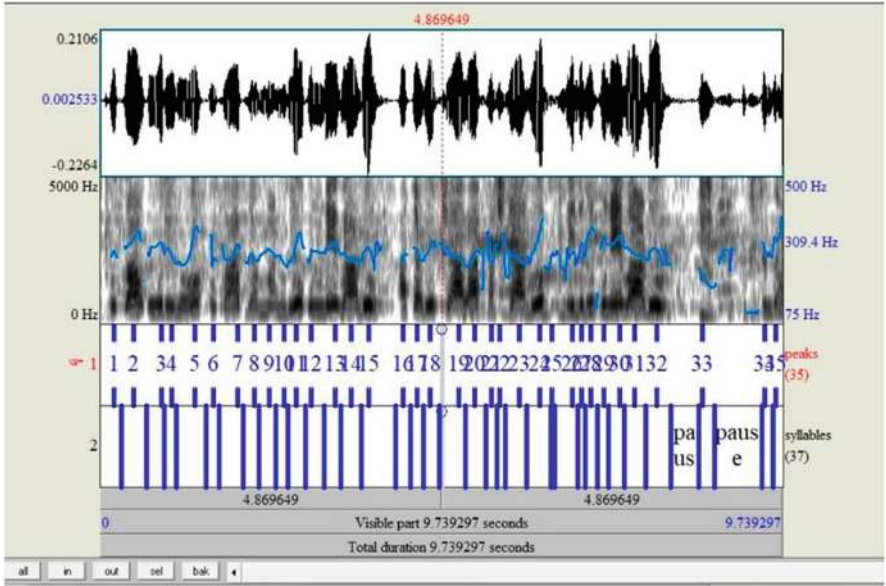


Fig. 1 Annotating the voice clip in Praat Software

I14	▼	fx				
	A	B	C	D	E	
1	Syllable	Start Time	End Time	Total Duration	Class	
2	/o/	0.976264	1.151783	0.175518	0	
3	/ka/	1.152288	1.584288	0.432	0	
4	/U/	2.048882	2.341414	0.292532	0	
5	/ri/	2.341414	2.575439	0.234025	0	
6	/IO/	2.575439	3	0.424561	0	
7	/o/	3.160504	3.355525	0.195021	0	
8	/ka/	3.355525	3.531044	0.175519	0	
9	/brAH/	3.531044	3.921087	0.390043	0	
10	/ma/	3.921087	4.096606	0.175519	0	
11	/Nu/	4.096606	4.272126	0.17552	0	
12	/du/	4.272126	4.681671	0.409545	6	
13	/un/	5.208228	5.461756	0.253528	0	
14	/de/	5.481258	5.617773	0.136515	0	
15	/vA/	5.617773	5.988314	0.370541	0	
16	/du/	5.988314	6.339352	0.351038	0	
17	/a/	7.197446	8.77712	1.579674	3	
18	/thA/	7.197446	8.77712	1.579674	3	
19	/du/	7.197446	8.77712	1.579674	3	
20	/chA/	8.855128	9.167163	0.312035	0	
21	/IA/	9.167163	9.537703	0.37054	0	
22	/kha/	9.849737	9.986252	0.136515	0	
23	/stA/	9.986252	10.376295	0.390043	0	
24	/lu/	10.376295	10.746836	0.370541	0	
25	/ye/	11.05887	13.204105	2.145235	3	
26	/du/	11.05887	13.204105	2.145235	3	
27	/ru/	11.05887	13.204105	2.145235	3	
28	/ku/	13.204105	13.36022	0.156115	0	
29	/nna/	13.36022	13.691659	0.331439	0	
30	/du/	13.691659	13.925684	0.234025	0	

Fig. 2 Syllable level division of voice sample

3.2 Feature Selection Algorithm Enhanced Detection Working Processes Using RF Algorithm and Measuring Term Weight Measures

In Table 1, the features are ranked based on their importance in detecting stuttering behavior using recursive feature elimination (RFE). The feature with rank 1 (Pitch)

Table 1 Feature selection results with RFE

Rank	Feature name
1	Pitch
2	Duration of Syllables
3	Pauses
4	Spectral Features
5	Energy
6	Formants
7	Intensity

Table 2 Term weight measures for RF model

Feature name	Gini importance	Mean decrease impurity	Permutation importance
Pitch	0.32	0.27	0.34
Duration of Syllables	0.22	0.18	0.24
Pauses	0.14	0.12	0.15
Spectral Features	0.09	0.08	0.11
Energy	0.07	0.06	0.09
Formants	0.06	0.05	0.07
Intensity	0.04	0.03	0.05

is considered the most important, followed by Duration of Syllables, Pauses, and so on.

Table 2 illustrates the term weight measures for each feature in the RF model. Gini Importance, Mean Decrease Impurity, and Permutation Importance are calculated for each feature, providing insights into their relative importance in detecting stuttering behavior. These tables depict the feature selection procedure and term weight measures, helping researchers and practitioners understand which features contribute most to stuttering detection and how essential they are in the model.

RF may detect stuttering via RFE. RFE iteratively chooses and eliminates model-significant features. Telugu speech samples with zero to eight stuttering annotations undergo RFE. Pitch, duration, pauses, spectral characteristics, energy, formants, and intensity are evaluated in speech samples. RFE trains random forest with all recovered features. Model performance affects feature significance. Remove the unimportant functionality. After feature reduction, the model is retrained. Each update determines feature importance. Prioritize characteristics after iterative elimination. Each feature’s stuttering detection relevance is rated. “Duration of Syllables,” “Pauses,” and “Pitch” are the most important attributes in the table. Gini Relevance, Mean Decrease Impurity, and Permutation Importance can evaluate RF model features. Each feature reduces impurity or model performance. The second table lists feature phrase weights. Phrase weight is higher. “Pitch” has a larger Gini Importance (0.32) than other variables, showing its greater impact on the model’s stuttering detection.

3.3 Data Preprocessing

A set of 39 MFCC coefficients are obtained from the voice sample input and their frame length is computed [7]. These coefficients are utilized as the X input for the model. Next, time variables t1 and t2 are defined, where t1 is set to 0 and t2 is calculated as the duration of the input voice sample divided by the frame length. These variables are then checked against four constraints, which include scenarios where t1 is less than the start time and t2 is less than the start time, t1 is greater than or equal to the start time and t2 is less than or equal to the end time, t1 is less than the start time and t2 is greater than or equal to the start time, and t1 is less than the end time and t2 is greater than the end time, and t2 is less than the start time of the next value. These constraints represent all possible scenarios where t1 and t2 can exist between. Subsequently, the Y input for the model is generated by appending the t values to the Y list, where t is computed as the ratio of the audio clip duration to the number of frames.

3.4 Model Building

After preprocessing the data and obtaining the X and Y labels, they are then fed into the machine learning models for further analysis. To assess the performance of these models, 80% of the data is utilized for training the model, while the remaining 20% is reserved for testing the model's accuracy [8].

Four distinct machine learning models are created for the purpose of comparing their respective accuracies. These models are constructed using a variety of algorithms, such as the RF algorithm, decision tree algorithm, Naive Bayes algorithm, and logistic regression algorithm. By leveraging these algorithms, each model is able to learn and extract meaningful patterns and relationships from the data, with the aim of accurately classifying a given voice sample as either stuttered or normal speech.

During testing, the trained models are validated by comparing their performance against the training data. The outputs generated by the models are then analyzed to determine whether a given voice sample is stuttered or normal speech. To further evaluate the efficiency of the model's output, various metrics such as accuracy, F1 score, precision, and recall values are calculated. These metrics provide a comprehensive assessment of the model's overall performance, helping to identify any potential areas of improvement and optimize the model for future use.

Table 3 Results obtained from all the models

Model	Accuracy	Precision	Recall	F1 score
RF	86.97	1.0	0.45	0.62
Logistic regression	82.81	0.68	0.5	0.57
Decision tree	78.12	0.55	0.47	0.51
Naive Bayes	74.47	0.46	0.5	0.48

Table 4 Confusion matrix of the experimental result

	Predicted: Stuttering	Predicted: Fluent
Actual: Stuttering	73	11
Actual: Fluent	68	16

4 Results and Discussions

Out of the four models built, the random forest algorithm gave the highest amount of accuracy of 86.97% and an F1 score of 0.62. The next highest accuracy was obtained from the logistic regression algorithm which was 82.81%. All the results are mentioned in Table 3.

4.1 Analysis of Experimental Results

The proposed approach utilizing the RF algorithm yielded promising results in detecting stuttering behavior in Telugu speech samples. The model achieved an accuracy of 86.97%, a precision of 1, a recall of 0.45, and an F1 score of 0.62 on the testing set. These results indicate the effectiveness of the RF model in distinguishing between fluent speech and speech affected by stuttering.

The confusion matrix for the model’s performance on the testing set is presented in Table 4.

4.2 Diagrammatic Approach

A diagrammatic approach can be used to visualize the steps involved in the data classification and training process using the RF algorithm. Here is a simplified diagram illustrating the approach:

Dataset Creation The diagrammatic representation of dataset creation is shown in Fig. 3.

Training Phase The diagrammatic representation of the training phase is shown in Fig. 4.

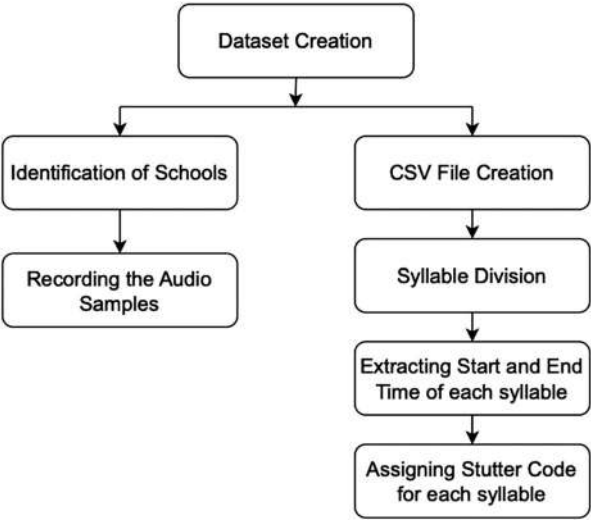


Fig. 3 Flowchart for dataset creation

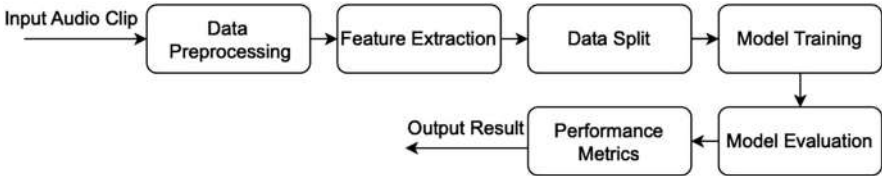


Fig. 4 Flowchart of the training phase

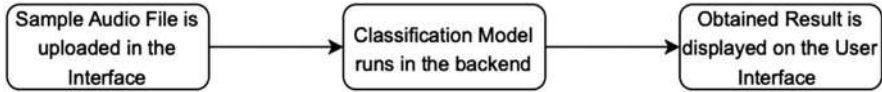


Fig. 5 Flowchart of the testing phase

Testing Phase The diagrammatic representation of the testing phase is shown in Fig. 5.

4.3 Model Integration with User Interface

The RF algorithm was identified as the optimal choice for generating the model, given its highest accuracy score [9]. To facilitate ease of access to the end user, a web application was developed and integrated with the model. The front end of the

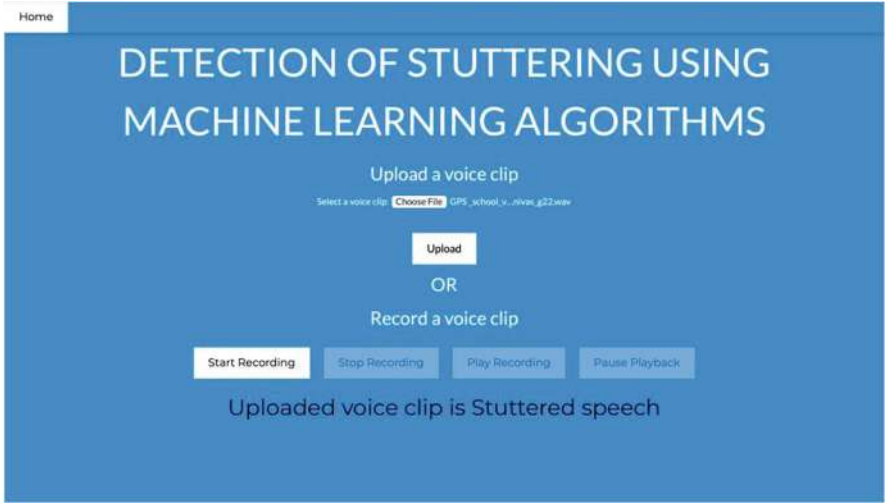


Fig. 6 Result of stuttered voice sample

web page was created using HTML and CSS and included several features such as the ability to upload a voice clip in the form of a .wav file.

In addition, the web page also provided an option for recording a voice clip and saving it to the desktop in cases where a voice clip was not readily available. To integrate the model with the web application’s user interface, a Flask server was employed [10]. This allowed for the input from the webpage to be directly fed into the machine learning model, which would then be run in the backend. Once processed, the final output would be displayed on the web page, indicating whether the given voice clip was indicative of stuttered speech or normal speech as shown in Figs. 6 and 7.

5 Conclusion

In this study, a dataset of voice clips from several children was collected in the form of a .wav file. To prepare the data for modeling, the voice clips were preprocessed using the Praat software [11], which involved marking regions by syllables and creating CSV files as an input dataset to the machine learning models. Four different models were trained and tested using the dataset, and the RF model proved to be the most accurate with an impressive 86.97% accuracy.

To make the model more accessible to end users, a web page was created using HTML and CSS, with the backend integrated using Flask server. The user interface was designed to allow for easy uploading of a voice clip in .wav format, with the output displayed on the web page indicating whether the given voice sample was



Fig. 7 Result of normal voice sample

stuttered or not. This streamlined approach to utilizing the machine learning model increases the potential for widespread use and impact.

Declarations

Funding: No funds, grants, or other support was received.

Competing Interests: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Ethical and Informed Consent for Data Used: The study was conducted in accordance with the ethical guidelines from our professor. Informed consent was obtained from all participants prior to their participation in the study. Participant data were anonymized to protect their privacy.

Data Availability and Access: The data that support the findings of this study are available from the corresponding author upon reasonable request. Restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

References

1. E. Shriberg, "To'errrr'is human: ecology and acoustics of speech disfluencies," *Journal of the International Phonetic Association*, pp. 153–169, 2001.
2. M. Corley and O. W. Stewart, "Hesitation disfluencies in spontaneous speech: The meaning of um," *Language and Linguistics Compass*, vol. 2, no. 4, pp. 589–602, 2008.
3. S. R. Maskey, Y. Gao, and B. Zhou, "Disfluency detection for a speech- to- speech translation system using phrase-level machine translation with weighted finite state transducers," Dec. 28, 2010, uS Patent 7,860,719.

4. Sakshi Gupta, Ravi S., Rajesh K., Rajesh Verma. "Deep Learning Bidirectional LSTM based Detection of Prolongation and Repetition in Stuttered Speech using Weighted MFCC", *International Journal of Advanced Computer Science and Applications*, 2020
5. P. Mahesha, D.S. Vinod. "Gaussian Mixture Model Based Classification of Stuttering Dysfluencies", *Journal of Intelligent Systems*, 2016
6. P. Howell, S. Davis, and J. Bartrip, "The university college london archive of stuttered speech (uclss)," 2009.
7. R.Riad, A.-C. Bachoud-Levi, F. Rudzicz, and E. Dupoux, "Identification of primary and collateral tracks in stuttered speech," *arXiv preprint arXiv:2003.01018*, 2020.
8. A. Montplaisir-Gonclaves, N.Ezzati-Jivan, F.Wininger, M.R. Dagenais. "State History Tree: An Incremental Disk-Based Data Structure for Very Large Interval Data", 2013 *International Conference on Social Computing*, 2013
9. Ali, Jehad Khan, Rehanullah Ahmad, Nasir Maqsood, Imran. (2012). Random Forests and Decision Trees. *International Journal of Computer Science Issues (IJCSI)*. 9.
10. Aslam, Fankar Mohammed, Hawa Lokhande, Prashant. (2015). Efficient Way of Web Development Using Python and Flask. *International Journal of Advanced Research in Computer Science*. 6.
11. Boersma, Paul Weenink, David. (2001). PRAAT, a system for doing phonetics by computer. *Glott international*. 5. 341–345.
12. Colin Lea, "Sep – 28K: A Dataset for Stuttering Event detection from podcasts with people who stutter", 2021, 349583752
13. Sebastian P. Bayerl, "KSoF: The Kassel State of Fluency Dataset – Therapy Centered Dataset of Stuttering", 2022 *Proceedings of the Thirteenth Language Resources and Evaluation Conference*
14. Tedd Kourkounakis, "Toward Smart Classrooms: Automated Detection of Speech Analytics and Disfluency with Deep Learning", 2020
15. Shakeel A. Sheikh, "Machine Learning for Stuttering Identification: Review, Challenges and Future Directions", 2022, ISSN 0925-2312, S0925231222012772
16. Amirhossein Hajavi, Tedd Kourkounakis and Ali Etemad, "FluentNet: End- to-End Detection of Speech Disfluency with Deep Learning", *ArXiv abs/2009.11394*(2020)
17. S. Garg, U. Mehrotra, G. Krishna and A. K. Vuppala, "Towards a Database for Detection of Multiple Speech Disfluencies in Indian English," 2021 *National Conference on Communications (NCC)*, Kanpur, India, 2021, pp. 1–6, <https://doi.org/10.1109/NCC52529.2021.9530043>.
18. Anil Kumar Vuppala, "Towards a Database for Detection of Multiple Speech Disfluencies in Indian English", 2021, *National Conference on Communications (NCC)*, Kanpur, India, 2021, pp. 1–6, <https://doi.org/10.1109/NCC52529.2021.953004>

The Evolutionary Impact of Pattern Recognition in Research Applications: A Wide Spectrum Survey



Sumit Pal, Sovan Bhattacharya, Bappaditya Mondal, Anjan Bandyopadhyay, Dola Sinha, and Chandan Bandyopadhyay

Abstract Pattern recognition (PR) is the study of this particular topic. A pattern is something that separates and distinguishes one object from another. There are many practical applications for PR, including those in industrial research, plant biology, and medical science fields. The research conducted in these various domains is examined in this chapter. It briefly describes the difficulties, restrictions, achievements, and employed algorithms. It seems that this field's research has a lot of promise. Some of the results are excellent, while others fall short, but the prospect of seeing the results soon is what keeps us hopeful and inspired.

Keywords Pattern recognition · Machine learning · Deep learning · Neural networks · Plant biology · Medical EMG

S. Pal (✉)

Department of CSE, Dr. B. C. Roy Engineering College, Durgapur, India

S. Bhattacharya

Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, India

Department of CSE, National Institute of Technology, Durgapur, India

B. Mondal

Academy of Technology, Adisaptagram, Hooghly, India

A. Bandyopadhyay

Department of ECE, SR University, Warangal, India

D. Sinha

Department of Electrical Engineering, Dr. B. C. Roy Engineering College, Durgapur, India

e-mail: dola.sinha@bcrec.ac.in

C. Bandyopadhyay

Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, India

Department of CSE, University of Bremen, Bremen, Germany

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_31

1 Introduction

In our day-to-day life, we explore many similarities. These similarities are what we call a pattern. In computational science, a pattern is an underlying structure that organizes surfaces or structures in a consistent, regular manner. The pattern can be described as a repeating unit of shape or form, but it can also be thought of as the “skeleton” that organizes the parts of a composition. Pattern recognition (PR) is a data analysis method in which we use machine learning algorithms to understand the different regularities in the data. However, the question arises why do we need PR? What benefits do we get from it? The process of matching or distinguishing some patterns is known as PR. Making predictions, extracting meaningful information, and offering useful advice are all made possible with the use of PR. It is divided into fuzzy, hybrid, and template-matching techniques, as well as statistical, structural, and neural models. Its applications can be found in the sectors of medicine, finance, law, industry, and research, automating complicated processes more accurately in real time. Thus, why so many articles chose the PR field needs no further explanation. Suffice it to say that we also chose this field and analyzed it further. The chapter is primarily divided into five main sections: the Introduction, the Background, the Work Journey, the Work Constraints, the Comparative Study, and the Future Work. The Conclusion comes last.

2 Background

The process of finding patterns in data using statistical, machine learning, or deep learning techniques is known as PR (Fig. 1). It has uses in a number of industries, including computer science, engineering, and healthcare. Deep learning, machine learning, and statistical methods are often utilized techniques. Robotics, natural language processing, fraud detection, medical diagnostics, and image/audio identification are all made possible by PR. PR is still a vital technique despite issues like overfitting and partial data, and its significance is predicted to expand with the emergence of big data and rising data complexity.

The selection of algorithms has a significant impact on how well a learning system performs given the inputs. The training set and the testing set are typically created from the data. By using training rules and algorithms that define the

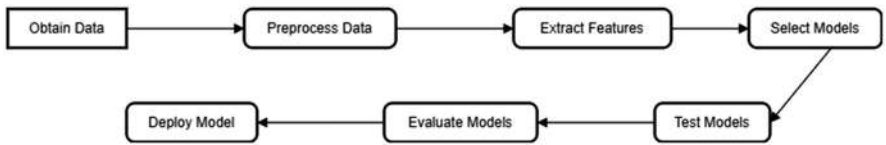


Fig. 1 PR general approach

connection between input data and output judgments, the training set is used to build the model. The system will learn from the data thanks to this procedure. On the other side, the system's performance is evaluated using the testing set after training. One can gauge a system's accuracy and precision by comparing the output to the testing data. Generally, training uses 80% of the data, whereas testing uses 20%.

3 Work Journey

The journey that started a long way back has seen much research and not in a single field. Thus, stating those research to explain the current exceptional progress is a much-appreciated choice.

3.1 Relevant Previous Studies

Abiodun et al.'s study [1] comprehensively examines the challenges faced by artificial neural networks (ANNs) in PR. It highlights the need to address issues such as object classification, orientation, scaling, and buried layer comprehension. The study emphasizes the performance of various ANN models, including GAN, SAE, DBN, RBM, RNN, CNN, and more, and suggests further development of existing models for continued advancements in PR.

In their publication, Namratha et al. [2] focus on machine learning and PR, particularly clustering methods such as K-means and hierarchical clustering. They provide a detailed overview of each technique, highlighting their steps and discussing their pros and cons. The authors offer valuable guidance for selecting the most appropriate clustering approach based on specific application requirements. This work serves as a significant resource for researchers and practitioners in machine learning and PR, advancing our understanding of clustering and its diverse applications.

A thorough survey [3] was undertaken by Antani et al. The survey focuses on the use of PR techniques for content-based access to image and video data in media archives. To build database indexes, the approaches included in the survey entail finding commonalities and removing visual data from low-level features. The study looks at the developments in this area with a focus on methods that make it possible to retrieve photographs and videos based on their content. This review offers an overview of the state-of-the-art in PR approaches for this application, making it an invaluable resource for researchers and practitioners working on content-based retrieval in media archives.

3.2 Works in Different Fields

In their paper [4], Bhattacharya et al. focus on the application of deep learning, specifically convolutional neural network (CNN) models such as rectified linear unit and Mask-CNN, for plant disease identification in smart agriculture. The study utilizes image datasets from agricultural lands to train deep learning models, enabling the detection of specific illnesses in different areas. Real-time data are collected using modified drones equipped with cameras. The technology provides regular updates to farmers, aiding in the timely identification and mitigation of plant diseases. This integration of deep learning and agriculture showcases the potential of advanced technologies to transform farming practices and enhance overall crop productivity.

Hargrove et al. [5] address the challenges faced by individuals with amputations in their daily activities and their impact on the surrounding environment. The paper proposes a surgical technique called targeted muscle reinnervation, which utilizes electromyography (EMG) PR to enhance conventional amputation surgeries. PR offers improved results compared with direct control methods, but clinical implementation faces limitations such as changes in electrode and limb positioning and force variation. The study's findings have the potential to benefit individuals through the development of prosthetic hands with myoelectric control, enabling more natural and precise movements for improved quality of life (Fig. 2).

Surface electromyography is emphasized by Zhai et al. in their paper [6] as an exciting area for neuroprosthetics. They suggest a self-recalibrating system based on a CNN classifier that continuously updates itself to maintain consistent performance and do away with the requirement for human retraining. The NinaPro database experiment fared better than a standard support vector machine (SVM)-based classifier. Yet, the database's application was hampered by restrictions in real-world situations. The results of this work could be used to create wearable medical monitors that give healthcare professionals access to real-time patient vital sign data or self-powered environmental monitoring equipment for remote locations (Fig. 3).

Similarly, some other researchers and their work are also listed for a better understanding of the journey of the PR field.

1. Kwon et al. in their paper [7] state that DNA-folded nanostructures can act as a template to display multiple binding motifs with precise and accurate spatial PR properties, which can identify viral inhibition.
2. He et al. in their paper [8] state how LYSMD3 binds chitin to the human respiratory system and evokes health issues such as asthma or allergic reactions.

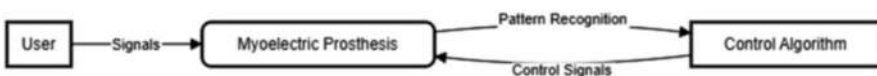


Fig. 2 Myoelectric prosthesis using PR



Fig. 3 Surface EMG Using PR

3. Saijo et al. in their paper [9] state plants with their PR receptor identify the microbes, which in terms generates two types of signals, PTI and ETI. However, microbes or pathogens using a wide array of virulence effectors overcome PTI. On the contrary, the effectiveness of ETI often leads to cellular destruction.
4. Yuan et al. in their paper [10] state how the plant immune system is affected by the PTI and ETI.
5. Uhlmann et al. in their paper [11] state the process and requirements for using updated methods like selective laser melting and its sensors for processing the industrial patterns.
6. Roshani et al. in their paper [12] state the effective way for combining the scale thickness with the flow patterns for the oil industry.
7. Seo et al. in their paper [13] state the research on artificial optical synapsis devices that behave like human pupils in identifying colors. The main problem regarding this research is recreating the biological neural network and applying it to artificial devices.
8. In the study [14] conducted by Bhattacharya et al., the authors' future prediction problem in the field of scientific research was addressed using machine learning models, with decision tree and random forest models demonstrating superior performance with an accuracy score of 1.0.
9. Parajuli et al. in their paper [15] state the advantages and disadvantages of the myoelectric prosthesis that comes with EMG PR.
10. Parrilla-Gutierrez et al. in their paper [16] state that the bottlenecks of classical Von-Neumann architecture can be overcome by the chemical process but lack of programmability holds them back.

4 Work Constraints

Till now, the motivation for the research is described in the work journey; now comes the time to dig a little deeper and analyze the work constraints acquired during the research by the researchers in Table 1.

Here, some of the restrictions are listed in a table manner. The majority of them list the absence of datasets as a limitation and mention how difficult it was to get them. Some of them explained how the research was somewhat constrained by the high computational value of the data collection. An obstacle is some described noise and unidentified elements. Nonetheless, those restrictions did not prevent the researchers from coming up with some worthwhile findings. Based on that, a comparison study was conducted in the section that follows.

Table 1 Different aspects of research

Article	Algorithm	Challenges	Results	Limitations
Classification of EEG signals based on pattern recognition approach [17]	SVM, MLP, NB, k-NN, K-fold cross-validation	Data gathering	Clinically useable	Small dataset, the difference in result
A pattern recognition algorithm for assessing trajectory completeness [18]	DTW, Bottom-up PRAVT	Lack of algorithms	Successfully performed tasks	Unknown and noise factors
Deep neural networks for pattern recognition [19]	DNN	Implementation	Enhanced the accuracy and enhancement and differentiated objects	Dataset size is large
Dynamic hand-writing analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective [20]	Mann-Whitney U Test, Relief algorithm	The lack of available dataset	Successful implementation of CAD system based dynamic handwriting method	Dataset availability
EMG pattern recognition in the era of big data and deep learning [21]	UPN, CNN, RNN	Use of unsupervised data to build the PR system	Successfully utilize the minimized dataset and elimination periodic recalibration	The high computational cost of the whole system, online methodologies associated with deep learning processes

5 Comparative Study

The study done on the research has some useful data associated. This section tries to describe those briefly.

5.1 Human Health and Medical Studies

PR is utilized in various fields to address different challenges. In the context of cancer detection, PR techniques using biological omics data and the ADTEx algorithm [22] help identify genetic changes associated with tumor cells, enabling early diagnosis and potentially improving cancer treatment outcomes. Moreover, PR receptors (PRRs) [23] can recognize changes in RNA cells, including those related to immunity and diseases. In the field of electromyotic prosthetics, PR methods using EMG (Fig. 4) [5] assist in identifying neurogenic responses for amputees. Overcoming computational and movement analysis challenges in this domain can be achieved through the application of modified machine-learning techniques combined with PR.

5.2 Plant Health and Immunity

PRRs or PR receptors are also responsible for immunogenic responses in plants. Plants depend on innate immunity [9]. The cellular damage done to it by microbial pathogenic cells triggers pattern-triggered immunity (PTI). The whole PTI system helps in natural infections where the PRR system engages with different microbial systems. Plants have two types of innate immune systems [10]: PTI and effector-triggered immunity (ETI). ETI is responsible for halting pathogen infections by directly invoking PTI systems in an immunogenic response.

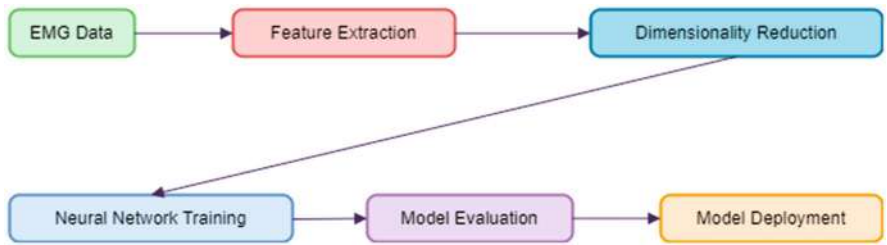


Fig. 4 EMG general PR scheme

5.3 Industrial Purpose

The industry keeps increasing in size and volume. To maintain its production with the most cost-effective methods and with advanced technologies that can affect the product quality and quantity alike, thus applying PR to identify the molecular structure for SLM [11] to identify the flow of characteristics in the oil industry [12] for better economic value. All of this uses some advanced algorithms to do the PR and they are as follows: (1) nearest neighbor, (2) Bayes classifier, (3) neural network, (4) SVM [11], (5) random forest [24], and (6) Levenberg–Marquardt algorithm used for MLP which in terms used for void friction measuring [12].

5.4 Practical Practices

The practicality of using PR is without a doubt an integral part of these papers. Vehicular trajectory models like CF [18] describe the pattern in the behavior of individual drivers. Using dynamic time warping and a bottom-up algorithm, this CF model picks up the different patterns in the different driving regimes (acceleration, deceleration, lane changing, etc.). Artificial optical synaptic device [13] uses PR methodologies to ideally detect color and color-mixed objects. It mimics human vision by using the optic-neural network. Brain-inspired parallel computing [25] mostly uses hardware-based configurations, like hardware neural networks, as conventional architectures like Von-Neumann bottlenecks restrict the processing of a large amount of data.

5.5 Research-Oriented

Ongoing research in the field of PR focuses on advancing the techniques and algorithms used in the process. Machine learning approaches, such as deep neural networks, are utilized to enhance the accuracy and refinement of PR systems [19]. ANNs, particularly CNNs, are commonly employed for tasks involving PR, especially in analyzing visual imagery [1]. PR is also being explored in the domain of 3D object detection, where algorithms are developed to detect and analyze objects in three-dimensional space [26]. Furthermore, the field of quantum PR involves the application of PR algorithms for reconstructing the trajectories of charged particles, leveraging the principles of quantum mechanics [27].

6 Applications and Future Impacts

Numerous industries use PR, including manufacturing (quality control), healthcare (medical diagnosis), finance (fraud detection), natural language processing (speech and text analysis), autonomous driving, consumer behavior analysis, document analysis, recommendation systems, and security systems. The application of PR across industries has increased thanks to developments in machine learning and deep learning.

PR has a bright future in a variety of fields. By combining many biomarkers and imaging data, PR in the medical industry can be extremely helpful in the early identification of cancer. The use of PR in healthcare will be improved even more by new algorithms and methods. The development of new rules and tactics for analyzing the stock market is made possible by PR in the financial sector. Unquestionably, further development in PR will result in considerable gains across a range of industries.

7 Conclusion

An actual data analysis technique, PR, is a subset of machine learning. We may conclude that PR has a wide range of practical applications after examining all the various papers. Fuzzy set theory, neural network field, signal processing, psychology, biological field, and other fields are involved. Image processing to capture image sensors, health systems to improve cancer treatment, plant immunity systems, manufacturing industries, finding out the quality of groundwater for reducing health risk assessment, and other fields are also involved. An early evaluation of an accurate result based on some numerical concept, some new algorithms, and some cutting-edge technology is expected from a PR model in scenarios like the selection of pharmaceuticals or the treatment process for cancer patients. After doing our research, we conclude that PR is important and will continue to receive assistance in order to help humans carry out any work quickly, accurately, and much more effectively.

Acknowledgments

Funding: We do not have any funding for this article.

Ethics approval: The authors confirm that the work described has not been published before, and it is not under consideration for publication elsewhere.

Consent to participate: The authors agree to participate in the conference.

Conflict of interest: We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted

Authors' contributions: All authors have equal contributions to the article.

References

1. Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Abubakar Malah Umar, Okafor Uchenwa Linus, Humaira Arshad, Abdullahi Aminu Kazaure, Usman Gana, and Muhammad Ubale Kiru. Comprehensive review of artificial neural network applications to pattern recognition. *IEEE Access*, 7:158820–158846, 2019.
2. M Namratha and TR Prajwala. A comprehensive overview of clustering algorithms in pattern recognition. *IOSR Journal of Computer Engineering*, 4(6):23–30, 2012.
3. Sameer Antani, Rangachar Kasturi, and Ramesh Jain. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern recognition*, 35(4):945–965, 2002.
4. Sovan Bhattacharya, Ayan Banerjee, Saikat Ray, Samik Mandal, and Debkanta Chakraborty. An advanced approach to detect plant diseases by the use of cnn based image processing. In *Innovations in Computer Science and Engineering: Proceedings of the Tenth IICSE, 2022*, pages 467–478. Springer, 2023.
5. Levi J Hargrove, Laura A Miller, Kristi Turner, and Todd A Kuiken. Myoelectric pattern recognition outperforms direct control for transhumeral amputees with targeted muscle reinnervation: a randomized clinical trial. *Scientific reports*, 7(1):1–9, 2017.
6. Xiaolong Zhai, Beth Jelfs, Rosa HM Chan, and Chung Tin. Self-recalibrating surface emg pattern recognition for neuroprosthesis control based on convolutional neural network. *Frontiers in neuroscience*, 11:379, 2017.
7. Paul S Kwon, Shaokang Ren, Seok-Joon Kwon, Megan E Kizer, Lili Kuo, Mo Xie, Dan Zhu, Feng Zhou, Fuming Zhang, Domyoung Kim, et al. Designer dna architecture offers precise and multivalent spatial pattern-recognition for viral sensing and inhibition. *Nature chemistry*, 12(1):26–35, 2020.
8. Xin He, Brad A Howard, Yang Liu, Aaron K Neumann, Liwu Li, Nidhi Menon, Tiffany Roach, Shiv D Kale, David C Samuels, Hongyan Li, et al. Lysmd3: A mammalian pattern recognition receptor for chitin. *Cell reports*, 36(3):109392, 2021.
9. Yusuke Saijo, Eliza Po-ian Loo, and Shigetaka Yasuda. Pattern recognition receptors and signaling in plant–microbe interactions. *The Plant Journal*, 93(4):592–613, 2018.
10. Minhang Yuan, Zeyu Jiang, Guozhi Bi, Kinya Nomura, Menghui Liu, Yiping Wang, Boying Cai, Jian-Min Zhou, Sheng Yang He, and Xiu-Fang Xin. Pattern-recognition receptors are required for nlr-mediated plant immunity. *Nature*, 592(7852):105–109, 2021.
11. Eckart Uhlmann, Rodrigo Pastl Pontes, Abdelhakim Laghmouchi, and Andr’e Bergmann. Intelligent pattern recognition of a slm machine process and sensor data. *Procedia Cirp*, 62:464–469, 2017.
12. Mohammadmehdi Roshani, Giang TT Phan, Peshawa Jammal Muhammad Ali, Gholam Hossein Roshani, Robert Hanus, Trung Duong, Enrico Corniani, Ehsan Nazemi, and El Mostafa Kalmoun. Evaluation of flow pattern recognition and void fraction measurement in two phase flow independent of oil pipeline’s scale layer thickness. *Alexandria Engineering Journal*, 60(1):1955–1966, 2021.
13. Seunghwan Seo, Seo-Hyeon Jo, Sungho Kim, Jaewoo Shim, Seyong Oh, Jeong-Hoon Kim, Keun Heo, Jae-Woong Choi, Changhwan Choi, Saeroonter Oh, et al. Artificial optic-neural synapse for colored and color-mixed pattern recognition. *Nature Communications*, 9(1):1–8, 2018.
14. Sovan Bhattacharya, Ayan Banerjee, Abhik Goswami, Subrata Nandi, and Dinesh Kumar Pradhan. Machine learning based approach for future prediction of authors in research academics. *SN Computer Science*, 4(3):1–11, 2023.
15. Nawadita Parajuli, Neethu Sreenivasan, Paolo Bifulco, Mario Cesarelli, Sergio Savino, Vincenzo Niola, Daniele Esposito, Tara J Hamilton, Ganesh R Naik, Upul Gunawardana, et al. Real-time emg based pattern recognition control for hand prostheses: a review on existing methods, challenges and future implementation. *Sensors*, 19(20):4596, 2019.

16. Juan Manuel Parrilla-Gutierrez, Abhishek Sharma, Soichiro Tsuda, Geoffrey JT Cooper, Gerardo Aragon-Camarasa, Kevin Donkers, and Leroy Cronin. A programmable chemical computer with memory and pattern recognition. *Nature communications*, 11(1):1–8, 2020.
17. Hafeez Ullah Amin, Wajid Mumtaz, Ahmad Rauf Subhani, Mohamad Naufal Mohamad Saad, and Aamir Saeed Malik. Classification of eeg signals based on pattern recognition approach. *Frontiers in computational neuroscience*, 11:103, 2017.
18. Anshuman Sharma, Zuduo Zheng, and Ashish Bhaskar. A pattern recognition algorithm for assessing trajectory completeness. *Transportation research part C: emerging technologies*, 96:432–457, 2018.
19. Kyongsik Yun, Alexander Huyen, and Thomas Lu. Deep neural networks for pattern recognition. *arXiv preprint arXiv:1809.09645*, 2018.
20. Donato Impedovo and Giuseppe Pirlo. Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. *IEEE reviews in biomedical engineering*, 12:209–220, 2018.
21. Angkoon Phinyomark and Erik Scheme. Emg pattern recognition in the era of big data and deep learning. *Big Data and Cognitive Computing*, 2(3):21, 2018.
22. Tingting Cheng and Xianquan Zhan. Pattern recognition for predictive, preventive, and personalized medicine in cancer. *EPMA Journal*, 8(1):51–60, 2017.
23. Megumi Tatematsu, Kenji Funami, Tsukasa Seya, and Misako Matsumoto. Extra-cellular rna sensing by pattern recognition receptors. *Journal of innate immunity*, 10(5–6):398–406, 2018.
24. Dmitry S Bulgarevich, Susumu Tsukamoto, Tadashi Kasuya, Masahiko Demura, and Makoto Watanabe. Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures. *Scientific reports*, 8(1):1–8, 2018.
25. Seunghwan Seo, Beom-Seok Kang, Je-Jun Lee, Hyo-Jun Ryu, Sungjun Kim, Hyeongjun Kim, Seyong Oh, Jaewoo Shim, Keun Heo, Saeroonter Oh, et al. Artificial van der waals hybrid synapse and its application to acoustic pattern recognition. *Nature communications*, 11(1):1–9, 2020.
26. Shilpa Rani, Kamlesh Lakhwani, and Sandeep Kumar. Three dimensional objects recognition & pattern recognition technique; related challenges: A review. *Multi-media Tools and Applications*, pages 1–44, 2022.
27. HM Gray. Quantum pattern recognition algorithms for charged particle tracking. *Philosophical Transactions of the Royal Society A*, 380(2216):20210103, 2022.

Prediction of GATE Examination Clearance for Fresh Graduate Candidates: An Advanced Machine Learning Approach



Ayan Banerjee, Rachana Das, Puja Kumari, Ankita, Syed Zahir Hasan, and Sovan Bhattacharya

Abstract The main objective of any educational institution is to provide students with the best education and knowledge possible. To do this, it is crucial to recognize the kids needing more help and take the necessary steps to raise their performance. In this chapter, we predicted how well students would perform on their GATE test on the first attempt. The linear regression machine learning (ML) model is our primary method for predicting the GATE score. We individually trained the models for each GATE CS paper's subjects. The earlier graduation marks for each topic are given special regard in this research. Now that we have analyzed our ML models for each subject based on their training and testing accuracy, we can conclude that the models for every individual subject give an accuracy of more than 90%. Hence, our approach and model can be implemented in real-time applications.

Keywords Machine learning · Linear regression · GATE score prediction · GATE CS marks prediction · GATE clearance prediction · Advanced ML approach

Rachana Das, Puja Kumari, Ankita, Syed Zahir Hasan and Sovan Bhattacharya contributed equally with all other contributors.

A. Banerjee (✉) · R. Das · P. Kumari · Ankita · S. Z. Hasan
Department of CSE, Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India
e-mail: Ayan.B@labvantage.com; zahir.hasan@bcrc.ac.in

S. Bhattacharya
Department of CSE, Data Science, Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India

Department of CSE, National Institute of Technology, Durgapur, West Bengal, India

1 Introduction

Students nowadays prefer to pursue higher education through various competitive examinations because there are more opportunities in many industries. GATE is one of the most challenging examinations to assess engineering graduates' competence. Thousands of people take the GATE examination yearly to increase their knowledge in engineering-related domains or other relevant areas and to apply for admission to the best higher education institutions for MS or M.Tech courses.

A multitude of variables is used to predict the applicant's success, deciding to admit a student to a graduate program difficult. We employed a machine learning (ML) method known as linear regression for our endeavor. Our whole approach is illustrated with the help of a block diagram in Fig. 1. Here we considered a GATE paper of code *PI*, which covers the syllabus of three subjects, *SUB 1*, *SUB 2*, and *SUB 3*. Now we collected the B.Tech score and GATE score of all three subjects from the students who have already appeared for the GATE examination once. In our case, the B.Tech score will be the input data, and the GATE score will be the label value. Now we trained three different regression models, *MODEL 1*, *MODEL 2*, and *MODEL 3* for three subjects. After training at the appropriate training ratio and data size, we out-sampled the models by giving the B.Tech score of an individual subject, and the ML model will detect the GATE score accordingly. Summing up all the particular subject's GATE scores, we can easily conclude the clearance result of a GATE-appearing candidate at the first attempt.

1.1 Motivation and Objectives

Many students seek to pass the GATE examination each year. The motivation of this research is to help them by decreasing the time they need to consider whether they can pass the examination based on their prior examination results. Students can evaluate themselves based on their previous B.Tech scores to upgrade their performance. Therefore, this can be considered as a carrier guidance for the young generation. This will help the fresh graduate understand their abilities and follow which path in their future to amend their career. Thus, our project will have an impact on society.

1.2 Challenges and Contribution

As far as we are aware, there is no perfect dataset online. Hence, we had to manually collect it through a Google Form, which took a huge amount of time. After the dataset of GATE CS appeared candidates was gathered, preprocessing was done to clean up and resolve the dataset by removing extraneous and irrelevant data. Finally, we trained the linear regression models to predict the GATE score and whether

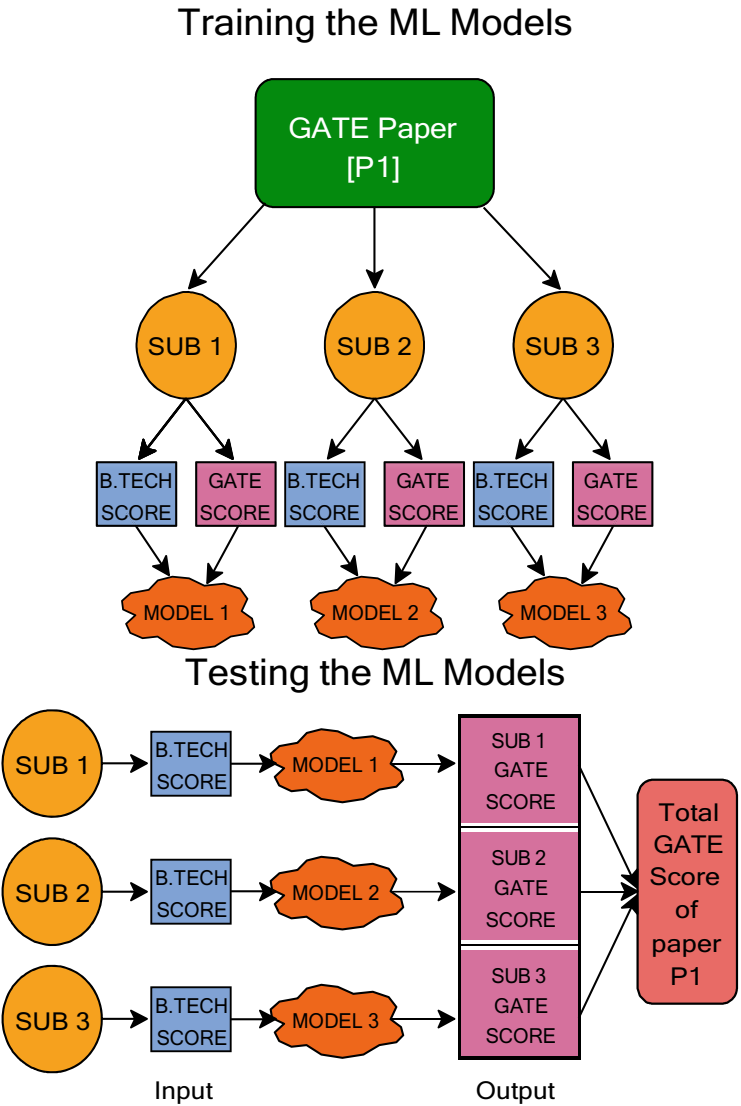


Fig. 1 Framework of our GATE clearance prediction system using ML models

cleared or not at the first attempt. We obtained training and testing accuracy for each subject, which is more than about 90%.

This chapter is organized and arranged by starting with the Introduction section, followed by the Research Review section, the Methodology section, and Experimental Setup, after which the results of the experimental setup are discussed in the Results and Discussion section, and finally finishing off with the Conclusion and Future Work section.

2 Research Review

A few existing works focused on a student's score prediction from their previous qualification result. In this section, we discussed some of those existing works. Agarwal et al. [1] suggested a model predicting students' academic achievement using neural networks, an ML technique. Anand et al. [2] proposed an ML-based algorithm that predicts students' overall performance and identifies those eligible to take examinations. Harshika et al. [3] raised awareness of the need for early at-risk student identification and preventive interventions for improving student performance. Rao et al. [4] demonstrate how challenging it is for educational institutions to raise student grades to boost employability. Arsad et al. [5] compared linear regression results with those of a neural network prediction model developed to predict students' academic progress. Mat et al. [6] gave an overview of the studies on using academic analytics by educational institutions to improve student success and retention. Kemper et al. [7] employed two ML methods, decision trees and logistic regression, to estimate student dropout using examination data from the Karlsruhe Institute of Technology. Feng et al. [8] concentrated on the assessment system's assessment component, which was utilized by more than 600 maths students in their class throughout the 2004–2005 school year. Hasmayni et al. [9] aimed to determine how well the National Examination Score of SMP can forecast high school students' academic progress in the city of Medan. Singh et al. [10] addressed utilizing data mining methods to glean relevant data from educational databases for forecasting student success rates. Zeweidy et al. [11] compared the efficacy of several data mining methods in predicting undergraduate engineering students' academic success from their high school grades. Sharma et al. [12] discuss utilizing a logistic regression model to construct a placement predictor system. Bhattacharya et al. [13] proposed different ML techniques' capabilities to predict the future of new budding researchers. Rusli et al. [14] centered on utilizing a student's demographic profile and first-semester cumulative grade point average (CGPA) as predictor factors to forecast their academic achievement in undergraduate courses. Arsad et al. [15] proposed a study that examines the CGPA growth for several matriculation and diploma student intakes. Saha et al. [16] showed that the school examination results (scores) of 1002 students are analyzed using the binary logistic regression model. Alfian et al. [17] showed the performance of undergraduate students in the University of Malaya's Faculty of Business and Accountancy and the factors affecting that performance.

From the above study, we can conclude that different ML techniques, including decision tree, logistic regression, and so on, were employed in the literature review of various researchers on students' performance prediction.

However, the existing work has yet to consider the GATE score and clearing prediction. Hence, we targeted preparing a model that will predict whether a candidate will clear the GATE examination on his first attempt using linear regression models.

3 Methodology

ML allows computer systems to learn without being explicitly programmed automatically. The steps that were taken to produce the desired outcomes for this article are covered in this section. Six essential steps—Data Collection, Data Preprocessing, Training and Testing, ML Training Model, ML Model Validation, and Deployment had been executed as illustrated in Fig. 2.

3.1 Problem Formulation

Initially, we predicted the individual subject’s GATE score with the help of our trained model. Then we formulated the individual GATE scores to obtain the total GATE score of a candidate out of 100. The formulation is given below:

$$\text{GATE} - \text{Score}_{\text{total}} = \frac{100}{N} \sum_{t=1}^N S_t$$

(1)

where N is the total number of subjects in a particular GATE paper and S_t is the t th number of subjects’ predicted GATE score.

3.2 Data Collection and Preprocessing

Using a Google Form, we gathered a manually curated dataset that includes a critical analysis of experimental and predicted data for each candidate for the previous

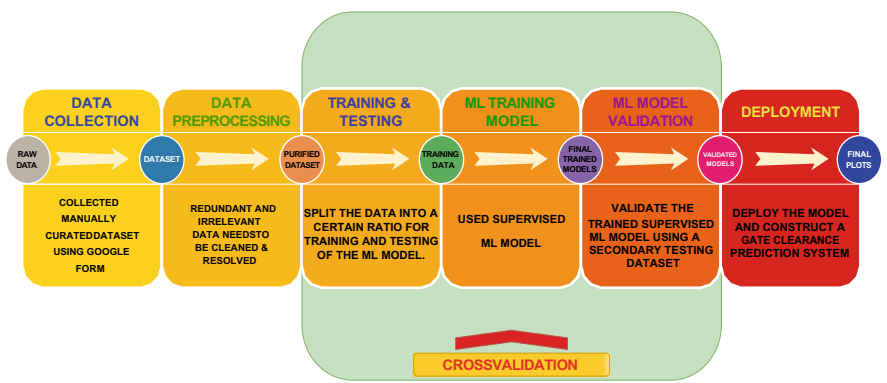


Fig. 2 Flowchart of the ML model of the GATE score prediction approach

year's GATE examination. Now the collected data need to be preprocessed for further experiments which includes two main steps:

- *Removal of Erroneous Data Elements:* Faults such as extra dots, symbols, or compatibility issues with the utf-8 format must be meticulously screened to guarantee an error-free experiment execution and trustworthy findings. It comprises identifying data errors and fixing them by changing, updating, or removing data.
- *Features Selection:* For this forecast, we selected graduate subjects that are frequent in the GATE test and collected both B.Tech scores and GATE scores from the previous year's GATE candidates to predict the likelihood of passing the GATE examination for the first time.

4 Experimental Setup

In this section, we aimed to find the best condition where the linear regression would give more precise and accurate prediction results by adapting several experimental cases. These cases are discussed below:

- *Varying Training Ratio at Fixed Data Size:* In this case, we kept the data size fixed at 500 and altered the training ratio to 80%, 70%, 60%, and 50% and obtained the testing accuracy at each training ratio. The results are discussed elaborately in Sect. 5.
- *Varying Data Size at Fixed Training Ratio:* Similarly, we fixed the training ratio at 80%, but the data size is varied to 500, 400, 300, and 200, and the testing accuracy at each data size is analyzed in Sect. 5.
- *Comparative Study:* From the literature survey, we analyzed that most of the student's performance prediction tasks considered logistic regression and decision tree ML algorithms also. Therefore, we performed a comparative study of these with our linear regression model.

4.1 Evaluation Metrics

In this chapter, we have only used accuracy evaluation metrics, which are mathematically represented as follows:

$$\text{Accuracy (A)} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

5 Results and Discussion

In our study, we employed linear regression to estimate the GATE score from the B.Tech score. We independently trained the models for each subject of the computer science paper and used them to estimate the GATE score for each subject based on their B.Tech score in that subject. Finally, we added up the results from each subject to arrive at the final score out of 100 and contrasted the obtained GATE score with the cutoff scores from the previous year.

5.1 Results Obtained at Different Training Ratio with Fixed Data Size

Here, we varied the training ratio to 80%, 70%, 60%, and 50%. The results are presented in tabular format in Table 1 and graphically represented in Fig. 3. The result table shows that the 80% training ratio experiment gives maximum testing accuracy and more than 90% in every subject’s linear regression models. Hence, we will consider the trained models with an 80% training ratio.

5.2 Results Obtained at Different Data Sizes with Fixed Training Ratio

Like the training and testing ratio, dataset size also impacts the accuracy of an ML model to a great extent. Hence, we varied data sizes to 500, 400, 300, and 200, displayed in Table 2 and illustrated in Fig. 4. On analyzing the result, we can

Table 1 Testing accuracy scores of all the subjects at a fixed data size of 500 at different training–testing ratios

Varying training and testing ratio keeping data size fixed (500)				
Training–testing ratio %	80–20	70–30	60–40	50–50
Discrete and engineering mathematics	0.93	0.91	0.90	0.89
Theory of computation	0.89	0.90	0.90	0.89
Digital logic	0.97	0.95	0.95	0.95
Computer organization and architecture	0.93	0.93	0.93	0.93
Programming and data structures	0.91	0.90	0.90	0.89
Algorithms	0.96	0.96	0.96	0.96
Compiler design	0.91	0.91	0.91	0.91
Operating system	0.87	0.84	0.83	0.83
Databases	0.95	0.95	0.95	0.95
Computer networks	0.96	0.94	0.94	0.94

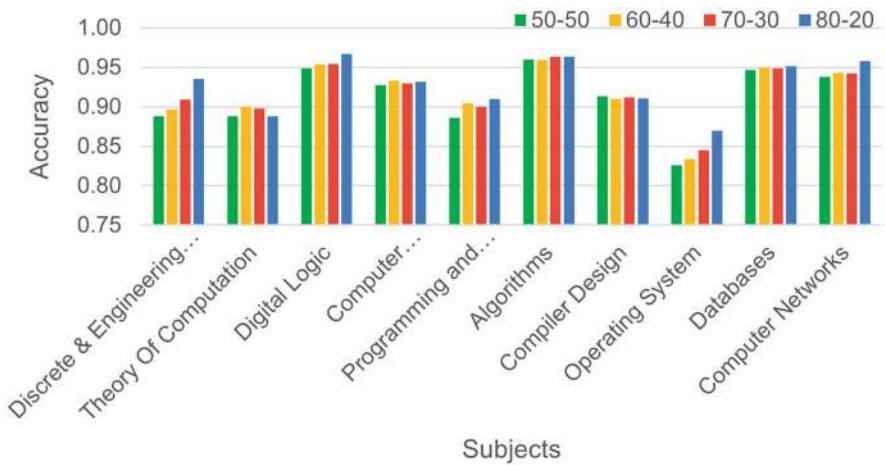


Fig. 3 Varying training–testing ratios at fixed data size

Table 2 Testing accuracy scores of all the subjects at a fixed training–testing ratio of 80–20 at different data sizes

Varying data size ratio keeping training testing ratio fixed (80–20)				
Data size	500	400	300	200
Discrete and engineering mathematics	0.93	0.89	0.87	0.82
Theory of computation	0.89	0.89	0.88	0.80
Digital logic	0.97	0.95	0.95	0.97
Computer organization and architecture	0.93	0.94	0.89	0.93
Programming and data structures	0.91	0.89	0.82	0.88
Algorithms	0.97	0.95	0.96	0.97
Compiler design	0.91	0.88	0.86	0.80
Operating system	0.87	0.81	0.64	0.77
Databases	0.95	0.94	0.93	0.90
Computer networks	0.96	0.94	0.94	0.94

perceive that a data size of 500 gives maximum accuracy which is more than 90% in all the subject’s ML models.

5.3 Result of Comparative Study

We tried to perform a comparative study of our linear regression model with *Logistic Regression* and *Decision Tree* models. However, when logistic regression is attempted, it throws an error: “ValueError: Unknown label type: ‘continuous’”. This suggests that the GATE score column in our dataset contains continuous numerical values, and logistic regression is not appropriate for predicting continuous values.

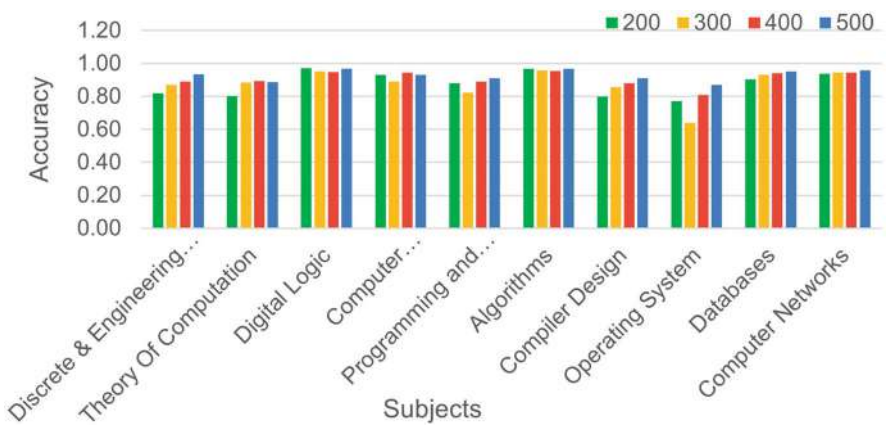


Fig. 4 Varying data size at fixed training ratio

Table 3 Accuracy scores using decision tree and linear regression models

Subject	Decision tree	Linear regression
Discrete and engineering mathematics	0.92	0.93
Theory of computation	0.85	0.89
Digital logic	0.96	0.97
Computer organization and architecture	0.93	0.93
Programming and data structures	0.90	0.91
Algorithms	0.95	0.97
Compiler design	0.91	0.91
Operating system	0.84	0.87
Databases	0.93	0.95
Computer networks	0.95	0.96

Logistic regression is typically used for binary classification problems where the output variable is categorical. Hence, logistic regression cannot be used for this prediction work.

Now, we attempted for decision tree model and obtained the testing accuracy with the same dataset, data size, and training ratio. Finally, we compared it with our proposed approach using the linear regression model. The comparative results are depicted with the help of Table 3 and also illustrated with the help of the plot in Fig. 5. The accuracy of the decision tree model is comparatively lower than the linear regression model.

Hence, from the above cases, we can conclude that the linear regression model will have the highest testing accuracy with an 80% training ratio and 500 data size, for a more accurate and precise prediction of the GATE score from the B.Tech score.

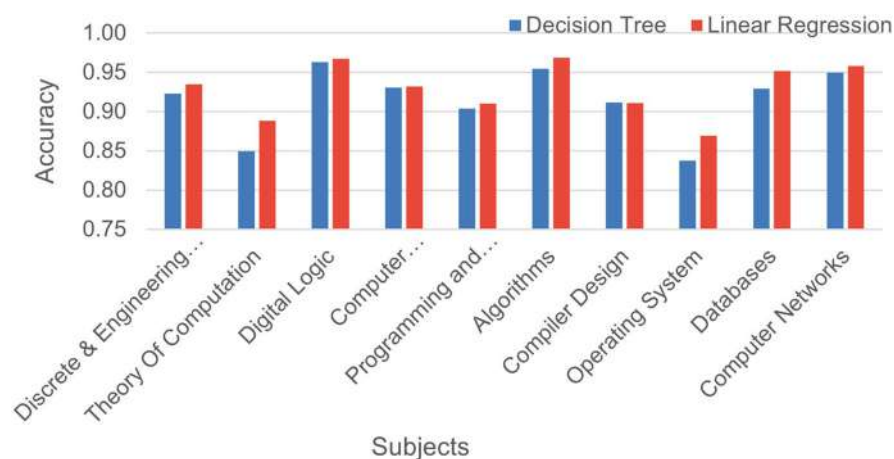


Fig. 5 Comparative study between linear regression and decision tree ML model

6 Conclusion and Future Works

We can infer from the result section that, based on the previous experiments, analysis, and results, linear regression is one of the most effective ML models for carrying out the task of the candidate's future result predictions. For more accurate answers in the future, we can train the algorithm using other GATE paper codes, similar to how we did for computer science papers. Additionally, we can train and test the dataset using different cutting-edge ML algorithms to increase accuracy. The ML model will then be implemented in a website.

Acknowledgments

Ethics Approval: The authors confirm that the work described has not been published before, and it is not under consideration for publication elsewhere.

Consent to Participate: The authors agree to participate in the conference.

Conflict of Interest: We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Agrawal, H., Mavani, H.: Student performance prediction using machine learning. *International Journal of Engineering Research and Technology* **4**(03), 111–113 (2015)
2. Anand, V., Kumar, S., Madheswari, A.N.: Students results prediction using machine learning techniques. *International Journal of Advanced Science and Applications* **3**(2), 325–329 (2016)

3. Harika, A., KA, A., A, S. Esha, Kulkarni: Student marks prediction using machine learning techniques. *International Journal of Engineering Development and Research (IJEDR)* **10** (2022)
4. Rao, G.N., Nagaraj, S., et al.: A study on the prediction of student's performance by applying straight-line regression analysis using the method of least squares. *International Journal of Computer Science Engineering* **3**(1) (2014)
5. Arsad, P.M., Buniyamin, N., et al.: Neural network and linear regression methods for prediction of students' academic achievement. In: 2014 IEEE Global Engineering Education Conference (EDUCON), pp. 916–921 (2014). IEEE
6. Bin Mat, U., Buniyamin, N., Arsad, P.M., Kassim, R.: An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention. In: 2013 IEEE 5th Conference on Engineering Education (ICEED), pp. 126–130 (2013). IEEE
7. Kemper, L., Vorhoff, G., Wigger, B.U.: Predicting student dropout: A machine learning approach. *European Journal of Higher Education* **10**(1), 28–47 (2020)
8. Feng, M., Heffernan, N.T., Koedinger, K.R.: Predicting state test scores better with intelligent tutoring systems: developing metrics to measure assistance required. In: *Intelligent Tutoring Systems: 8th International Conference, ITS 2006, Jhongli, Taiwan, June 26–30, 2006. Proceedings 8*, pp. 31–40 (2006). Springer
9. Hasmayni, B.: Prediction of junior high school national examination score on the learning achievement in high school students in medan. In: *Proceedings of the First Nommensen International Conference on Creativity & Technology, NICCT, 20–21 September 2019, Medan, North Sumatera, Indonesia* (2020)
10. Singh, W., Kaur, P.: Comparative analysis of classification techniques for predicting computer engineering students' academic performance. *International Journal of Advanced Research in Computer Science* **7**(6) (2016)
11. El Zeweidy, M., Osman, E., Elhennawy, A.P.M.E.: A comparative analysis of techniques for predicting academic performance. *Journal of the ACS* **7** (2013)
12. Sharma, A.S., Prince, S., Kapoor, S., Kumar, K.: Pps—placement prediction system using logistic regression. In: 2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), pp. 337–341 (2014). IEEE
13. Bhattacharya, S., Banerjee, A., Goswami, A., Nandi, S., Pradhan, D.K.: Machine learning based approach for future prediction of authors in research academics. *SN Computer Science* **4**(3), 1–11 (2023)
14. Rusli, N.M., Ibrahim, Z., Janor, R.M.: Predicting students' academic achievement: Comparison between logistic regression, artificial neural network, and neuro-fuzzy. In: 2008 International Symposium on Information Technology, vol. 1, pp. 1–6 (2008). IEEE
15. Arsad, P.M., Buniyamin, N., Ab Manan, J.-L., Hamzah, N.: Proposed academic students' performance prediction model: A malaysian case study. In: 2011 3rd International Congress on Engineering Education (ICEED), pp. 90–94 (2011). IEEE
16. Saha, G.: Applying logistic regression model to the examination results data. *Journal of Reliability and Statistical Studies*, 105–117 (2011)
17. Alfian, E., Othman, N.: Undergraduate students' performance: the case of university of malaya. *Quality assurance in education* **13**(4), 329–343 (2005)

Foreseeing Worker Attrition Using Machine Learning



P. LaxmiKanth , P. Maruthi Vara Prasad, S. Jitendra, and A. Yashwanth

Abstract Foreseeing Worker Attrition could be a think about that investigates the utilization of machine learning calculations to foresee worker attrition. The ponder will examine past information from a company to hunt for designs and patterns that will be utilized to decide which workers are most likely to take off. Foreseeing worker steady loss includes building a show that employments verifiable worker information to foresee the probability of a representative taking off the company. The proposed model typically incorporates a variety of factors such as job satisfaction, salary, tenure, job title, and performance metrics to identify patterns and trends that might indicate a higher risk of attrition and also machine learning algorithms such as AdaBoost, XGBoost algorithms and artificial neural network used in the study, as well as the data sources and features used to train the models. Overall, the abstract highlights the potential advantages of accurate Foreseeing Worker Attrition, such as lowering turnover costs and increasing employee retention, and gives a brief summary of the study's objectives, methods, and conclusions. It also discusses the study's potential implications for businesses interested in managing employee attrition.

Keywords Attrition · People analytics · XGBoost (optimized distributed gradient boosting library) · ANN (artificial neural network) · AdaBoost (adaptive boosting)

P. LaxmiKanth (✉) · P. M. V. Prasad · S. Jitendra · A. Yashwanth
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,
Springer Proceedings in Mathematics & Statistics 441,
https://doi.org/10.1007/978-3-031-51338-1_33

429

1 Introduction

1.1 Motivation

Organizations confront a genuine issue with representative whittling down or intentional turnover since it has an effect on their efficiency, capacity to support their workforces, and long-term development plans.

1.2 Problem Definition

Worker maintenance may be a noteworthy trouble for both companies and recruiters on this street as it comes about within the misfortune of trade plans prospects in expansion to aptitudes, encounters, and people. By making strides in the strategies of drawing in and holding ability, individual analytics helps businesses and their human assets (HR) experts in diminishing turnover. To improve decision-making, it thus represents the systematic quantification and identification of the human factors influencing business outcomes.

1.3 Scope

The research would discuss the value of employee retention in businesses and the costs associated with high levels of employee turnover. It would also give a brief overview of people analytics and its expanding role in human resources management [7].

1.4 Objective

The research's particular center is on the utilization of machine learning calculations to foresee worker steady loss. It could explain how machine learning is useful, such as its ability to detect complex patterns and relationships in data and its potential to provide more accurate predictions than traditional statistical methods.

2 Literature Survey

2.1 Prediction Employee Turnover Using Machine Learning

Punnoose and Ajit [1] predicted employee turnover in organizations utilizing machine learning calculation. By utilizing Calculated Relapse. K-nearest neighbor (KNN), Naive Bayesian = 0.64 0.59, Logistic Regression = 0.66, 0.50,[52 s]20%,[59 s] Random Forest = (Profundity controlled) 0.79, 0.51,[23 min 10 s]29%. Although random forest believes that its randomization steps will help it achieve far better generalization, the table shows that it is still insufficient to prevent over-fitting in this situation.

2.2 An Improved Random Forest Algorithm for Predicting Employee Turnover

R. Colomo-Palacios et al. [2] proposed distant better improved random forest calculation for predicting employee turnover. The random forest calculation is utilized to produce support vector machines (SVMs). The ponder presents an unused expository strategy that can offer assistance to human asset divisions to foresee employee turnover more precisely, and the test comes about giving extra bits of knowledge to diminish employee turnover deliberately. The precision of these calculations is deficient and highlights of representative aim seldom work well in these models.

2.3 Predicting Employee Attrition Utilizing Random Forest and SVN Algorithms

D. Angrave et al. [3] predicted employee attrition using random forest and SVM algorithms. The study presents a modern expository strategy that can offer assistance to human asset divisions to predict employee turnover more precisely, and the exploratory comes about giving extra bits of knowledge to decrease worker turnover deliberately. The most reduced rate of untrue positives is around 4.5%. It was missed identifying 20 employees who cleared out the company, gaining the leading review score of 0.541; the most elevated genuine positive rate of around 72%, accurately anticipating 51 out of 71 laborers who cleared out the company.

2.4 Foreseeing Employee Turnover Utilizing Machine Learning Procedures

Bhartiya et al. [4] proposed a precise stream to foresee employee turnover utilizing machine learning procedures. Python tools were used to connect the models of machine learning Naive Bayes, K-nearest neighbor, random forest, decision tree, and SVM. The approach advocated, random forest, had an 83% precision rating. Information analysis identified and reduced the majority of sources of employee mutability.

2.5 Big Data in an HR Context: Exploring Organizational Change Readiness Employee Attitudes

Shah et al. [5] utilized a normative conceptual model to examine employee behaviors and attitudes in the context of organizational change readiness. A data sample from a large public sector organization was analyzed using structural equation modeling to identify the influences of salary, job promotion, organizational loyalty, and organizational identity on employee job satisfaction. The researchers propose a framework that brings together big data principles, implementation approaches, and management commitment requirements to assess employee attitudes and behaviors as part of wider HR predictive analytics approaches [8].

3 Existing System

For attrition prediction, the system presented accuracy as the primary evaluation standard. In various datasets, various machine learning approaches such as decision tree, logistic regression, and SVM are used and evaluated. It is difficult to determine which model is best for attrition prediction. Employee attrition is always less than the employee who stays with the organization [6]. As a result, datasets are always unbalanced. As a result, an accurate model is desired to improve the prediction accuracy of the models; which results in better results for employers. Employers and HR departments are aware of their employees' behavior based on accurate prediction results [10].

The existing system has some limitations:

The predictive models training must be online as data will be dynamic and new data can be added whenever required.

Considering unbalanced data is a real challenge especially for organizations and companies with high turnover rates because the adopted predictive models are experimentally not suitable for unbalanced data in machine learning.

4 Proposed System

Foreseeing Worker Attrition is drawn nearer as an administered learning issue, particularly as a problem of binary classification. Alternatively, we need to identify and affirm the presence or nonattendance of an employee’s purpose to take off. To achieve this, we tried different supervised machine methods. We centered on the taking after classifiers specifically: artificial neural networks (ANNs), AdaBoost, and XGBoost algorithms. The models will at that point be prepared on the dataset utilizing their best setup and adaptability, including extraction for improved precision and interpretability. The proposed methodology can be seen in Fig. 1.

4.1 Data Collection

Several research studies that were compiled from employee attrition records provided the data for this essay. We place more emphasis on this stage than on selecting the subset of all pertinent data that was used.

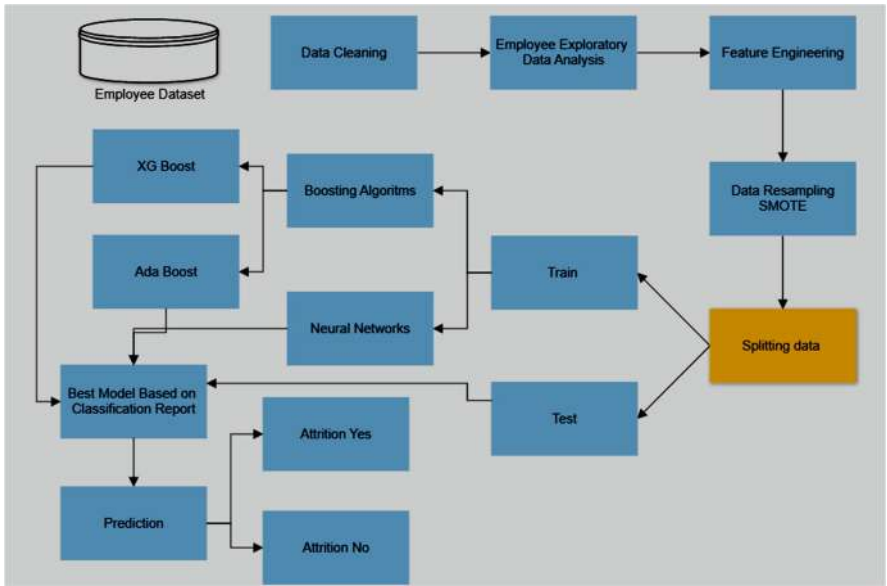


Fig. 1 Proposed methodology

4.1.1 Description of the Dataset

The employee attrition dataset is considered to evaluate the performance of the proposed framework with various feature selections and classifiers. The employee attrition dataset is taken from the Kaggle repository. Table 1 represents the dataset used in this research work [9].

- Data Source: Open Source
- Data Collected From: Kaggle

Data Source Link: <https://www.kaggle.com/datasets/patelprashant/employee-attrition>

4.2 Data Preprocessing

The information chosen is organized by organizing, cleaning, and exploring it. There are three typical information pre-management steps:

- Formatting: The data you chose might not be in a relationship that is viable for you to deal with, concurring with the course of action. The data may be in a social educational file and you would like it in a level record, or it may be in a particular report arranged and you would like it in a content report or social instructive list.
- Cleaning: The expulsion or filling in of missing data is known as cleaning data. There can be data situations that need certain pieces of information or do not pass along the information you know you really have to make a decision. These individuals have to be put to death. Also, a few of the attributes may contain delicate data that have to be anonymized or evacuated from the fabric.
- Information Cleaning Steps and Methods

There could be a six-step information cleaning process to form beyond any doubt your information is prepared to go.

 - Step1: Expel unessential information
 - Step2: Deduplicate your information
 - Step 3: Settle basic errors
 - Step 4: Deal with lost information
 - Step 5: Channel out information exceptions
 - Step 6: Approve your data
- Sampling: Clearly, there may be more carefully chosen information available than you actually wish to use. Longer evaluation times and greater processing and memory demands can be directly attributed to additional information. Before taking into account the complete dataset, you can conduct a very thorough delegation preliminary of the selected data, which may be much faster for researching and prototyping techniques.

Table 1 IBM Employee Attrition dataset

Sl. No.	Feature name
1	Attrition
2	Department
3	Employee number
4	Daily rate
5	Age
6	Distance from home
7	Work-life balance
8	Education field
9	Job satisfaction
10	Business travel
11	Environment satisfaction
12	Hourly rate
13	Stock option level
14	Standard hours
15	Years in current role
16	Job role
17	Employee count
18	Marital status
19	Job level
20	Relationship satisfaction
21	Education
22	Percentage salary hike
23	Education
24	Over18
25	Performance rating
26	Monthly income
27	Job involvement
28	Total working years
29	Gender
30	Training time last year
31	Overtime
32	Years at organization
33	Monthly rate
34	Number of organizations worked
35	Years with current manager

4.3 Feature Engineering

Feature engineering is an important part of learning models. The inspiration is to speed up the information change handle and increment the exactness of the learning show. In our inquiry about ponder, we connected highlight designing methods to handle the highlights of HR worker records to discover the highlights that best fit our

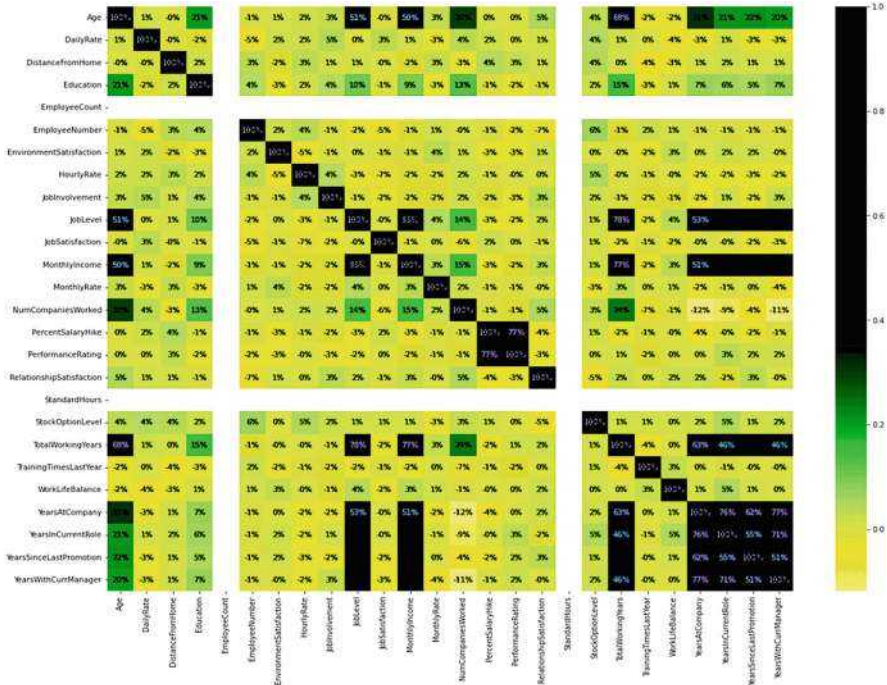


Fig. 2 Correlation matrix of features

learning demonstration. We performed an information relationship investigation to discover the leading highlights, as appeared in Fig. 2 (correlation matrix of names).

Through examination, we have taken the following characteristics: Employee Count, Distance From Home, Employee Number, Job Involvement, Environment Satisfaction, Hourly Rate, Job Satisfaction Training Times Last Year, Monthly Rate, Standard Hours, Stock Option Level, Daily Rate, Work-Life Balance, Relationship Satisfaction, and Standard Hours. Overlooked highlights are due to tall negative relationships among the remaining dataset highlights.

Chosen highlights were encoded utilizing the name encoding procedure. Achieving high accuracy scores have proven exceptionally useful in investigating thinks to realize tall exactness scores.

4.4 Resampling

Figure 3 represents the SMOTE used in Foreseeing Worker Attrition to address the problem of imbalanced datasets. In this case, the minority class would be employees who leave the company (attrition), and the majority class would be employees who stay. An imbalanced dataset in Foreseeing Worker Attrition can be problematic as

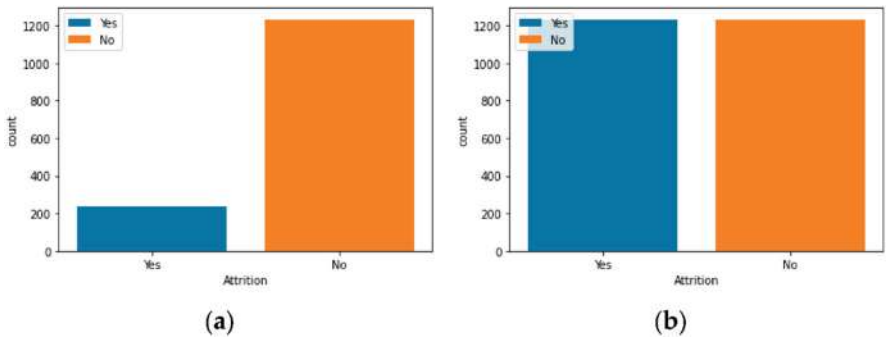


Fig. 3 The resampling of the dataset utilizing the SMOTE strategy was applied: (a) the unequal target category and (b) the target category after adjusting the dataset

a classifier trained on such data may be biased toward predicting the majority class (employees who stay) and may perform poorly in predicting the minority class (employees who leave). SMOTE can be used to generate synthetic samples of the minority class, making the dataset more balanced and improving the classifier’s ability to predict employee attrition accurately. However, it is essential to keep in mind that SMOTE may present manufactured tests that do not represent the genuine fundamental distribution of the minority course. As such, it is vital to carefully assess the execution of the classifier and maintain a strategic distance from overfitting to the engineered tests created by SMOTE.

4.5 Dataset Splitting

The dataset spitting reduces the complexity and overfitting of the show. We integrated the dataset component to create our demonstration generalization. The ratio used to perform the part was 75:25. The 25% fragment of the dataset was used for show testing assessments, and the remaining 75% fragment was used for our suggested demonstration preparation purpose.

5 Proposed Machine Learning

5.1 Artificial Neural Networks

ANN is a machine learning algorithm that is used for prediction tasks. An ANN consists of interconnected nodes (also called neurons) that process and transmit information. It was inspired by the design and operation of biological neural networks.

Fig. 4 Single-input neuron

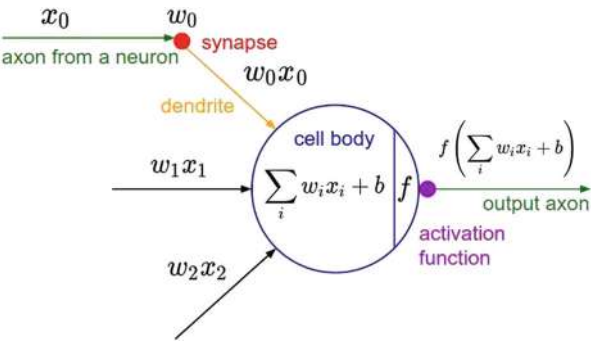
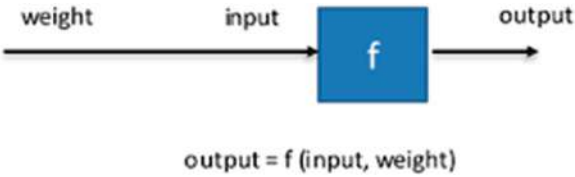


Fig. 5 Multilayer perceptron

ANN Classifier

The organic neuron employed for prediction formed the basis for neural network—network of neurons. Let us see if we can get a single neuron to learn. Figure 4 depicts a single neuron with a single input.

When a neuron has many inputs, as shown in Fig. 5, the MLP is made up of inputs that are weighted and connected to the layer. The neuron then takes many inputs and forms a multilayer perceptron as a result. The diagram depicts a multilayer experience.

$$OT = (\varepsilon_1\omega_1 + \varepsilon_2\omega_2 + \dots\dots + \varepsilon_k\omega_k) + \Theta$$

(1)

- where OT is the output;
- σ is the sigmoid function or transformed function;
 - ξ is the input to the neuron;
 - ω is the weight of input (1 to k);
 - Θ is the bias.

5.2 AdaBoost

A machine learning technique known as AdaBoost, or adaptive boosting, can be applied to situations involving binary categorization, such as forecasting worker attrition. An ensemble method called AdaBoost combines a number of weak learners to produce a strong learner.

In the context of Foreseeing Worker Attrition, AdaBoost works by combining multiple decision trees, where each decision tree is a weak learner that predicts whether an employee is likely to leave the company or not based on a subset of the available features. During the training process, AdaBoost assigns higher weights to misclassified samples in each iteration, so that the subsequent decision trees focus more on the difficult samples. The final prediction is made by taking a weighted average of the predictions from all the decision trees.

One of the advantages of AdaBoost is that it is less prone to overfitting than some other machine learning algorithms because it assigns higher weights to misclassified samples and focuses more on the difficult samples. Additionally, AdaBoost can handle imbalanced datasets, which are common in Foreseeing Worker Attrition.

5.3 *XGBoost*

The machine learning technique XGBoost is frequently employed to address classification and regression issues. A powerful model is produced using an ensemble method that combines different decision trees. A variation of the gradient boosting algorithm, XGBoost, stands for extreme gradient boosting.

In the Foreseeing Worker Attrition study, XGBoost can be used to train a model that predicts whether an employee is likely to leave the company based on a set of input characteristics. The algorithm works by training the decision tree iteratively, and each new tree is designed to improve the overall accuracy of the model. Trees are added to each trained iteration to reduce errors caused by previous trees. The projections of all trees are averaged to provide the final prediction. The ability of XGBoost to manage missing data and outliers is one of its important characteristics and also allows for feature importance estimation, which can be useful in determining which features are the most important for foreseeing worker attrition. Additionally, XGBoost can handle imbalanced datasets, which are common in Foreseeing Worker Attrition [11].

6 Experimental Results

The complete flow of applications is designed by using Python programming language in the Jupiter notebook. All of the tests were conducted on a device that details Intel® Core™ i5-1135G7 @ 2.40GHz, 8GB RAM, 2419 MHz, 4 Core(s), 8 Consistent Processor(s).

As we all know, Python language contains several modules, third-party packages, and a lot of modules that are required for constructing the application. To execute the program, one should first install the necessary packages which are required to execute. Once the packages are installed, now we can foresee worker attrition. The outcomes of predicting worker attrition were closely examined. The preparation

accuracy, testing precision, accuracy score, review score, and f1 measure score are some of our research study's assessment measures. The key elements of assessment measurements are as follows:

True positive (TP): the value where the actual numbers are higher than the projected ones.

True negative (TN): value where genuine values are also negative, as expected values are.

False positive (FP): value where although the true value is actually negative, both approaches anticipate a value to be positive.

False negative (FN): value where although both approaches anticipate a value to be negative, the actual value is actually positive.

The accuracy score was calculated when our model was being developed and tested. We discovered that our best XGBoost achieved an 88% exactness score on concealed information, demonstrating how accurate our machine learning demonstration is at producing and testing information. At that moment, we generalized our demonstration. The equation condition is conveyed in Eq. (1) in order to determine the precision score:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

The degree to which the display correctly identifies values as being positive among all values is referred to as its precision. The model's recall measures how well it can distinguish between real positive values. The recall ratings were 95% for Adaboost, 99% for XGBoost, and 68% for ANN, and the precision scores were 89% for Adaboost, 88% for XGBoost, and 90% for ANN. Equations (2) and (3), respectively, provide the mathematical notations for calculating the precision and recall:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

The execution of our vision demonstration was evaluated using the F1 score. The precision and recall score values are combined to create the F1 score. Adaboost's F1 score was 92%, XGBoost's F1 score was 93%, and ANN's F1 score was 78%. The F1 score was determined using Eq. (4):

$$F1 - \text{score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

Table 2 represents the accuracy of different proposed models.

Table 2 Accuracy of different models

Algorithm	Accuracy
XGBoost	88%
AdaBoost	86%
ANNs	87%

XGBoost Model (Fig. 6)

```
import xgboost as xgb

# Convert the data to DMatrix format for XGBoost
dtrain = xgb.DMatrix(X_train, label=y_train)
dtest = xgb.DMatrix(X_test, label=y_test)

# Define the parameters for the XGBoost classifier
params = {
    'max_depth': 3,
    'eta': 0.1,
    'objective': 'binary:logistic',
    'eval_metric': 'error'
}

# Train the XGBoost classifier on the training data for 50 rounds
num_rounds = 50
xgb_model = xgb.train(params, dtrain, num_rounds)

# Predict the labels of the test data
y_pred = xgb_model.predict(dtest)
y_pred = [round(value) for value in y_pred]

# Calculate the accuracy of the XGBoost classifier
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)

Accuracy: 0.8777173913043478
```

Fig. 6 Sample code for XGBoost model

Classification Report (Fig. 7)

```
# Print the classification report
print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	0.67	0.12	0.21	48
1	0.88	0.99	0.93	320
accuracy			0.88	368
macro avg	0.77	0.56	0.57	368
weighted avg	0.85	0.88	0.84	368

Fig. 7 Evaluation metric results for the proposed model

Prediction (Fig. 8)

Fig. 8 Sample code for prediction

```
In [79]: a=X.iloc[[3]]
         predictions =xgb_reg.predict(a)
         if predictions <=0.5:
             print("Not Attrition")
         else:
             print("Attrition")

Attrition
```

7 Conclusion

Foreseeing Worker Attrition is an important task for organizations as it can help them identify potential problem areas and take proactive measures to retain their employees. Machine learning models can be used to analyze historical data and identify factors that are strongly associated with attrition. By leveraging these insights, organizations can develop targeted retention strategies to improve employee satisfaction and reduce turnover rates. However, it is important to note that no model can predict employee attrition with 100% accuracy and that there may be other unmeasured factors that contribute to turnover.

8 Future Enhancements

In future enhancements, incorporate natural language processing (NLP) techniques to analyze employee feedback and sentiment data. By integrating NLP, the model could better understand the underlying reasons behind employee attrition and provide more targeted recommendations for improving employee retention. Additionally, the model could be enhanced with a real-time dashboard that provides stakeholders with up-to-date insights on employee attrition and retention trends, enabling them to take proactive measures to address potential issues before they escalate.

Declaration

1. All authors do not have any conflict of interest.
2. This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. “Forecasting employee attrition in organizations using artificial intelligence algorithms,” by R. Punnoose and P. Ajit Global Journal of Advanced Research in AI, 5, no. 9, p. 5, 2016, doi: 10.14569/IJARAI.2016.050904.
2. “Intention to leave career among software professionals,” by R. Colomo-Palacios, C. Casado-Lumbreras, S. Misra, and P. Soto-Acosta. Human Factors and Ergonomics in Manufacturing and Service Industries, Vol.24, No.6, Nov. 2014, pp.641–655, doi: <https://doi.org/10.1002/hfm.20509>.
3. “Analytics in HR: Why HR is likely to fail the challenge of big data,” by D. Angrave, A. Charlwood, I. Kirkpatrick, M. Lawrence, and M. Stuart Journal of Human Resource Management, vol. 26, no. 1, p. 1–11, Jan. 2016, doi: <https://doi.org/10.1111/1748-8583.12090>.
4. [Amazon.fr](https://www.amazon.fr) —Jean Paul Isson, Livres, People Analytics in the age of big Data: Revolutionizing the way you Attract, Recruit, Develop, and Retain Talent. The date: December 15, 2019. [Online]. There is a link to the page below:<https://www.amazon.fr/People-Analytics-Era-Big-Data/dp/1119050782>.
5. An Evaluation of Conceptual Boundaries and Benefits of People Analytics, International Journal of Information Management, volume 43, pages 224–247, December 2018.
6. T. Pape, “Data Item Prioritization for Business Analytics: Framework and its Application to Human Resources,” European Journal of Operational Research, volume 252, issue 2, July 2016, pages 687–698.
7. “Human Resource Predictive Analytics (HRPA) for Enterprise HR Management,” International Journal of Science and Technology Research, volume 5, issue 5, pages 33–35, 2016, authored by S.N. Mishra, D.R. Lama, and Y. Pal.
8. “HR Value Proposition through Predictive Analytics: A Comprehensive Overview,” in New Paradigm in Decision Science and Management, by P. Likhitkar and P. Verma. 2020, Singapore: Springer, doi:https://doi.org/10.1007/978-981-13-9330-3_15, pages 165–171.
9. T. Peeters, J. Paauwe, and K. Van De Voorde, “Creating a Framework for Effective People Analytics,” Journal of Organizational Effectiveness, People, and Performance, volume 7, issue 2, July 2020, pages 203–219, doi:<https://doi.org/10.1108/JOEPP-04-2020-0071>.
10. Investigating the readiness of organizations for transformation, as well as the attitudes and behaviors of employees, in the context of HR using big data, N. Shah, Z. Irani, and A. M. Sharif, J. Bus. Res., vol. 70, pp. 366–378, Jan. 2017, doi: <https://doi.org/10.1016/j.jbusres.2016.08.010>.
11. “Using big-deep-smart data in imaging to guide materials design,” S. V. Kalinin, B. G. Sumpter, and R. K. Archibald Nature Materials, Oct. 2015, vol. 14, no. 10, p. 973–980, doi:<https://doi.org/10.1038/nmat4395>.

Mouse Controlling Using Eyeball Action



S. Kranthi Reddy, D. Shivananda Reddy, B. Suresh, and B. Pavan Kumar

Abstract There are certain persons who cannot utilize computers due to medical conditions. The design of eyeball controlling is extremely beneficial, especially for those with disabilities and natural inputs in the future. Furthermore, putting the control system in place enables that, without the aid of a second person, they could operate the computers. This system is highly useful for physically handicapped people who will control the mouse cursor with their heads and eyes. The camera is used in this project to capture the image and video of eyeball movement. First, it determines the pupil's middle spot of the eye. Then, a change in pupil spot results in a different action in the cursor, and then the pupil detection process is carried out using the free and open-source OpenCV library for image processing and computer vision. It can be used to identify objects, faces, and other features in images and videos. In this process, we connect to the webcam and command the mouse cursor to move in accordance with eyeball movement, after which each image from the camera is extracted. Then transmit it to the library called OpenCV to find the exact location of the eye center position. After detecting the center location of the eye, we can extract the coordinates (x and y) of the eyeball from a Python library named OpenCV and we can direct the cursor to move to the eyeball's specified x and y coordinates.

Keywords Eyeball movement · Open CV · Mouse · Computer · Camera

1 Introduction

Personal computers are becoming an integral part of our daily lives because we utilize them for business, education, and recreation. All of these applications share the fact that keyboard and mouse input is the primary method used while using a

S. K. Reddy (✉) · D. S. Reddy · B. Suresh · B. P. Kumar
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,
Springer Proceedings in Mathematics & Statistics 441,
https://doi.org/10.1007/978-3-031-51338-1_34

445

personal computer. Although it is not an issue for people who are physically strong, it can be an impassable barricade for those with restricted limb range of motion of the eye. In these circumstances, the use of process would be a preferred method that relies on the brain region's stronger capabilities, like eyeball movements. To make such alternate input methods possible, a system that employs a low-cost technique to handle a mouse pointer on a screen was developed. The eye tracker uses images taken by a modified webcam to track the user's eye movements. These eye movements are then graphed on the computer screen to determine where the mouse pointer should be placed. By automatically changing the mouse movement and the position of the eyes, eye movement is captured via a camera. There are normally three phases to every algorithm used to process digital images: input, processor, and output. The image is captured by a camera during the input stage. It was sent to a certain system, and that system's output was a processed image.

A combination of software and hardware makes up an existing system. An existing system may be a standalone unit or a material of a huge system. An embedded system is a system that is built on a microcontroller or microprocessor and is designed to perform a specific function. For example, a fire fear is a built-in device that only detects smoke. Python is a high-level programming language. This indicates that Python source code is written in generally recognizable English, giving the Pi instructions in a simple and fast-to-learn manner. This is not highlighted. In contrast, middle-level languages, such as assembly, are more similar to how a machine thinks but are nearly impossible for a person to understand without training. The significance of human-machine interaction is attractive and evident as computer automation advances quickly. Some persons with disorders are not able to access computers with their hands. People with the condition primarily employ eyeball movement control. By integrating this eye-controlling technology with computers, it will be possible for them to function independently of human assistance. The human-computer interface (HCI) field focuses on using machine technology to create a human-computer interaction (HCI). Finding the best technology to enable efficient communication between people and computers is necessary. The crucial function is played by HCI. To allow persons with disorders the same opportunity to be involved in the information organization, it is necessary to update a mechanism that disseminates an alternative means of communication between persons and machines. HCI has stimulated the attraction of researchers all over the world in recent days. The approach machine for eyeball movement observation for impaired persons is implemented using HCIs. The suggested system for managing an inconspicuous HCI includes eye detection, eye tracking, face tracking, face detection, mouth detection, mouth tracking, and real-time analysis of a series of head moves and eye blinks. Human eye movements have taken the place of the traditional mouse-based way of computer interaction. This process will make it simple and highly efficient for those who are paralyzed, physically impaired, or not having hands to operate. First, the web camera takes the picture and then, using OpenCV source code for pupil-of-eye identification, focuses on the eyeball in the picture. The center of the human pupil is the outcome of this. The user or human

will operate the cursor by moving left and right using the pupil center location as an illusion point.

The goal of HCI is to create intuitive, user-friendly, and aesthetically pleasing interfaces. Keep HCI principles in mind when creating the interface for your eye movement analysis project, such as uniformity, simplification, and efficient use of visual components. Make sure that the interface is simple to use and offers consumers clear instructions. To communicate with computer systems, HCI investigates several input modalities. You might think about using other input modalities, such as mouse, keyboard, or touch-based interactions, in this project in addition to eye movement. This can give consumers freedom and allow for the accommodation of people with various needs or preferences.

2 Literature Review

This project explains how we looked into current systems and used them as a model for our own system. The face is taken by the MATLAB vfm program as part of the system used in [1]. By separating the face into three equal parts and locating the top third, the eyes are located. The tracking of eye movements is made possible by referencing the iris corners and computing the iris shift. An experimental web camera-based eye-tracker method designed specifically for low-level powered devices is described in this study [2]. It comprises five processes: background suppression, a face detection method based on the haar cascade features, geometric estimation of the eye position, monitoring the eyeball center using the gradient vector's mean, and user gaze detection. This research [3] proposes a book eyeball motion analysis model found on five eye characteristic positions. Convolutional neural networks are trained for the detection of eye quality points in this study. A dataset with about 0.5 million settings from 38 patients was created. This machine [4] demonstrates the implementation of a trade eye tracker as a pointer that just requires the user to move their eyes as a peripheral device. The tests were carried out to gauge the individuals' comfort level and willingness. Using a consumer-grade depth sensor, the system illustrated in [5] performs gaze direction estimates from human eye movement. Preprocessing is done on the photos to get rid of extraneous details like background objects. The user's head location is determined via the suggested method using depth pictures.

3 Related Work

The OpenCV tool picture-handling mechanism is worn to recognize eyeballs. The cursor can move based on eyeball movement. Controlled closing of the eye is used to find out the eye aspect ratio (EAR) in order to carry out various actions such as scrolling, clicking, and selecting operations.

In the current machine, a MATLAB tool is used to detect the eyeball and control the computer. An eyeball motion-controlled wheelchair is currently operated by observing the motion of the eyes. Because predicting the center of the eye in the MATLAB tool is difficult, we are using OpenCV. The drawbacks of the current system are having less precision and that it cannot use a regular camera [6–13].

4 Proposed System

In this project, the mouse cursor movement is handled by eye movement via OpenCV. The web camera detects eye movement, which OpenCV then processes. This process can be used to move the cursor. The person must view in front of a computer screen, with a specialized video camera set up above the screen to examine the user’s eyes and face. The computer is constantly analyzing the focus video photo. To grab a key, the person must stare at it for a specific amount of time, whereas to “press” any key, the person must simply blink the eyes. The advantages of the proposed system are having excellent precision and that people with physical disabilities can use computers.

In this project, they mainly predict eyeball movement. At first, we want to identify the face landmark position. By the use of landmarks, we can easily identify the movements of the eye. We can detect eyeball movement and eye blinking in a video and also predict feelings. Understanding the Dlib’s facial landmark finder: Dlib’s model allows us to estimate the 68 2D face landmarking accurately in Fig. 1 clearly, and the 68 facial landmarks picture is shown in Fig. 1.

When determining eyeball motions, only the eyes are examined. As seen in the image, the eye size is designated by six coordinates (x,y) in Fig. 2, initiating the far toward the left side and continuing clockwise toward the right side, encompassing the balancing part of the eye region.

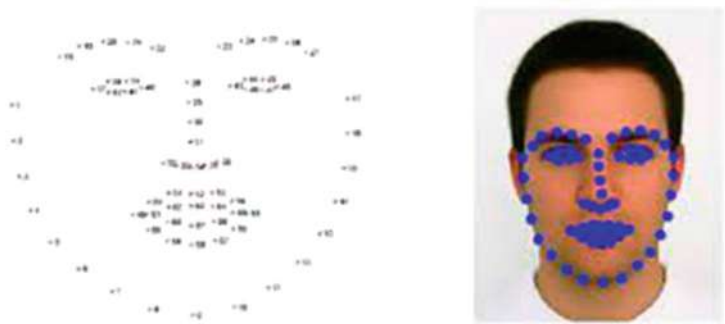


Fig. 1 68 face marker positions

Fig. 2 Six face points associated with the eye

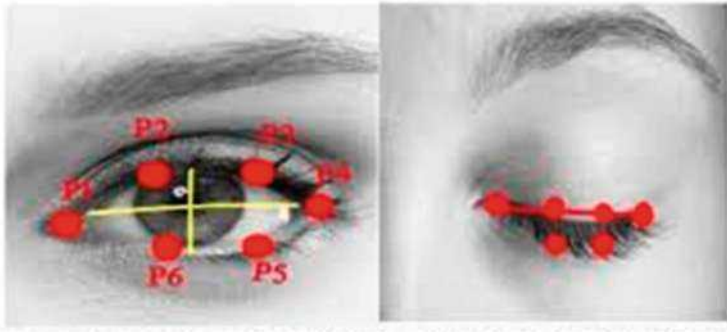
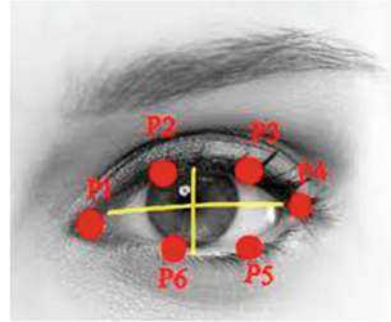


Fig. 3 Points of eye (open eye, closed eye)

After long research is done in real time of eye observation using face landmarks, an equation that makes certain the coordination between all six face coordinates known as EAR would be derived as follows:

$$EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2 ||p1 - p4||} \quad (1)$$

Points p1 through p6 in the equation above represent 2D face landmark positions. The gap between the eye's vertical points is substituted for the numerator of the equation, and the gap between the eye's horizontal points is substituted for the denominator. The EAR is nearly the same when the eyes are widely open, but it drops to null when the user blinks the eyes.

If the eye is completely open, then the EAR is greater and leftovers are the same over a long period, as shown in Fig. 3 (left side). The EAR drops dramatically and approaches null when a user blinks. The EAR remains the same over a period of time and approaches null gradually. The number then increases to indicate that the individual in question is blinking.

Eyes are only deemed open if the minimum EAR value is satisfied; otherwise, they are regarded as closed. If necessary, you might want to change this setting

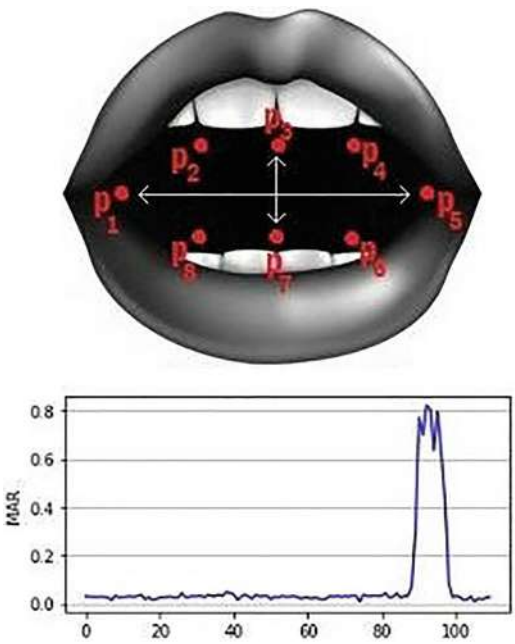
to suit your requirements. After attempting to locate EAR in a range of different circumstances, decide on the value. It is critical to remember that this EAR refers to the combined EAR for both eyes, not simply the EAR of one eye. EAR’s value is susceptible to abrupt changes. Even if you blink twice and close your eyes, the EAR will dramatically drop. However, blinking indicates that one should always flag their drowsiness. A person would be considered groggy if they had closed one or both of their eyelids for a duration of at least 10 video frames, indicating that their ear was very low. Therefore, the use of this fixed value is the high number of settings in a sequence where the EAR equation may remain below MINIMUM EAR without setting off a drowsiness warning.

$$\text{MAR} = \frac{||p2 - p8|| + ||p3 - p7|| + ||p4 - p6||}{2 ||p1 - p5||} \tag{2}$$

There are a total of 20 xy coordinates in the mouth aspect ratio (MAR) algorithm, as shown in Fig. 4. The MAR is used to determine if the mouth is fully open (or) fully closed, which is similar to the EAR.

The preceding algorithm depicts the complete procedure for controlling a cursor with ocular movement using a Raspberry Pi tool and Python OpenCV tool. The Raspberry Pi tool is at the heart of the processing block, which tracks facial expressions using an Internet Protocol camera. After waiting for the eyes to be recognized, the camera captures the image. The photo handling method of OpenCV

Fig. 4 Mouth aspect ratio when the mouth is open



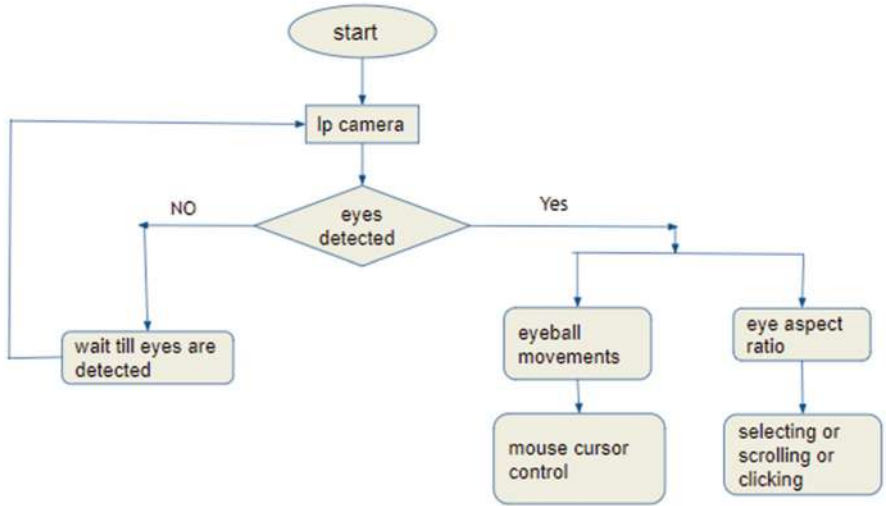


Fig. 5 Flowchart of mouse controlling using eyeball

technology is used to distinguish between eyes. Eyeball movement can be used to control the mouse pointer, and action of the eyes is used to calculate the specific EAR to do various activities such as click operation, select operation, and scrolling as shown in Fig. 5.

5 Results and Discussion

Change the current working directory before launching the Python source code, and then launch the Python code by entering the instructions shown in Fig. 6 clearly. After typing the commands in the terminal, we get a frame window, and the camera begins capturing the face that reads the input. When the user opens their lips, the input begins to be read, and if the user wants to handle the mouse cursor upward, the Anchor point should be moved up.

If the user wants to handle the mouse cursor downward, the Anchor point should be moved down as shown in Fig. 7.

If the user wants to handle the mouse cursor right side, the Anchor point should be moved right as shown in Fig. 8.

If the user wants to handle the mouse cursor left side, the Anchor point should be moved left as shown in Fig. 9.

If the user wants to handle the scrolling mode, the user needs to minimize and maximize the eyes as shown in Fig. 10.

Fig. 6 Cursor moving upward



Fig. 7 Cursor moving downward

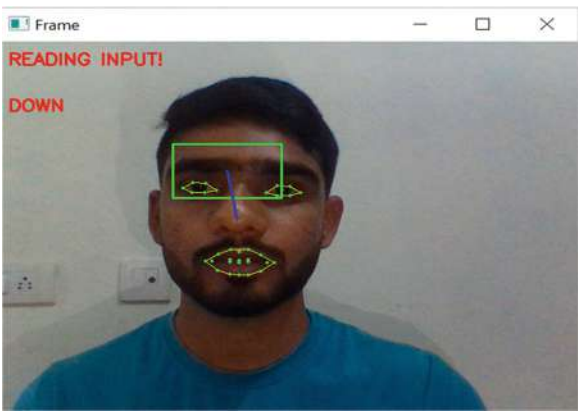
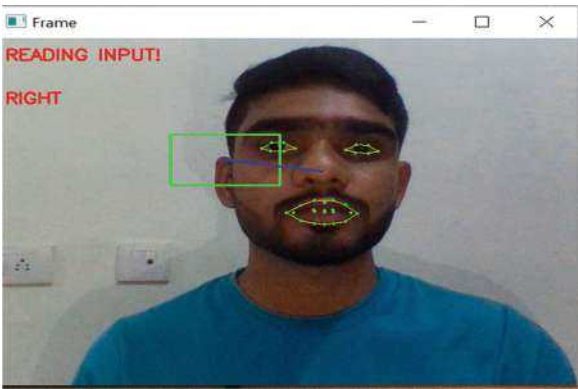


Fig. 8 Cursor moving right



Detection failure is the major criterion used to calculate the efficiency of the proposed system. Images can be obtained by searching the BioID(library) database.

Fig. 9 Cursor moving left

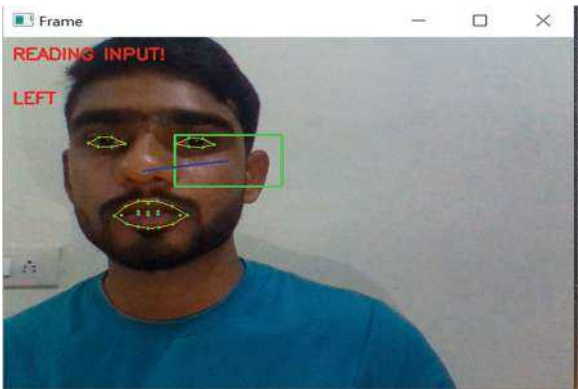


Fig. 10 Scrolling mode activated

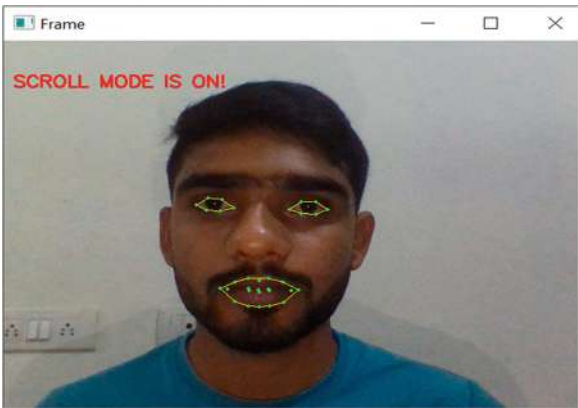
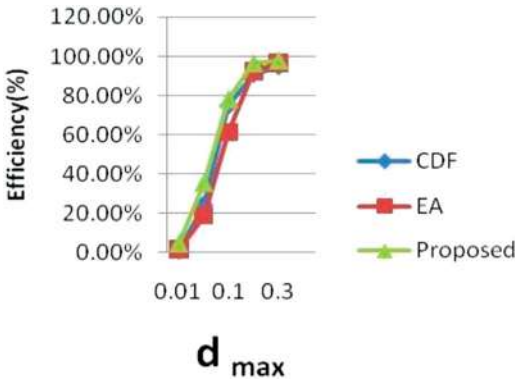


Fig. 11 Analysis of efficiency



In terms of accuracy, the suggested method performs better than those currently in use. Figure 11 shows the efficiency in various error levels.

6 Conclusion and Future Scope

The process clearly conveys that the computer cursor can be handled by ocular movement rather than using hands and fingers on the system. This project will be useful for humans who have difficulty controlling cursor points on computers using the machine's physical components. Because the eyeballs can be used to move the cursor, people with disabilities can use computers without the assistance of others. In the future, this technology can be improved by incorporating new approaches such as clicking events, mouse motions, and human interface machines that use eye-blinking operations.

In addition to showing the possibility of natural HCI, the implementation of mouse control via ocular action is also true to the fundamentals of HCI. This project contributes to the development of technology that is user-centric, inclusive, and ethically conscious by focusing on usability, user experience, accessibility, personalization, error handling, privacy, and interface design. This technology has the potential to provide more seamless and natural interactions between users and computers with additional research and development. This technology has increased due to ocular motion and eye blinking to achieve a high degree of efficiency and accuracy in eye movement compared with existing systems.

References

1. Geetha and Mangaiyarkarasi, ["Cursor Control System Using Face Expressions for Human-Computer-Interaction"], Volume 8 Issue 1-4-2014, ISSN: 0976-1353 (IJET) International Journal of Emerging Technology in CSE, pp 30–34.
2. Caceres, Enrique, Miguel Carasco, and Sebastian Rios. ["Evaluation of an eye-pointer interaction device for (HIC)human-computer interaction"] Heleyon 4.3 (2018): E00574.
3. Cech and Soukupova, ["Real-Time Eye Blink Detection using Facial Landmarks"], Center for System Perception, Feb 2016.
4. EniChul Lee-Kang Ryoung Park ["A robust eye gaze tracking method based on a virtual eyeball model"], Springer, pp.319–337, April (2008).
5. Florina Ungreanu, Robert Gabrel Lupu and Valentin Siriteanu, ["Eye tracking mouse for human-computer-interaction"], IEEE Conference Publication, aug 2013.
6. Hegde, Veena , Ramya ,S. Ullagaddimath, and S. Kumuda. ["Low cost eye based human-computer interface system (Eye controlling mouse)."] India Conference (INDICON), 2016 Annual IEEE, 2016.
7. Lee, Jun Seok, Kyung hwa Yu, Sang won Leigh, Jin Yong Chung, and Sung Goo Cho. ["Method for controlling device on the basis of eyeball motion, and device there for."] U.S. Patent 9,864,429, issued Jan 9, 2018.
8. Nasor, Mohamed, et al. ["Eye controlled mouse cursor for physically disabled individuals."] Advances in Science and Eng Technology International Conferences (ASETIC), feb 2018. IEEE, 2018.
9. Neil Castelino and Michele Alva, ["An image based eye controlled assistive system for paralytic patients"], IEEE Conference Publications of 2017.
10. Prof. Prashant Salunkhi, Ashwini R Patil [" Device Controlling Using Eye Movement"], International Conference of Electrical, Electronics and Optimization Techniques, (ICEEOT) 2016.

11. Sure Pumrin and Chaerat kraichan, ["Face and eye tracking for controlling computer functions"], iee Conference Publications, Mar 2014.
12. Samuel Epstein Eric Misimer Margrit-Betke ["Using Kernels for avideobased mouse-replacement interface"], Springer link, (2012).
13. Valenti, Robirto, Nicu Sibe, and Theo Gevirs. ["Combining head-pose and eye location information for gaze estimation."] IEEE Transaction on Image Processing 21..2 (2012): 802–815.

Power Quality Improvement by Using Shunt Hybrid Active Power Filter



D. V. Kiran, G. Neetha, G. Gowtham, K. Anusha, K. Ravali,
and A. Bharath Kumar

Abstract To alleviate the power quality problem with the alternating current (ac) mains in an ac–direct current (dc) power supply delivering to a nonlinear load, this project report describes the design, modeling, and development of a passive shunt filter and a shunt hybrid power filter (SHPF). The power filter consists of an active power filter and a passive shunt filter connected in series. The first step was to design a passive filter capable of eliminating harmonic distortion. By addressing concerns like fixed compensating features and the resonance problem, SHPF is an effort to improve upon passive filters. Simulations of a typical distribution system with an SHPF were performed to validate the study results presented. To demonstrate how the harmonic extraction circuit influences the SHPF's harmonic compensation characteristic, the harmonic contents of the source current have been calculated and compared for different scenarios. This was done to demonstrate the effect of the circuit on the harmonic contents. Shunt passive filter, SHPF, harmonic correction, and modeling are some of the keywords here.

Keywords Power quality · Active power · THD · SHPF

1 Introduction

These days, machinery that is driven by power electronics may be found in a variety of locations, including factories as well as private homes. These pieces of equipment have considerable influence on the quality of the voltage that is provided and have led to an increase in the pollution produced by harmonic current in distribution systems. Additionally, these pieces of equipment have contributed to a rise in the amount of pollution generated by harmonic current. They have a wide range of unfavorable effects on the power system's equipment and on the customers,

D. V. Kiran (✉) · G. Neetha · G. Gowtham · K. Anusha · K. Ravali · A. B. Kumar
Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering
and Technology, Chittoor, India

including increased losses in overhead and underground cables, transformers, and rotating electric machines; operational issues with protection systems; overvoltage and shunt capacitors; errors in measuring instruments; and malfunction of low efficiency of customer sensitive loads. Low measurement accuracy, broken or inaccurate measuring devices, and inefficient customer-sensitive loads all contribute to these outcomes.

Passive filters have a long history of use in industrial power systems for the aim of lowering the amount of distortion brought on by harmonic current. However, they also have a few disadvantages, such as the resonance problem, the fact that their performance is dependent on the impedance of the system, and the fact that they absorb harmonic current from nonlinear loads, which can cause further harmonic propagation in the power system [1]. These drawbacks could be overcome, however, by designing the system so that it has a higher impedance than the harmonics that are being absorbed.

Active power filters (APFs) are a solution that was created in an effort to find a solution to this problem. It does not suffer from any of the drawbacks that would be associated with a passive filter. To do this, they will inject harmonic voltage or current into the system at the appropriate amplitude and phase angle. This will result in the cancellation of the harmonics that are created by nonlinear loads. However, it does have a few drawbacks, such as a high initial cost and high-power losses, both of which prevent it from having a wide application. These drawbacks together prevent it from having a wide application. This is particularly true for systems that have a high-power rating in their specifications [2].

The implications of these limits have led to the development of hybrid power filters, which have been used in a wide range of real-world system applications [3–7]. This was done in order to offset the impacts of these constraints. The active filter, passive filter, and three-phase PWM inverter make up the shunt hybrid filter. The shunt hybrid filter is therefore formed. The shunt hybrid filter also incorporates an active filter. Both passive and active filters present difficulties, but this filter effectively mitigates the impacts of both. Harmonic correction is one of the many services it provides, and it is particularly useful for high-power nonlinear loads [4] because it is both effective and cost-efficient. There are a few various control techniques that might be employed in order to get rid of the harmonic components that are present in the source current. One of these options could be to use a filter. Two of the many methodologies that are included in this category are the synchronous reference frame (SRF) transformation and the instantaneous power (p-q) theory. Both of these are instances of the SRF transformation. The use of high-pass filters or the separation of harmonic source current components from fundamental components may both be used in these procedures [5].

In this thesis, a shunt hybrid power filter, often commonly referred to as an SHPF, is represented by using an “a-b-c” reference frame that is kept static throughout the process. Following this step, the model is transformed into the rotating “d-q” reference frame so that the level of control complexity can be decreased. To accomplish the desired effect of having the filter’s current automatically follow its reference value, two distinct ways of decoupled current control are used. In

the course of the execution of these methods, one after the other, a controller of the proportional–integral (PI) type and a controller of the hysteresis type are each utilized. On the contrary, a PI controller is used to manage the dc voltage that is being generated by the filter. This is done so in order to prevent any unwanted side effects. This is done in order to guarantee the highest possible level of performance. It is possible to control the harmonic current that is drawn from the alternating current (ac) mains by passing the harmonic current that is drawn from the nonlinear load via the passive filter. As a result of this, the harmonic currents that are drawn from the ac mains are eliminated. When a shunt active filter and an LC passive filter are used together, the necessary rating for the active filter is much lower than what it would be if it were used on its own as a stand-alone component. This is because the passive filter is acting as a load on the active filter. As the LC circuit is able to filter out the switching ripple on its own, a switching ripple filter is not necessary in this scenario because it is capable of doing so. As a direct consequence of this, the installation of the filter is not required.

The results of the simulation are analyzed, and it is determined that the SHPF is effective in reducing and softening the effects of harmonics.

1.1 Objective

- The planning and development of a passive filter for the ac–direct current (dc) supply system that will be used to power a nonlinear load.
- The development of an SHPF for an ac–dc supply system with a nonlinear load to be fed by it.
- In the MAT LAB and SIMU LINK environment, simulate both of the different approaches.
- Carry out the comparison research of the two different strategies based on the response of the simulation.

The power quality at the ac mains is negatively impacted as a result of the nonlinear characteristics shown by the diode rectifier interface, which is connected to the electric utility. The objective of this effort is to find a solution to the harmonic issue that has been affecting the ac mains.

In the course of this investigation, a number of different approaches, including passive filter, active filter, hybrid filter, and six pulses PWM converter, are put into practice.

1.2 Importance of Power Quality

- The metrics that reflect reactive power, harmonic pollution, and load imbalance are what determine power quality.

- A sinusoidal voltage waveform that is constant in amplitude and frequency would be the optimal shape for an electrical supply since it is the most efficient. However, the reality is frequently quite different because the supply system has a non-zero impedance. There is a possibility of coming across a broad range of loads, in addition to other occurrences, such as transients and outages. All of these factors contribute to the fact that the reality is frequently quite different.
- Any load that is linked to the network will run smoothly and effectively if the power quality of the network is excellent. Both the expense of installation and the carbon footprint it leaves behind will be modest.
- If the power quality of the network is bad, then loads that are attached to it will either cease functioning altogether or will have a shortened lifetime, and the efficiency of the electrical infrastructure will drop as a result.

The start-up costs, operating costs, and carbon footprint will all be quite high, and it is highly probable that the business will not even be profitable when those costs are included.

2 Literature Survey

In today's world, there are a great deal of problems with power quality, many of which have gotten so serious that they affect a great deal of nonlinear apparatus. As a result of the widespread use of these systems, they are becoming more polluted, and the environment as a whole has grown increasingly sensitive. These occurrences take place because there is an excessive quantity of harmonics present in the system, and these harmonics are the root cause of the unintended power losses that occur in electrical equipment. By removing harmonics, the power quality issue may be mitigated in a number of different ways using the approaches that have been established.

Because of their ease of use and lower cost, passive filters are often a viable option in circumstances like these. However, this filter suffers from a number of shortcomings, the most notable of which are its fixed compensation, its cumbersome components, and its difficulty in avoiding the resonance issue that plagues L-C filters. Because of this, something called an APF was invented to completely compensate for distortions.

We have decided to go with the shunt hybrid filter architecture, which is a mix of a passive filter with parallel connections and an active filter with a low rating [2, 3]. The effectiveness of this configuration's functioning has been taken into consideration throughout this article. There is a wide variety of control methods that may be used to get the desired outcomes. However, the SRF method is one of the methods that is the most conventional and can be applied in a practical setting [4–9]. A phase-locked loop (PLL) circuit is required for the synchronization to be done by the SRF. A control system of the SRF technique with the modified PLL has been provided in the study, despite the fact that the performance of traditional PLL is subpar. This is because ordinary PLLs have such poor performance.

3 Methodology and Proposed Work

The configuration of the proposed SHPF-TCR compensator is depicted in Fig. 1. The SHPF is a fifth-tuned passive filter connected in series with a small rating APF. The APF has a three-phase full-bridge pulse width modulated voltage source inverter with an input boost inductor (L_{pf} , R_{pf}) and a dc bus capacitor (C_{dc}). The combination of the passive filter connected in parallel with TCR forms a shunt passive filter. The objective of a tuned passive filter is for harmonic compensation and power factor correction, and the objective of the active filter is to isolate the system impedance from the filter impedance to avoid resonance. The active filter also participates in suppressing harmonics due to the load and TCR by enhancing the compensation characteristics of the shunt passive filter. Though SHPF suppresses the harmonics effectively, reactive power cannot be compensated due to the passive filter. Hence, to regulate reactive power, TCR is preferred. The nonlinear loads considered are three-phase diode rectifier with RL load, three-phase diode rectifier with RC load, and three-phase voltage controller.

3.1 Modeling and Control Strategy

3.1.1 Modeling of SHPF

The system equations are first written in 123 reference frames. Using Kirchhoff's voltage law,

$$v_{s1} = L_{PF} \frac{di_{c1}}{dt} + R_{PF} i_{c1} + v_{CPF1} + v_{1M} + v_{MN} \quad (1)$$

$$v_{s2} = L_{PF} \frac{di_{c2}}{dt} + R_{PF} i_{c2} + v_{CPF2} + v_{2M} + v_{MN} \quad (2)$$

$$v_{s3} = L_{PF} \frac{di_{c3}}{dt} + R_{PF} i_{c3} + v_{CPF3} + v_{3M} + v_{MN} \quad (3)$$

$$v_{CPF1} = L_T \frac{di_{c1}}{dt} - C_{PF} L_T \frac{d^2 V_{CPF1}}{dt^2} \quad (4)$$

$$v_{CPF2} = L_T \frac{di_{c2}}{dt} - C_{PF} L_T \frac{d^2 V_{CPF2}}{dt^2} \quad (5)$$

$$v_{CPF3} = L_T \frac{di_{c3}}{dt} - C_{PF} L_T \frac{d^2 V_{CPF3}}{dt^2} \quad (6)$$

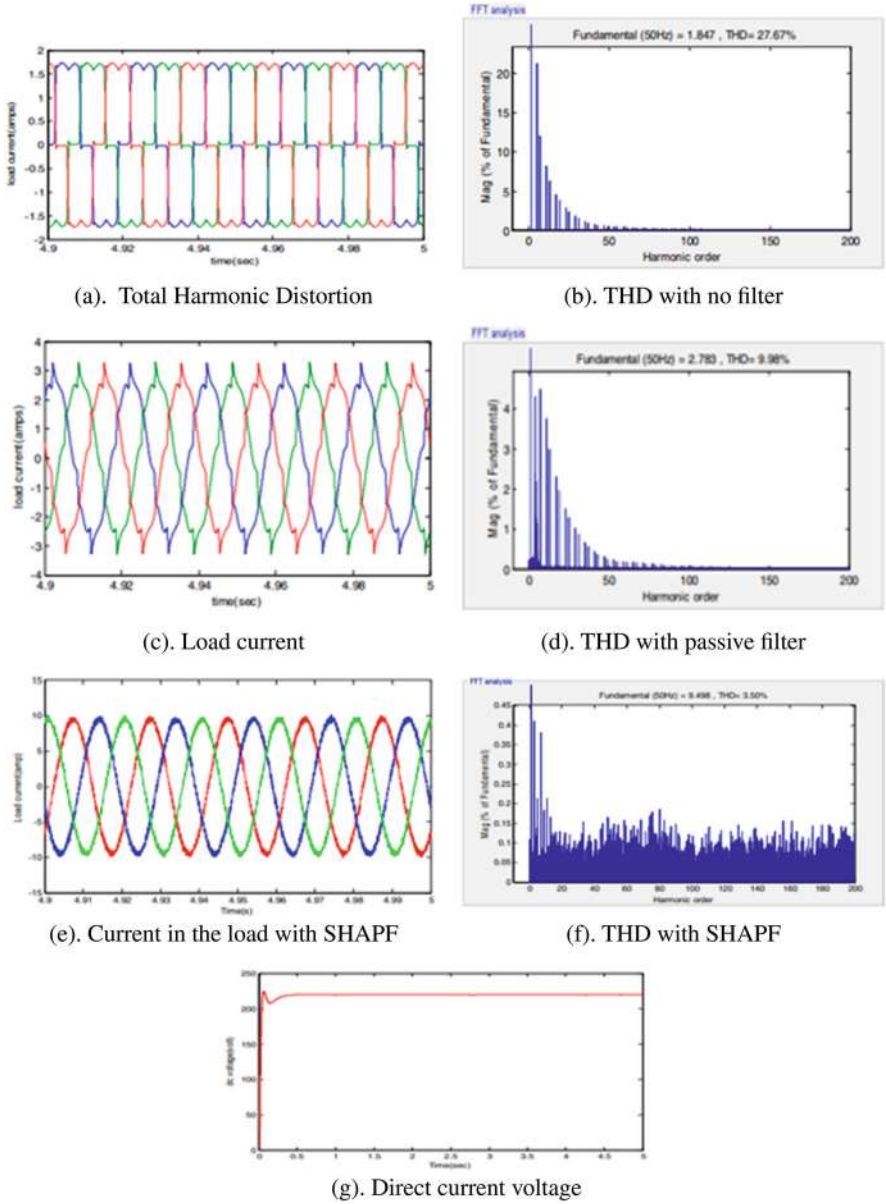


Fig. 1 Optimal source balance load current (a), total harmonic distortion (THD) (b) with no filter, load current (c), THD (d) with passive filter, (e) current in the load with SHAPF, (f) THD with SHAPF, and (g) direct current voltage

Table 1 System specification parameters

L-L source voltage, and frequency	$V_{s\text{ L-L}} = 440\text{ V}, f_s = 60\text{ Hz}$
Line impedance	$L_s = 0.5\text{ mH}, R_s = 0.1\text{ }\Omega$
Diode Rectifier with RL load	$L_{L1} = 10\text{ mH}, R_{L1} = 40\text{ }\Omega$
Diode Rectifier with RC load	$R_{L2} = 26\text{ mH}, C_{L2} = 15\text{ }\mu\text{F}$
Voltage controller	$L_{L1} = 20\text{ mH}, R_{L1} = 27\text{ }\Omega$
Passive filter parameters	$L_{PF} = 2.2\text{ mH}, C_{PF} = 120\text{ }\mu\text{H}, R_{PF} = 0.55$
Active filter parameters	$C_{dc} = 3000\text{ }\mu\text{F}, R_{dc} = 1000\text{ }\Omega$
DC bus voltage of APF of SHPF	$V_{dc} = 50\text{ V}$
Switching frequency	1920 Hz
Inner controller parameters	$K_{P1} = 43.38, K_{P2} = 43.38: K_{i1} = 374, K_{i2} = 374$
Outer controller parameters	$K_1 = 0.26, K_2 = 42$
Cut off frequency of the low pass filters	$F_C = 1000\text{ Hz}$
TCR inductance	$L_T = 25\text{ mH}$
TCR controller parameters	$K_P = 0.98, K_I = 15$

4 Simulation

Using the power system block set in MATLAB-Simulink for a three-phase four-wire power network with a shunt active and shunt passive filter, the simulation results for both balanced and unbalanced source situations of a power system network may be created. Table 1 outlines the SHAPF parameters as they are now configured to be used.

4.1 Balanced Source Condition

The suggested controller effectively compensates for current harmonics when a balanced source and load are assumed. From $t = 4.9$ to $t = 5$ seconds is taken from the simulation in a steady state with a constant load. Figure 1a and b show the extremely nonlinear load current and its FFT spectrum, respectively.

Figure 1c shows that even after attaching the passive filter, nonlinearity and distortion still occur, even if the lower-order harmonics are decreased. Figure 1e depicts the load current after the active filter complete compensation with the PI controller has been implemented. Figure 1f provides conclusive evidence that the THD dropped below 5% in accordance with the IEEE standard [10].

4.2 Unbalanced Source Condition

Figure 2a is a depiction of the source voltage, which demonstrates that it is entirely unbalanced for the unbalanced source. This can be seen by looking at the graph. The

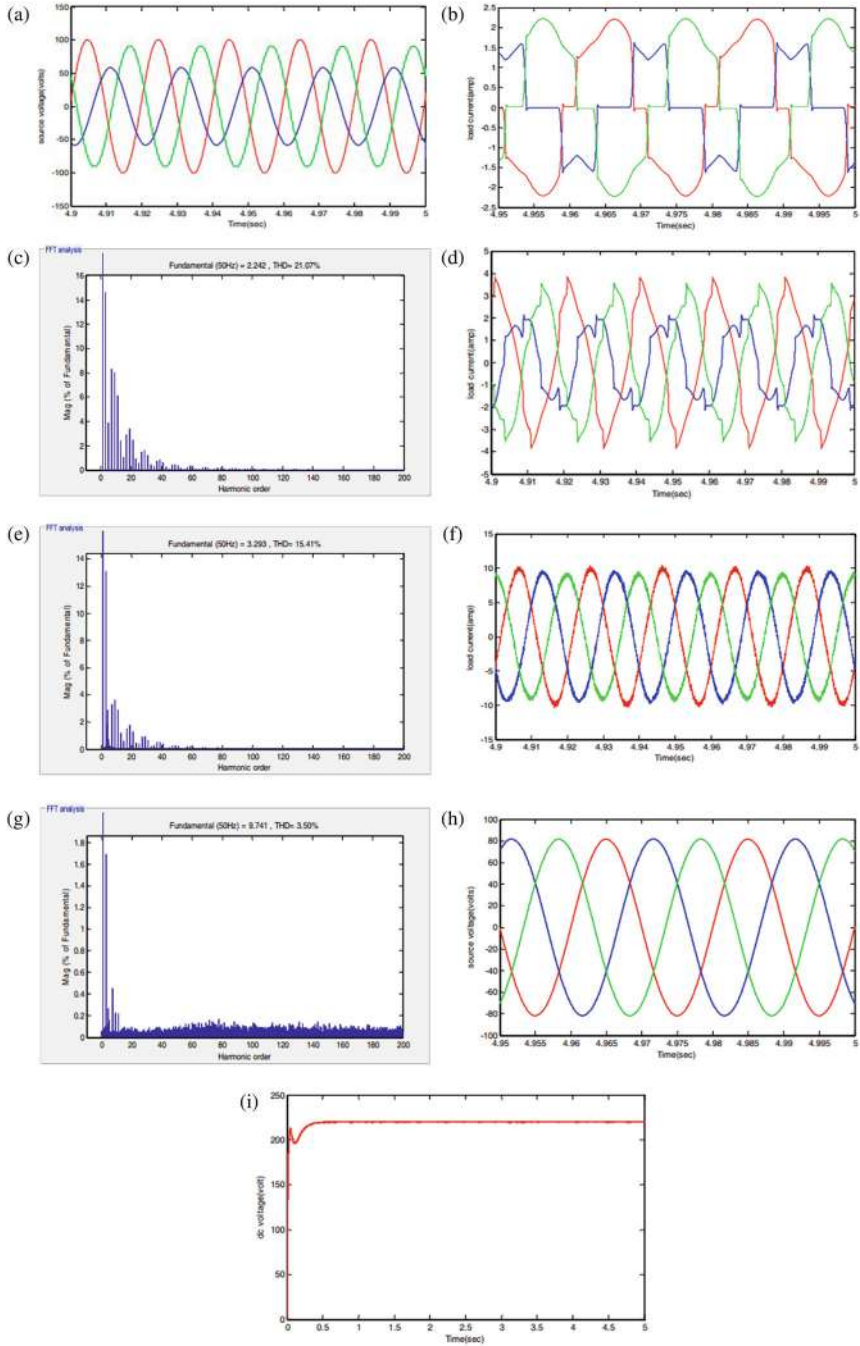


Fig. 2 Condition of an unbalanced source. (a) Unbalanced source voltage, (b) load current without any filter, (c) THD without any filter, (d) the load current with a passive filter, (e) the THD with a passive filter, (f) the load current with SHAPF, (g) the THD with SHAPF, and (h) the source voltage became balanced with SHAPF. (i) DC voltage

entirely nonlinear load current may be observed in Fig. 2b, which depicts the current in the load. As a consequence of doing the FFT analysis, the THD was computed, and its result can be shown in Fig. 2c. As can be observed in Fig. 2d, the nature of the load current has started to become linear after the passive filter has been attached, and Fig. 2e demonstrates that the THD has begun to manifest after the passive filter has been connected. There is a representation of both of these figures included inside the figure. Figure 2f displays the load current that was measured after the SHAPF full compensation with the PI controller had been installed. You can get this information on the Internet. The FFT analysis that was done after connecting SHAPF can be seen in Fig. 2g, which displays the appropriate THD, which is less than 5%. This analysis was completed successfully. As can be seen in Fig. 2h, as soon as SHAPF was attached, the source voltage immediately achieved a condition in which it was completely in equilibrium. Fig. 2i displays the dc voltage.

5 Conclusion

As a result of the work done in this study, a control strategy for SHAPF has been developed to improve power quality. The method that was used to build the control technique was the SRF method. To achieve effective grid voltage synchronization in both balanced and unbalanced scenarios, a modified PLL was designed and put into use. The THD was determined to be 3.50% under both balanced and unbalanced source circumstances. After applying SHAPF to both of the different sources, it was discovered that the load current is linear. Consequently, SHAPF, which is based on PI and a modified SRF theory approach, is a credible harmonic reducer due to its fast response and high-quality filtering. This is because SHAPF has an SRF theory that has been modified.

References

1. J. C. Das, "Passive filters; potentialities and limitations," *IEEE Trans. on Industry Applications*, Vol. 40, pp. 232–241, 2004.
2. B. Singh, K. Al-Haddad, and A. Chandra, "Are view of active filters for power quality improvement," *IEEE Trans. Ind. Electron.*, vol. 46, no.5, pp.960–971, Oct. 1999.
3. L. Asiminoaei, E. Aeloiza, P. N. Enjeti, and F. Blaabjerg, "Shunt active-power-filter topology based on parallel interleaved inverters," *IEEE Trans. Ind. Electron.* vol. 55, no.3, pp. 1175–1189, Mar. 2008.
4. H. Fujita, H. Akagi, "A practical approach to harmonic compensation in power systems; series connection of passive and active filters," *IEEE Trans. on Industry Applications*, Vol.27, pp. 1020–1025, 1991.
5. L. Asiminoaei, E. Aeloiza, P. N. Enjeti, and F. Blaabjerg, "Shunt active-power-filter topology based on parallel interleaved inverters," *IEEE Trans. Ind. Electron.* vol. 55, no.3, pp. 1175–1189, Mar. 2008.

6. B. Singh, V. Verma, A. Chandra, K. Al-Haddad, "Hybrid filters for power quality improvement," *IEEE Proc. on Generation, Transmission and Distribution*, Vol.152, pp. 365–378, 2005.
7. Salem Rahmani, Abdelhamid Hamadi, Nassar Mendalek, and Kamal Al-Haddad, "A New Control Technique for Three-Phase Shunt Hybrid Power Filter," *IEEE Transactions on industrial electronics*, vol.56, no.8, pp. 606–805, august 2009.
8. B. Singh, K. Al-Haddad, and A. Chandra, "Are view of filters for power quality improvement," *IEEE Trans. On Industrial Electronics*, Vol.46, pp.960–971, 1999.
9. S. Rahmani, K. Al-Haddad, and H. Y. Kanaan, "A comparative study of two PWM techniques for single-phase shunt active power filter sem-ploying direct current control strategy," *J. IET Proc.—Elect. Power Appl.*, Vol.1, no.3, pp. 376–385, Sep.2008.
10. N. Mendalek, K. Al-Haddad, L. A. Dessaint, and F. Fnaiech, "Nonlinear control technique to enhance dynamic performance of a shunt active power filter," *Proc. Inst. Elect. Eng.—Elect. Power Appl.*, vol. 150, no.4, pp.373–379, Jul.2003.

Integration of Renewable Energy Systems Into Utility Grid: A Review on Power Quality Issues, Mitigating Devices, and Control Algorithms



Joddumahanthi Vijaychandra, Santi Behera, and Lingraj Dora

Abstract Integration of renewable energy systems (RESs) into the utility grid is the trend in electrical power systems distribution networks. It is a fact that RESs are the major contributors to generating electricity and meeting the power demand. The main reasons for this integration include an increase in energy demand, environmental impacts due to the emission of carbon gases, extinguishment of fossil fuels, and so on. In fact, several power quality (PQ) issues will emerge after integrating into the utility grid. This study aims to provide various PQ issues and their mitigation methods of grid-integrated RESs.

Keywords Control algorithms · DFACTS devices · Power quality · Renewable energy · Soft computing · Solar energy

1 Introduction

The efforts have been tremendous across the globe in the development of RESs and their associated technologies from the past decades to meet the energy demand of every end user. If the RESs are installed at appropriate geographical locations, then the necessity for fossil fuels gradually gets diminished, thereby resulting in an increase in the sustainability of power supply across the globe. When there is a proper integration of RESs with conventional energy sources, the reliability of the electrical power system will be improved in the end. There has been a huge rise in the emergence of RES technologies over the past three decades. It is because these sources have the following advantages: (a) the atmosphere cannot be polluted; (b) these can be available in huge quantities; and (c) these can be well suited for

J. Vijaychandra (✉) · S. Behera · L. Dora

Department of Electrical and Electronics Engineering, Veer Surendra Sai University of Technology, Burla, Odisha, India

e-mail: ldora_eee@vssut.ac.in

Table 1 Applications of RESs [1]

Types of RESs	Common applications
Solar	Electricity generation, water heating, solar farms
Wind	Wind pumps, wind battery chargers, wind electricity generators
Hydro	Electricity generation, control of floods, irrigation
Geothermal	Heat pumps, urban heating
Biomass	Electricity generation, gasification, and pyrolysis

decentralized use [1]. The various RESs are solar, wind, hydro, geothermal, and biomass. The applications of these sources are tabulated below.

Among all the RESs, solar energy systems (SESs) have gained their own importance due to several advantages that include abundance availability (Table 1) of sunlight across the globe, zero emission of harmful gases, and so on [2]. The SESs have become popular in meeting the power demand of remote areas in which a conventional plant fails to serve due to several factors primarily concerned with economic feasibility [3]. Integration of SES into the grid will result in several PQ issues such as power fluctuations, voltage swells and sags, harmonics, and so on.

Power quality (PQ) is a major concern in any power system network. It is generally referred to as the ability of an electrical power system network to create a perfect supply of power that is free from other disturbances in the sinusoidal wave shape [4]. Analyzing and assessing the PQ issues is always a challenging task while integrating SES into the grid. As renewable energy systems (RESs) possess a highly intermittent nature, especially SESs, in turn, it will result in huge fluctuations in power. Hence, while integrating into the main grid, several PQ issues will emerge in the system [5], which in turn affect the system’s stability and reliability. The major fact is that the inherent disturbances in the power grid side like voltage sags and frequency fluctuations would interact with the interconnected RESs and thus result in severe PQ issues at the end. The major objective of this work is to provide a brief analysis of PQ issues and their mitigation techniques of grid-integrated RESs.

The various research activities happening across the globe on PQ mitigation techniques are presented in [6, 13, 14, 22]. A detailed discussion of various PQ disturbances in the utility grid with the integration of RESs is presented in [7–9]. A brief review analysis of FACTS technologies besides the use of three-phase APF/STATCOM technologies for PQ improvement is presented in [10, 11, 15, 16, 25]. Various novel control algorithms for the enhancement of PQ of a grid-connected SPVS are explained in [12]. Mitigation techniques for minimization of voltage dips and flickers in electricity networks are addressed in [14, 18]. Voltage profile improvement in power transmission systems using the dragonfly algorithm and fuzzy logic control is addressed in [17, 19]. Harmonic mitigation and PQ improvement DSTATCOM is addressed in [20, 40, 42]. A detailed discussion over standalone PV-BES-DG-based microgrid with PQ improvements is given in [21, 24].

Table 2 Summary of author contributions

Reference number	Proposed area	Proposed work/techniques
[7–9]	Integration of RESs to grid	Various PQ issues and mitigation techniques
[10, 11]	RESs	FACTS devices and three-phase APF/STATCOM technologies
[52]	Voltage profile improvement in transmission lines	Dragonfly optimization algorithm and Fuzzy logic controller
[26–28]	Grid-penetrated RESs	DVR, APF, and other FACTS devices
[29–37]	Control of grid integration of residential solar photovoltaic systems	IPT, ANN, PBT, ISCT, SRFT techniques
[43–51]	PQ mitigation techniques	Soft computing techniques

A stability enhancement method based on an adaptive virtual resistor for an electric-hydrogen hybrid DC microgrid grid-connected inverter under a weak grid is explained in [22]. PQ improvement and low-voltage ride dynamic voltage restorer (DVR) [26] and active power filter (APF) for mitigation of PQ issues in grid integration of wind and photovoltaic energy conversion system [27] and various FACTS devices for voltage stability enhancement of renewable-integrated power grid was reviewed [28]. A detailed discussion on instantaneous power theory [29], artificial neural network (ANN) [30, 38], power balance theory [31], instantaneous symmetrical components theory [32, 34], and synchronous reference frame theory [37] for control and grid integration of residential solar photovoltaic systems is addressed.

A technique for FFT harmonics compensation and leakage current suppression in a 10-kW PV inverter is discussed in [33]. Simulation of synchronous reference frame PLL for grid synchronization using Simulink [35, 36] for single-phase grid converter applications is addressed (Table 2). A detailed analysis on grid-connected dual-voltage source inverters with PQ improvement features is given in [39]. Mitigation of PQ issues for grid-connected photo voltaic system using soft computing techniques is given in [43–52].

The chapter is organized as follows: PQ and its issues in grid-connected RESs are discussed in Sect. 2. Section 3 provides information about the mitigation techniques employed to alleviate PQ issues, followed by Conclusions in Sect. 4.

2 Power Quality

The term PQ refers to the ability to maintain the magnitude and frequency of the original sinusoidal wave of voltage and current of the electrical network. Under many operating conditions, it mainly associates with the control of voltage only because current control is very difficult when compared with that of controlling the voltage [4]. Some of the PQ issues that arise due to the integration of RESs

into the grid are listed as follows: Fluctuations in voltage, voltage sags and swells, long-duration voltage variations, noise, waveform distortions, power frequency variations, harmonics, voltage spikes, transients, flickers, and so on.

The causes and effects of these issues are explained in this section. Fluctuations in voltage indicate sudden changes in the amplitudes of the voltage in an electrical power system network, and these are mainly caused by the load-switching phenomenon. The severity of this issue is very high and results in effects such as over and under voltages, light flickering, and so on. Voltage sags refer to the reduction in the magnitude of supply voltage, whereas voltage swells refer to the transitory increase in the voltage levels above the specified rated values. The severity of this issue is moderate and these are mainly caused due to excessive sudden increases in loads and heavy inrush currents and results in effects such as problems due to overloading, loss of data, damage of equipment, and so on.

Transients are the ones that are caused by sudden changes in the voltages and currents and their severity is very high and causes more damage to the power system network even though they last for a few milliseconds. Voltage spikes refer to the increase of voltage above the specified tolerance levels. Long-time voltage interruption refers to either complete interruption or decline in the voltage or load current magnitudes for a duration range from a few milliseconds to 2 s.

The superimposition of high-frequency signals with the reference waveform of the power system network results in noise, and due to these data, the loss will occur. The fluctuations in voltage levels in between the range from 90% to 110% of the rated voltage waveform result in Flicker and cause damage to equipment in the power system network. Waveform distortions are the deviations of signals from an ideal sinusoidal waveform, whereas power frequency variations refer to the deviations in fundamental power frequency from its rated frequency level (50 Hz). The distortions in the waveforms of voltage and currents due to non-linear loads result in harmonics [6–14].

3 PQ Mitigation Techniques

As we have discussed in the previous section, SESs encountered various PQ issues when they were being integrated with the grid. When there is a necessity to connect the DG sources like SESs to the three-phase grid, the major concern is the PQ (Fig. 1). There are several parameters to be taken into account in integrating these sources into the grid to make the interconnections free from PQ problems. Some of the methodologies involved in mitigating the PQ issues associated with grid-connected RESs are discussed in this section.

The FACTS devices play a major role in providing good solutions to encountering PQ issues. There are many types of FACTS devices available. In this section, the primary focus is on giving an idea over the role of custom power devices (CPDs) in improving the voltage and current quality of a grid-connected SES followed by DFACTS devices and their different control algorithms in mitigating PQ issues.

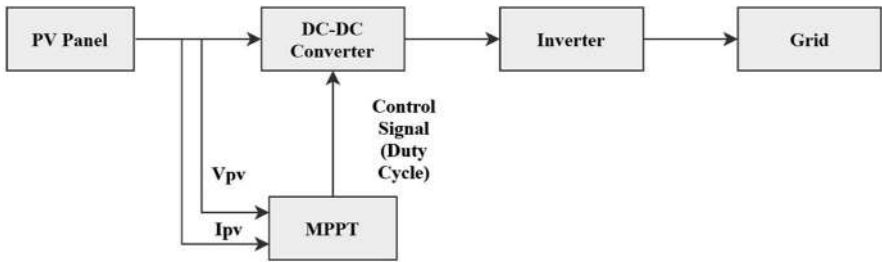


Fig. 1 Grid integration of the PV system

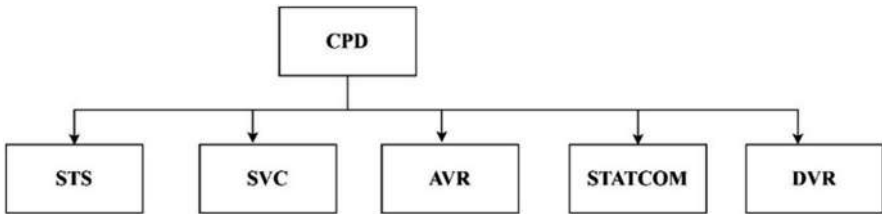


Fig. 2 Custom power devices

Table 3 FACTS devices for voltage quality improvement

FACTS device	Application
STS	Voltage control
SVC	Voltage control
AVR	Voltage control
STATCOM	Voltage control
DVR	Voltage control
VSC	Improves power system reliability
Dynamic distribution system compensator	Voltage stability enhancement
STATCOM power filter compensator	Stabilization of AC voltage, Power factor improvement
D-SSC	Mitigation of low-frequency oscillations

Besides, the role of soft computing techniques in the alleviation of PQ issues is proposed in this section.

Voltage Quality Improvement

CPDs gained their popularity for providing effective solutions for mitigating PQ issues in grid-connected RESs. There are several CPDs available, and these devices are incorporated with power electronic devices to alleviate PQ issues (Fig. 2). A brief overview on various CPDs which are shown in Fig. 2 will be discussed here (Table 3). Static transfer switch (STS) is used to automatically transfer loads between two independent AC power supplies without any interruption irrespective of any disturbance due to voltage occurring in any one of the supply sides.

Table 4 Performance comparison of DFACTS devices for voltage quality improvement [27]

DFACTS devices				
Parameters	DSTATCOM	DSSSC	IPQC	UPQC
Compensation of reactive power (Q)	2	1	3	3
Suppression of harmonics	Limited	Limited	Limited	Limited
Load balancing	2	2	3	2
Transient and steady-state stabilities	2	2	4	2
Voltage control	2	2	3	2
Number of switching components	9–12	6	6	12
Power rating of the converter	M	L	L	M
Cost	2	2	4	3

Poor—1, Good—2, Best—3, Better—4

In other words, STS will be switched to a backup power source so quickly when a failure or disturbance occurs in the main source [15, 16].

Static VAR compensator (SVC) is capable of supplying or absorbing reactive power control or regulating the voltage stability of the system. Due to some mismatch between power generation and demand, it results in voltage instability which is a major concern in electrical power systems. SVC is used for voltage profile improvement [17–20].

Automatic voltage regulator (AVR) is a device employed in the power system network in order to track the reference voltage. It may be either static or rotary type. Due to discrete changes in the voltages, the response time of AVRs is very small. A static compensator is one that also belongs to the class of FACTS devices that will be integrated at the point of common coupling (PCC) and provide voltage control by injecting reactive power into the system. Besides, transient stability and damping can be improved by imparting this device into the system. STATCOM has gained popularity in alleviating PQ issues, especially voltage fluctuations associated with RESs (Table 4).

DVR is the one that generally injects voltage into the power system to meet the load voltage requirement. The other methods include energy storage systems, energy conversion optimization, spinning reserve, and other special techniques [10, 21–26].

DFACTS Devices

DFACTS devices play a major role in mitigating the PQ issues of grid-connected RESs. These devices are used to maintain PQ in the distributed grid, and it can be done by controlling various parameters such as phase angle, harmonics of voltage and currents, line impedance, and so on. Better voltage regulation can be achieved, thereby resulting in improved stability in the system [2, 3, 7]. Types of the devices are shown in Fig. 3: series, shunt, and series-shunt combination. The uses of all the devices are compared in eight areas. The comparison is shown in Table 2.

- (1) VAR
- (2) Harmonics suppression

- (3) Load balancing
- (4) Transient and steady-state stability
- (5) Voltage control
- (6) Switching components
- (7) Power rating of the devices
- (8) Cost

DFACTS Control Algorithms

To enhance the performance of a grid-connected RES, DFACTS control algorithms were found to be the best alternative. The classification diagram is shown in Fig. 3. As the frequency-domain algorithms cannot provide proper alleviation of PQ issues as these are associated with some drawbacks like inaccuracy and so on, conventional control algorithms will be preferred [35–40]. The main idea is to generate switching signals to the DFACTS devices, and to achieve this (Fig. 4), PI controllers and necessary transformations are necessary. The reference and actual signals are compared and sent to the PI controller. The components obtained after doing necessary transformations (Clark and reverse Clark) will be sent to a low-pass filter to get fundamental DC components [27–34].

Current Quality Improvement Enormous research is going on in all the research institutes across the globe to find the mitigation techniques to eliminate all kinds of harmonics in the grid side due to the integration of RESs into the power grid. Some

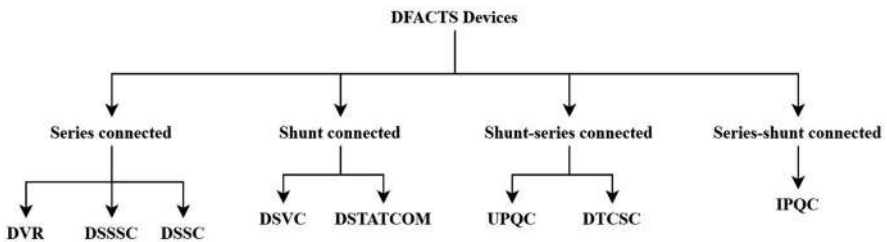


Fig. 3 DFACTS devices classification [27]

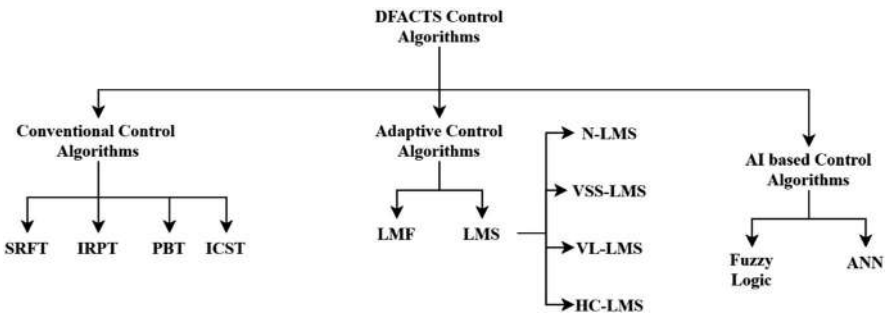


Fig. 4 DFACTS control algorithms [3]

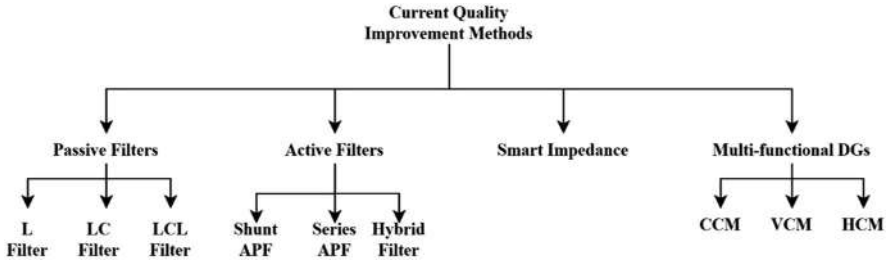


Fig. 5 Current quality improvement methods [3]

of the mitigation techniques have been discussed here. *Passive filters* include various combinations of inductors and capacitors and are extensively used in blocking the undesired harmonic components in a power system.

They offer very low impedance to the objectionable or unwanted harmonic frequency components that emerge from the supply input side which leads to block them effectively. In the grid integration of RESs, the fact is that the power system will be affected by various high-frequency harmonic components, and hence various grid-integrated passive filters such as L, LC, and LCL filters will be employed at the feeding end of the grid interfacing inverters (Fig. 5).

The *L filter* is a simple first-order filter and is employed at the output side of the grid interfacing inverters to block the high-frequency components. However, the researchers found that it is not giving effective results and has not met the IEEE standards, and hence another filter called *LC filter* which is a second-order filter has been employed. It was observed that the damping performance is found to be good in this filter when compared with the simple L filter and the major advantage is it was more efficient and cost-effective after placing a capacitor in parallel. There is, of course, another type of filter called *LCL filter* which is a third-order filter that gives better mitigation or attenuation of current harmonics in the grid side.

The major problem with passive filters is that they will cause severe negative constraints in the grid integration of RESs if their design is not proper. Owing to this severe constraint, *active power filters* have emerged into the picture where some promising results have been seen by the researchers as they found that active power filters have overcome all the demerits of passive filters. Besides, the active power filters provide good reactive power support and attenuate variations in voltage thereby improving the power factor. Active power filters were classified as *Shunt* and *series* active power filters. To handle PQ issues such as load imbalance, reactive power compensation, and current harmonics, shunt active power filters will be treated as the better ones.

In actual practice, harmonic elimination from the source current can be obtained by shunt active power filters and it is achieved by injecting compensating current at the PCC into the distribution grid, whereas the active power filters eliminate the voltage imbalance issues at PCC by injecting their components of harmonic voltage

Table 5 FACTS devices for current quality improvement [3]

CQI device	Applications
Passive filters	Reactive power support
Active power filters	Harmonic mitigation, reactive power compensation, load balancing
Smart impedance	Harmonic mitigation, reactive power compensation
Multifunctional DGs	Enhancement of voltage profile

in series with the grid [1, 2, 4–14, 33]. The drawbacks of series active power filters were overcome by employing a new filter which was formed by collaborating the passive filters with the series active filters (Table 5). To overcome the limitations related to tuned passive filters, a new tuned passive filter named *smart impedance* was proposed.

In smart grids, this method gives better stability enhancement and better voltage regulation and also mitigates the voltage distortions caused due to non-linear loads when the source impedance is not significant. It can offer very low impedance to the unwanted harmonic and thereby provide a short-circuited path for the harmonics of load current. Besides, the harmonics currents generated owing to the voltage source distortions can be mitigated by offering infinite impedance to the source voltage harmonics.

A *multifunctional DG* can be used as a PQ improvement device at the point of grid integration. There are three control schemes available in this method, namely, the *current control method (CCM)*, *voltage control method (VCM)*, and *hybrid control method (HCM)*. The CCM can be used for employing harmonic current compensation in a distribution generator system and this control can only be employed for current-controlled DGs. The VCM is also employed for grid interfacing inverters but cannot be used for solving any system harmonic issues due to the absence of closed-loop line current regulation, but it is highly effective in both grid-integrated and islanded modes. Hence, harmonic compensation that is not available with CCM can be feasible with VCM.

There are some limitations to be overcome in CCM and VCM and to achieve this, HCM has been employed. It is employed by controlling both the fundamental voltage of the capacitor and the line harmonic currents of an LCL filter. In this control method, by controlling or regulating the fundamental capacitor voltage, the control of DG output power can be achieved. Besides, with the help of a closed-loop harmonic current compensator, the line current harmonics can be controlled. FACTS devices for current quality improvement are shown in Table 5 [41–43].

Soft Computing Techniques in Grid-Connected RESs

Soft computing techniques can provide extensive solutions for many complex problems concerning various areas of electrical power systems. The issues in PV systems such as sudden temperature and irradiation changes can be addressed by these techniques effectively when compared with the conventional techniques. While analyzing the soft computing techniques in grid-connected SES, various stages such as a reconfiguration of the PV panel, maximum power point tracking

in the converter employed, mitigation of harmonics in inverters, and detection of islanding modes will be considered. There are some other parameters such as convergence, the complexity of the algorithm, oscillations around the operating point, and so on. To avoid the effects of partial shading on the PV panels, mitigate or alleviate nonlinear I–V curves and unwanted peaks by generating a duty cycle, mitigate harmonics in the DC–AC conversion stage, and verify the voltage, frequency, and phase angle for detecting Islanding mode in the grid-side inverters, the soft computing techniques are mostly used. In large electricity-generating plants, many solar PV panels will be connected in series and parallel to meet the power demand at the end. Partial shading is one of the serious issues to be analyzed and mitigated at the earliest, as it can reduce the total panel power if at least one panel is shaded. The two most popular methods are total cross-tied and Su Do Ku configurations are mostly used in *panel reconfiguration*.

In conventional methods, manpower is highly required to shift the panels by interchanging the connections, which also results in conduction losses. Soft computing techniques such as genetic algorithm, differential evolution, particle swarm optimization (PSO), fire fly algorithm (FFA), ant colony optimization (ACO), and so on are used to avoid the abovementioned issues by designing a proper fitness function that results in maximizing the current and the voltage generated per string. The harmonic elimination process is very important as the harmonics affect the PQ of an electric power system. Harmonics are generated mainly due to nonlinear loads. Due to the presence of harmonics, the voltage and current waveforms will get distorted and hence result in poor PQ at the end [39–48]. There are several conventional methods like SPWM and SVM, respectively, and artificial intelligence techniques like adaptive fuzzy-neural-network, adaptive fuzzy sliding control, and adaptive fuzzy backstepping to alleviate the harmonic content in the inverters, but due to some major drawbacks in conventional techniques, the researchers proposed soft computing techniques as an alternative to mitigate harmonics in grid-side inverters. For large power system networks, a simple two-level inverter cannot be used, and hence multilevel inverters such as diode clamped, flying capacitor type, and cascaded multilevel inverters will be used. ANNs are used to mitigate higher-order dominant frequency components. The FLC is employed to improve the performance of the grid-side inverter by mitigating harmonics generated due to non-linear loads. Similarly, evolutionary computing techniques are being employed to alleviate the harmonic content [2, 12, 31–38].

4 Conclusions

Due to the penetration of RESs in today's modern power grids, the world has come up with the best alternative for meeting the power demand. At the same time, some major challenges like mitigation of PQ issues are to be faced. In this chapter, a brief analysis of PQ issues associated with grid-connected RESs is presented. The types of RESs and their applications are discussed. The PQ issues classification

and its causes are addressed. The major focus is on discussing the PQ mitigation techniques by using various FACTS devices and their control algorithms, soft computing techniques, and so on. Various applications of FACTS devices besides their performance comparison in mitigating the PQ issues were presented in this chapter.

References

1. M. Bajaj and A. K. Singh, "Grid integrated renewable DG systems: A review of power quality challenges and state-of-the-art mitigation techniques," *Int. J. Energy Res.*, vol. 44, no. 1, pp. 26–69, 2020, <https://doi.org/10.1002/er.4847>.
2. K. W. Kow, Y. W. Wong, R. K. Rajkumar, and R. K. Rajkumar, "A review on performance of artificial intelligence and conventional method in mitigating PV grid-tied related power quality events," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 334–346, 2016, <https://doi.org/10.1016/j.rser.2015.11.064>.
3. K. Jha and A. G. Shaik, "A comprehensive review of power quality mitigation in the scenario of solar PV integration into utility grid," *Electron. Energy*, vol. 3, no. January, p. 100103, 2023, <https://doi.org/10.1016/j.prime.2022.100103>.
4. E. Hossain, M. R. Tur, S. Padmanaban, S. Ay, and I. Khan, "Analysis and Mitigation of Power Quality Issues in Distributed Generation Systems Using Custom Power Devices," *IEEE Access*, vol. 6, pp. 16816–16833, 2018, <https://doi.org/10.1109/ACCESS.2018.2814981>.
5. X. Liang, "Emerging Power Quality Challenges Due to Integration of Renewable Energy Sources," *IEEE Trans. Ind. Appl.*, vol. 53, no. 2, pp. 855–866, 2017, <https://doi.org/10.1109/TIA.2016.2626253>.
6. F. G. Montoya, A. García-Cruz, M. G. Montoya, and F. Manzano-Agugliaro, "Power quality techniques research worldwide: A review," *Renew. Sustain. Energy Rev.*, vol. 54, pp. 846–856, 2016, <https://doi.org/10.1016/j.rser.2015.10.091>.
7. G. S. Chawda et al., "Comprehensive Review on Detection and Classification of Power Quality Disturbances in Utility Grid with Renewable Energy Penetration," *IEEE Access*, vol. 8, pp. 146807–146830, 2020, <https://doi.org/10.1109/ACCESS.2020.3014732>.
8. A. Q. Al-Shetwi, M. A. Hannan, K. P. Jern, A. A. Alkahtani, and A. E. P. G. Abas, "Power quality assessment of grid-connected PV system in compliance with the recent integration requirements," *Electron.*, vol. 9, no. 2, pp. 1–22, 2020, <https://doi.org/10.3390/electronics9020366>.
9. A. Q. Al-Shetwi, M. A. Hannan, K. P. Jern, M. Mansur, and T. M. I. Mahlia, "Grid-connected renewable energy sources: Review of the recent integration requirements and control methods," *J. Clean. Prod.*, vol. 253, 2020, <https://doi.org/10.1016/j.jclepro.2019.119831>.
10. F. H. Gandoman et al., "Review of FACTS technologies and applications for power quality in smart grids with renewable energy systems," *Renew. Sustain. Energy Rev.*, vol. 82, no. August 2017, pp. 502–514, 2018, <https://doi.org/10.1016/j.rser.2017.09.062>.
11. W. U. K. Tareen et al., "Mitigation of power quality issues due to high penetration of renewable energy sources in electric grid systems using three-phase APF/STATCOM technologies: A review," *Energies*, vol. 11, no. 6, 2018, <https://doi.org/10.3390/en11061491>.
12. N. P. Babu, J. M. Guerrero, P. Siano, R. Peesapati, and G. Panda, "A Novel Modified Control Scheme in Grid-Tied Photovoltaic System for Power Quality Enhancement," *IEEE Trans. Ind. Electron.*, vol. 68, no. 11, pp. 11100–11110, 2021, <https://doi.org/10.1109/TIE.2020.3031529>.
13. M. Obi and R. Bass, "Trends and challenges of grid-connected photovoltaic systems—A review," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 1082–1094, 2016, <https://doi.org/10.1016/j.rser.2015.12.289>.

14. O. Ipinnimo, S. Chowdhury, S. P. Chowdhury, and J. Mitra, "A review of voltage dip mitigation techniques with distributed generation in electricity networks," *Electr. Power Syst. Res.*, vol. 103, pp. 28–36, 2013, <https://doi.org/10.1016/j.epsr.2013.05.004>.
15. P. A. Desale, V. J. Dhawale, and R. M. Bandgar, "Brief Review Paper on the Custom Power Devices for Power Quality Improvement," *Int. J. Electr. Electron. Eng.*, vol. 7, no. 7, pp. 723–733, 2014.
16. S. Agalar and Y. A. Kaplan, "Power quality improvement using STS and DVR in wind energy system," *Renew. Energy*, vol. 118, pp. 1031–1040, 2018, <https://doi.org/10.1016/j.renene.2017.01.013>.
17. J. Vanishree and V. Ramesh, "Optimization of size and cost of Static VAR Compensator using Dragonfly algorithm for voltage profile improvement in power transmission systems," *Int. J. Renew. Energy Res.*, vol. 8, no. 1, pp. 56–66, 2018, <https://doi.org/10.20508/ijrer.v8i1.6933.g7281>.
18. S. Dhivya, "Various model of flicker measurement methods and static var compensation for power quality improvement," no. 2, pp. 3392–3397, 2020.
19. A. Naderipour et al., "Optimal designing of static var compensator to improve voltage profile of power system using fuzzy logic control," *Energy*, vol. 192, p. 116665, 2020, <https://doi.org/10.1016/j.energy.2019.116665>.
20. O. P. Mahela, B. Khan, H. H. Alhelou, S. Tanwar, and S. Padmanaban, "Harmonic mitigation and power quality improvement in utility grid with solar energy penetration using distribution static compensator," *IET Power Electron.*, vol. 14, no. 5, pp. 912–922, 2021, <https://doi.org/10.1049/pel2.12074>.
21. V. Narayanan, S. Kewat, and B. Singh, "Standalone PV-BES-DG Based Microgrid with Power Quality Improvements," *Proc.—2019 IEEE Int. Conf. Environ. Electr. Eng. 2019 IEEE Ind. Commer. Power Syst. Eur. IEEEIC/I CPS Eur. 2019*, pp. 3–8, 2019, <https://doi.org/10.1109/IEEEIC.2019.8783251>.
22. L. Li, W. Chen, Y. Han, Q. Li, and Y. Pu, "A Stability Enhancement Method Based on Adaptive Virtual Resistor for Electric-hydrogen Hybrid DC Microgrid Grid-connected Inverter Under Weak Grid," *Electr. Power Syst. Res.*, vol. 191, no. September 2020, 2021, <https://doi.org/10.1016/j.epsr.2020.106882>.
23. Y. Naderi, S. H. Hosseini, S. Ghassem Zadeh, B. Mohammadi-Ivatloo, J. C. Vasquez, and J. M. Guerrero, "An overview of power quality enhancement techniques applied to distributed generation in electrical distribution networks," *Renew. Sustain. Energy Rev.*, vol. 93, no. March, pp. 201–214, 2018, <https://doi.org/10.1016/j.rser.2018.05.013>.
24. V. Narayanan, Seema, and B. Singh, "Multi-Operational PV-BES-DG Based Microgrid with Power Quality Improvement," *2020 IEEE Int. Conf. Power Electron. Smart Grid Renew. Energy, PESGRE 2020*, pp. 1–6, 2020, <https://doi.org/10.1109/PESGRE45664.2020.9070508>.
25. B. Singh and R. Kumar, "A comprehensive survey on enhancement of system performances by using different types of FACTS controllers in power systems with static and realistic load models," *Energy Reports*, vol. 6, pp. 55–79, 2020, <https://doi.org/10.1016/j.egyr.2019.08.045>.
26. A. Benali, M. Khiat, T. Allaoui, and M. Denaï, "Power Quality Improvement and Low Voltage Ride Through Capability in Hybrid Wind-PV Farms Grid-Connected Using Dynamic Voltage Restorer," *IEEE Access*, vol. 6, pp. 68634–68648, 2018, <https://doi.org/10.1109/ACCESS.2018.2878493>.
27. G. S. Chawda, A. G. Shaikh, O. P. Mahela, S. Padmanaban, and J. B. Holm-Nielsen, "Comprehensive Review of Distributed FACTS Control Algorithms for Power Quality Enhancement in Utility Grid with Renewable Energy Penetration," *IEEE Access*, vol. 8, pp. 107614–107634, 2020, <https://doi.org/10.1109/ACCESS.2020.3000931>.
28. F. Shaikh and B. Joseph, "Simulation of synchronous reference frame PLL for grid synchronization using Simulink," *Int. Conf. Adv. Comput. Commun. Control 2017, ICAC3 2017*, vol. 2018-Janua, pp. 1–6, 2018, <https://doi.org/10.1109/ICAC3.2017.8318790>.
29. S. Chandrasekaran and K. Ragavan, "Phase Locked Loop technique based on sliding DFT for single phase grid converter applications," *PEDES 2012—IEEE Int. Conf. Power Electron. Drives Energy Syst.*, pp. 1–4, 2012, <https://doi.org/10.1109/PEDES.2012.6484311>.

30. T. Sathiyarayanan and S. Mishra, "Synchronous reference frame theory based model predictive control for grid connected photovoltaic systems," *IFAC-PapersOnLine*, vol. 49, no. 1, pp. 766–771, 2016, <https://doi.org/10.1016/j.ifacol.2016.03.149>.
31. P. N. Jyothy Lakshmi and M. R. Sindhu, "Residential Solar Photovoltaic System with Artificial Neural Network Based Controller," 2018 Int. Conf. Control. Power, Commun. Comput. Technol. ICCPCCT 2018, pp. 416–420, 2018, <https://doi.org/10.1109/ICCPCCT.2018.8574296>.
32. M. V. Manoj Kumar, M. K. Mishra, and C. Kumar, "A Grid-Connected Dual Voltage Source Inverter with Power Quality Improvement Features," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 482–490, 2015, <https://doi.org/10.1109/TSTE.2014.2386534>.
33. W. U. Tareen, S. Mekhilef, M. Seyedmahmoudian, and B. Horan, "Active power filter (APF) for mitigation of power quality issues in grid integration of wind and photovoltaic energy conversion system," *Renew. Sustain. Energy Rev.*, vol. 70, no. January 2016, pp. 635–655, 2017, <https://doi.org/10.1016/j.rser.2016.11.091>.
34. B. B. Adetokun and C. M. Muriithi, "Application and control of flexible alternating current transmission system devices for voltage stability enhancement of renewable-integrated power grid: A comprehensive review," *Heliyon*, vol. 7, no. 3, p. e06461, 2021, <https://doi.org/10.1016/j.heliyon.2021.e06461>.
35. F. Wang and C. Zhang, "Protection and control for grid connected photovoltaic power generation system based on instantaneous power theory," *Proc. 2009 Pacific-Asia Conf. Circuits, Commun. Syst. PACCS 2009*, pp. 356–359, 2009, <https://doi.org/10.1109/PACCS.2009.166>.
36. Y. Sun, S. Li, B. Lin, X. Fu, M. Ramezani, and I. Jaithwa, "Artificial Neural Network for Control and Grid Integration of Residential Solar Photovoltaic Systems," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1484–1495, 2017, <https://doi.org/10.1109/TSTE.2017.2691669>.
37. B. Singh, D. T. Shahani, and A. K. Verma, "Power balance theory based control of grid interfaced solar photovoltaic power generating system with improved power quality," *PEDES 2012—IEEE Int. Conf. Power Electron. Drives Energy Syst.*, 2012, <https://doi.org/10.1109/PEDES.2012.6484359>.
38. A. S. Kumar and K. Prakash, "Analysis of split Capacitor based D-STATCOM with LCL filter using Instantaneous Symmetrical Components Theory," *Proceeding IEEE—2nd Int. Conf. Adv. Electr. Electron. Information, Commun. Bio-Informatics, IEEE—AEEICB 2016*, pp. 103–108, 2016, <https://doi.org/10.1109/AEEICB.2016.7538406>.
39. H. Bin, Q. Hua, and H. Cui, "A technique for FFT harmonics compensation and leakage current suppression in 10kW PV inverter," *Conf. Proc.—2012 IEEE 7th Int. Power Electron. Motion Control Conf.—ECCE Asia, IPEMC 2012*, vol. 2, pp. 836–840, 2012, <https://doi.org/10.1109/IPEMC.2012.6258954>.
40. F. Harirchi and M. G. Simoes, "Enhanced Instantaneous Power Theory Decomposition for Power Quality Smart Converter Applications," *IEEE Trans. Power Electron.*, vol. 33, no. 11, pp. 9344–9359, 2018, <https://doi.org/10.1109/TPEL.2018.2791954>.
41. M. K. A. Panda and S. P. S. Pathak, "us cr t," vol. 7217, no. October, 2015, <https://doi.org/10.1080/00207217.2015.1092600>.
42. L. Ashok Kumar and V. Indragandhi, "Power quality improvement of grid-connected wind energy system using facts devices," *Int. J. Ambient Energy*, vol. 41, no. 6, pp. 631–640, May 2020, <https://doi.org/10.1080/01430750.2018.1484801>.
43. R. Aljendy, R. R. Nasyrov, A. Y. Abdelaziz, and A. A. Z. Diab, "Enhancement of Power Quality with Hybrid Distributed Generation and FACTS Device," *IETE J. Res.*, vol. 68, no. 3, pp. 2259–2270, 2022, <https://doi.org/10.1080/03772063.2019.1698321>.
44. T. A. Jumani et al., "Computational intelligence-based optimization methods for power quality and dynamic response enhancement of ac microgrids," *Energies*, vol. 13, no. 15, 2020, <https://doi.org/10.3390/en13164063>.
45. A. Ouai, L. Mokrani, M. Machmoum, and A. Houari, "Control and energy management of a large scale grid-connected PV system for power quality improvement," *Sol. Energy*, vol. 171, no. November 2017, pp. 893–906, 2018, <https://doi.org/10.1016/j.solener.2018.06.106>.

46. J. R. Vazquez, A. D. Martin, and R. S. Herrera, "Neuro-fuzzy control of a grid-connected photovoltaic system with power quality improvement," IEEE EuroCon 2013, no. July, pp. 850–857, 2013, <https://doi.org/10.1109/EUROCON.2013.6625082>.
47. S. R. Das, P. K. Ray, A. K. Sahoo, K. K. Singh, G. Dhiman, and A. Singh, "Artificial intelligence based grid connected inverters for power quality improvement in smart grid applications," Comput. Electr. Eng., vol. 93, no. September 2020, p. 107208, 2021, <https://doi.org/10.1016/j.compeleceng.2021.107208>.
48. B. S. Goud, B. L. Rao, and C. R. Reddy, "An intelligent technique for optimal power quality reinforcement in a grid-connected HRES system: EVORFA technique," Int. J. Numer. Model. Electron. Networks, Devices Fields, vol. 34, no. 2, pp. 1–30, 2021, <https://doi.org/10.1002/jnm.2833>.
49. A. Krama, L. Zellouma, B. Rabhi, S. S. Refaat, and M. Bouzidi, "Real-time implementation of high performance control scheme for grid-tied PV system for power quality enhancement based on MPPC-SVM optimized by PSO algorithm," Energies, vol. 11, no. 12, 2018, <https://doi.org/10.3390/en11123516>.
50. C. R. Reddy, B. S. Goud, F. Aymen, G. S. Rao, and E. C. Bortoni, "Power quality improvement in hres grid connected system with fopid based atom search optimization technique," Energies, vol. 14, no. 18, 2021, <https://doi.org/10.3390/en14185812>.
51. M. Balamurugan, S. K. Sahoo, and S. Sukchai, "Application of soft computing methods for grid connected PV system: A technological and status review," Renew. Sustain. Energy Rev., vol. 75, no. January, pp. 1493–1508, 2017, <https://doi.org/10.1016/j.rser.2016.11.210>.
52. O. E. Okwako, Z. H. Lin, M. Xin, K. Premkumar, and A. J. Rodgers, "Neural Network Controlled Solar PV Battery Powered Unified Power Quality Conditioner for Grid Connected Operation," Energies, vol. 15, no. 18, 2022, <https://doi.org/10.3390/en15186825>.

Traffic Control System-Based Congestion Control and Emergency Vehicle Clearance



K. Krishna Reddy, S. Noor Mahammad, J. Divya, C. Vamsi, S. Ameer Basha, and B. Suresh Reddy

Abstract During a time crunch, time can act as either a helpful ally or a dangerous adversary. Because of this, it is very necessary to eke out every last second. Because of the recent rise in the number of cars, there has been a significant increase in the amount of congestion on the roads. As a direct consequence of this, a significant number of individuals have been killed because ambulances were unable to get to their destinations on time. Arriving at the location without incident and as promptly as feasible is the primary objective of every emergency vehicle. The odds of surviving are diminished by each second that is wasted due to stopped traffic or red lights along the road to the goal. Therefore, it is absolutely necessary to give ambulances signal priority in order to ensure that they are able to navigate the heavy traffic with the least amount of delay possible. In this chapter, a novel system that makes use of Internet of things to assist in accomplishing this objective is offered.

Keywords Traffic Management System · Intelligent Traffic Control · Real Time Lights

1 Introduction

INDIA is the country with the world's second-largest population and has an economy that is expanding at a rapid rate. It is experiencing serious issues with traffic congestion in its major cities. Because of limited space and high costs, the expansion of infrastructure is occurring at a far slower rate than the increase in the number of cars [1]. In Fig. 1, there is no system of lanes in use; therefore, the traffic is quite chaotic. It requires methods for traffic control that are distinct from those used in wealthy countries. Congestion may have detrimental effects that can

K. K. Reddy (✉) · S. N. Mahammad · J. Divya · C. Vamsi · S. A. Basha · B. S. Reddy
Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering and Technology, Chittoor, India
e-mail: krishnareddy206@mtieat.org



Fig. 1 The Bangalore traffic situation

be mitigated by the intelligent management of traffic flows. In recent years, wireless networks have gained widespread usage in the road transport industry as a result of the more cost-efficient possibilities they give [2]. ZigBee, RFID, and GSM are examples of technologies that might be leveraged to give cost-effective solutions in the field of traffic management.

The radio-frequency identification (RFID) system is a wireless technology that transfers data between an RFID tag and an RFID reader by means of radio frequency electromagnetic energy. While some RFID systems will only function within a range of inches or centimeters, others may function for a distance of 100 m (300 feet) or more. A GSM modem is a specialized kind of modem that, similar to a mobile phone, may be equipped to read a SIM card and run on the basis of a subscription to the services of a mobile operator. Controlling modems requires using the AT commands. These instructions originate from Hayes commands, which were used by Hayes smart modems in the past. ZigBee is a low-power technology that may be used in any of the possible job configurations to carry out a set of predetermined responsibilities. It makes use of the ISM bands, which are 868 MHz in Europe, 915 MHz in the United States and Australia, and 2.4 GHz everywhere else. The 868-MHz frequency band has a data transmission rate of 20 kilobits per second, whereas the 2.4-GHz frequency band has a data transmission rate of 250 kilobits per second [3, 4]. In the event of a radio frequency of 868/915 MHz, the ZigBee utilizes 11 channels, and in the case of a radio frequency of 2.4 GHz, it utilizes 16 channels. In addition to this, it makes use of two other channel topologies, namely, CSMA/CA and slotted CSMA/CA [5]. This chapter is broken down into five sections. The literature survey is discussed in Sect. 2. In Sect. 3, the present challenges that must be faced while making room for an ambulance and other vehicles are dissected and discussed. In addition to this, it discusses how the suggested model would solve the challenges that emerging nations as well as rich ones are now experiencing. In Sect.

4, the specifics of how the suggested model would be implemented are presented. And had a Focuses on the **improvements** made to the work, presenting the changes for Extension implementation.

2 Literature Survey

In cities of emerging countries like India, a key issue that has to be addressed is traffic congestion. A large portion of the increase in the number of automobiles seen in urban areas may be attributed to the expanding middle class as well as the growing urban population [6]. Congestion on roadways always leads to sluggish moving traffic, which increases the amount of time it takes to travel, and as a consequence, stands out as one of the most significant problems that may occur in urban areas. The green wave system is mentioned in [7]. This system is used to offer clearance to any emergency vehicle by changing all of the red lights on the route of the emergency vehicle to green, therefore delivering a full green wave to the desired vehicle. This system is used to provide clearance to any emergency vehicle. A “green wave” refers to the synchronization of the green phase of different traffic lights. A “green wave” configuration ensures that every vehicle that drives through an area with a green light will continue to be given green lights as it proceeds down the road. In addition to following the car as it travels along the green wave route, the system will also follow it when it travels through a traffic signal. One of the benefits of using this technique is that the global positioning system (GPS) device that is installed in the car does not need any extra electricity. The most significant drawback of green waves is that, in the event that the wave is disrupted, this might lead to traffic issues, which, in turn, can be made worse by the synchronization of the signals.

When this happens, the green wave’s line of cars builds until some of them are unable to make it to the green lights in time and must halt. It is a condition known as “over-saturation” [12, 13]. RFID traffic management is mentioned in [8] as a means of avoiding issues with conventional traffic control systems, particularly those using image processing and beam interruption methods. This RFID method accommodates busy intersections with many different types of traffic. It is a useful method for keeping track of time, as it calculates an adaptive plan for the movement of each traffic column in real time. The algorithm operates in real time and makes decisions that are analogous to those made by a police officer. The calculations and evaluations are based on the number of cars in each column and the route. The study has a flaw in that it does not elaborate on how emergency vehicles and the person in charge of the traffic lights exchange information. There was a proposal for an automated lane-clearing system for ambulances using RFID and GPS in [9]. The goal of this project is to speed up the time it takes for an ambulance to get to a hospital by clearing the lane it is in before the vehicle hits a traffic light. To do this, the traffic light in the ambulance’s route may be programmed to turn green at a certain distance from the intersection. To avoid needless traffic backups, RFID

is used to differentiate between emergency and non-emergency situations. Using transceivers and GPS, an ambulance can talk to a traffic light pole.

No human interaction is needed at the intersections because the technology is totally automated. One drawback of this technique is that it requires precise knowledge of the journey's beginning and conclusion points. It may not function if the beginning place is unknown or if the ambulance has to take a different route. In most major cities around the world, traffic is a major transportation problem. Countries like India and China, whose populations are growing at a faster pace than the global average, will feel this impact more acutely. The number of cars in Bangalore, India, for instance, has increased dramatically during the last several years. Therefore, many arterial roads and intersections are running at or above their capacity (i.e., v/c is greater than 1), and peak-hour travel times on some of the most important roads in the core are below 10 km/hour. Managing over 36,000 vehicles, an annual increase of 7–10% in traffic, roads operating at higher capacity ranging from 1 to 4, travel speeds of less than 10 km/hour in some central areas during peak hours, insufficient or no parking space for vehicles, and a limited number of policemen are just some of the main challenges outlined in [10]. As of [11], the city of Bangalore has installed a video traffic surveillance and monitoring system. The traffic management staff must manually analyze the data in order to calculate the appropriate length of each traffic signal at each intersection. It will relay such information to law enforcement in the area so that appropriate measures may be taken.

Congestion management, emergency vehicle clearance, stolen car identification, and so on are all examples of challenges that present technology is ill-equipped to solve. We recommend using our Intelligent Traffic Control System as a solution to these issues. There are essentially three sections to it. The automated signal control system is in the first section. In this area, RFID tags are installed in every car. As soon as it is within the RFID reader range, a signal will be sent. The level of traffic congestion may be calculated by reading the RFID tags on passing cars during a certain time period. As such, it determines how long that particular green light will remain on. The second section is set aside for emergency vehicle access. Here, a ZigBee receiver will be installed at a major intersection and ZigBee transmitter modules will be installed in all emergency vehicles. When the vehicle is being utilized in a life-or-death situation, the alarm will be activated. The signal will be sent over ZigBee and received by the corresponding ZigBee device. The light at the intersection will turn green as a result. When the ambulance has passed, the ZigBee signal is no longer received by the receiver, and the light changes to red. The third section is for finding stolen cars. Here, the RFID scanner checks the tag's information against a stolen RFID database. If a match is detected, the traffic light will turn red and an SMS will be sent to the police station so that the driver will be forced to stop at the intersection and the authorities may take action. These are the parts utilized in the experiment: CC2500RF module, Microchip PIC16F877A, RFID Reader-125KHz-TTL, and SIM300 GSM module. The component pinouts are shown in Fig. 2.

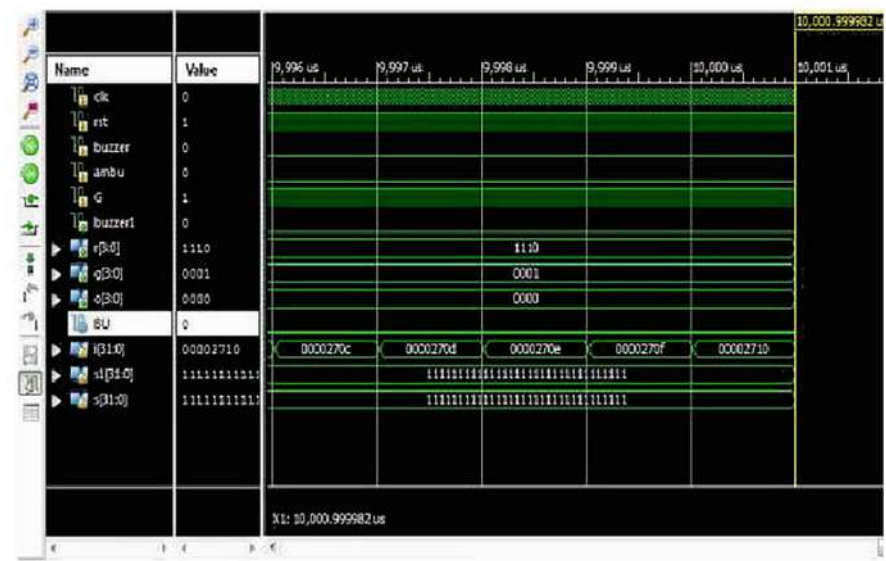


Fig. 3 Behavior modeling with congestion and reset = “1”

The Reader module simultaneously receives 8-bit data that is used to control the junction’s traffic lights’ brightness.

The primary tasks of a traffic control system may be broken down into the following three groups:

- (1) Managing traffic flow
- (2) Clearance for ambulance
- (3) Vehicle theft

3.2 Working of Congestion Control

The RFID Reader in the Traffic Control system begins reading RFID Tags at a predetermined spot on the vehicle.

When the reset is left at zero, the traffic lights will not blink, but when an RFID tag approaches the reader, the car will transmit data to the Reader module, which the reader will pick up. When the reset is set to High, the RFID Reader will keep track of how many cars have passed over the intersection and compare that number to a predetermined threshold; if the count is higher than the threshold, the lights at that intersection will turn green to let traffic through.

The next graphic depicts the simulated behavior of Congestion Control between Resets “0” and “1” of the Traffic Control System:

When $\text{reset} = "1"$ and $G = "1,"$ congestion is shown above a specific intersection in the simulation. Whenever there is a transition between states, the junction begins counting cars again.

4 Conclusion and Enhancements

There is a reduction in the amount of human labor required of the traffic police officer as a result of the implementation of automated traffic signal control that is based on the traffic density along the route. Because the entire system is automated, there is very little need for direct involvement from humans. As soon as the system determines that a car has been reported stolen, the traffic light will switch to the red position, allowing the officer who is now stationed at the intersection to take the proper action. In addition, SMS messages will be issued so that the authorities can be ready to track down the stolen car at the next available intersections. It is essential that vehicles used in emergency situations, such as ambulances and fire engines, get to their locations as quickly as possible. The priceless lives of a great number of individuals may be put in jeopardy if they spend a significant amount of time stuck in traffic. Once the emergency vehicle has passed through the intersection, the traffic light will turn green to indicate that it is safe for drivers to proceed. This will continue for as long as the emergency vehicle is still present. Only when the emergency vehicle has passed through, will the light change to the red position. Testing the prototype with RFID readers that have a further reading range is one way to make more improvements to it. Additionally, a GPS can be installed within the stolen vehicle detection module so that the precise location of the stolen vehicle can be determined. At this time, we are in the process of implementing the system by taking into consideration one route of the traffic intersection. Extending it to include all of the roads at a multi-road junction is one way to make it better.

References

1. Varaprasad G., Wahidabanu R.S.D., "Flexible Routing Algorithm for Vehicular Area Networks", in Proc. IEEE Conference on Intelligent Transport Systems Telecommunications, Osaka, Japan, pp.30–38, 2010.
2. Gokulan B.P., Srinivasan D., "Distributed Geometric Fuzzy Multi-agent Urban Traffic Signal Control", IEEE Transactions on Intelligent Transportation Systems, vol.11, no.3, pp.714–727, 2010.
3. Kumar S., Abhilash P.G., Jyothi D.G., Varaprasad G., "Violation Detection Method for Vehicular Ad Hoc Networking", ACM/Wiley Security and Communication Networks, 2014, <https://doi.org/10.1002/sec.427>.
4. Abdoos M., Mozayani N and Bazzan A.L.C., "Traffic Light Control in Non-Stationary Environments based on Multi agent Q-learning", in Proc. IEEE Conference on Intelligent Transportation Systems, pp. 580–1585, 2011.

5. ZigBee Alliance, "ZigBee Specifications", IEEE Standards 802.15.4k2013, <http://www.zigbee.org/Specifications.aspx>
6. "Traffic Congestion in Bangalore-A Rising Concern", <http://www.commonfloor.com/guide/traffic-congestion-in-bangalore-arising-concern27238.html>
7. Ayush K.R. Mittal, Deepika Bhandari, "A Novel Approach to Implement Green Wave system and Detection of Stolen Vehicles", in Proc. IEEE Conference on Advance Computing, pp. 1055–1059, 2013.
8. Suresh Sharma, Alok Pithora, Gaurav Gupta, Mohit Goel, Mohit Sinha, "Traffic Light Priority Control For Emergency Vehicle Using RFID", International Journal of Innovations in Engineering and Technology, vol.2, no.2, pp. 363–366, 2013.
9. Rashmi Hegde, Rohith R. Sali, Indira M.S., "RFID and GPS based Automatic Lane Clearance System for Ambulance", International Journal of Advanced Electrical and Electronics Engineering, vol.2, no.3, pp. 102–107, 2013.
10. Praveen Sood, "Bangalore Traffic Police-Preparing for the Future", <http://www.intranse.in/its1/sites/default/files/D1-S2->
11. http://www.bangaloretrafficpolice.gov.in/index.php?option=com_content&view=article&id=87&btp=87
12. Varaprasad G., "High Stable Power Aware Multicast Algorithm for Mobile Adhoc Networks", IEEE Sensors Journal, vol.13, no.5, pp. 1442–1446, 2013.
13. <http://phys.org/news/2013-05-physics-green-city-traffic-smoothly.html>

QR-Based Authentication for Login and Payment



J. Bibiana Jenifer, S. Sivaramakrishnan, Akhil Raula Satish, V. Preran, S. Chirag, and Shamolima Dutta

Abstract This chapter presents an authentication system that requires users to register and provide basic details before accessing the login module. Upon successful registration, users can proceed to authenticate their accounts by entering the email ID and password used during registration. Once the email ID and password are verified, users can proceed to the next authentication step, which involves selecting a QR (Quick Response) Code as the authentication method. The system generates a QR code and displays it on the user's phone or device. It is important to note that the QR code remains valid for only 30 s, after which a new and refreshed QR code is generated to enhance security. The effectiveness and security of this authentication system were achieved through the use of registration, email ID and password authentication, and time-limited QR code generation.

Keywords Authentication system · QR code · Security · Phishing attack · Geofencing

1 Introduction

User authentication and verification have become crucial in ensuring secure transactions and login processes. While methods like two-factor authentication (2FA) provide an additional layer of security compared with single-layer password authentication, they are not entirely foolproof. SMS-based one-time password (OTP) verification, a popular 2FA method, relies on the delivery of OTPs via SMS to verify users. However, this method has vulnerabilities, such as interception and phishing attacks, which raise concerns about its security [1–4].

To address these concerns, there is a need to either enhance the second layer of protection or introduce a third layer of security for logins and payments. In this

J. Bibiana Jenifer (✉) · S. Sivaramakrishnan · A. R. Satish · V. Preran · S. Chirag · S. Dutta
Department of Information Science and Engineering, New Horizon College of Engineering,
Bangalore, India

context, it is important to explore new technologies that have emerged alongside the resurgence of QR codes. These technologies include near-field communication (NFC), Bluetooth beacons, and geofencing.

NFC technology enables wireless communication between devices at short distances, typically less than 10 cm [5]. It is commonly used in mobile payment systems such as Google Pay and Apple Wallet, providing a secure means of exchanging data. NFC offers advantages over technologies like RFID, which are susceptible to attacks due to their longer communication range. However, NFC adoption is limited to NFC-enabled devices, unlike QR codes that can be scanned by any camera-enabled device.

Bluetooth beacons are small transmitters that broadcast data to nearby Bluetooth devices. They have been widely adopted in various industries, offering benefits such as indoor localization, proximity detection, and personalized experiences. BLE beacon-based services can provide contextual information and trigger events based on proximity to objects or places. While QR codes and NFC serve similar purposes, QR codes are more accessible to a larger audience because they can be printed or installed without specific hardware requirements [6–8].

Geofencing is a technology that defines a virtual boundary and triggers actions when a mobile device enters or exits the specified area. It utilizes GPS or Wi-Fi to deliver real-time messages or perform location-based actions. Combining geofencing with QR codes allows for dynamic redirection based on the user's GPS location, providing personalized experiences and targeted messaging.

These emerging technologies present opportunities to enhance authentication and verification processes beyond traditional methods. The subsequent sections will focus on QR code-based authentication and its integration with these technologies to augment security and user experience in login and payment systems.

2 Authentication Methods

2.1 SMS-Based Authentication Methods

In the registration stage of this method, a user registers a phone number and an account ID. A long-term password will also be set up by the user in order to generate a string of OTPs for subsequent logins. The user inputs his or her ID into an untrusted web browser to access an online service. The user then enters the long-term password after starting the pass app on their phone. The software will generate an OTP, which will be delivered straight to the server through SMS and used to verify the user's identity [9]. There are many methods where OTPs are transmitted to the user's smartphone through SMS, and similar methods have been proposed in other works [10, 11]. For many years, SMS security vulnerabilities have been widely and publicly discussed. Despite this, many businesses continue to rely on SMS for 2FA. The reason is that SMS authentication is simple to set

up and utilize. Furthermore, both customers and staff have been accustomed to utilizing it to obtain access to their numerous programmers, whether it is Slack or exchanging payments. End customers expect quick, frictionless authentication experiences and regard SMS as the ideal solution, without taking into account the security implications. If enterprises choose to abandon SMS authentication, they must find alternative options that are just as simple to use.

2.2 Token-Based

The use of security tokens is another approach to authentication. A physical object, such as a smart card or even a phone, might serve as a token. For instance, the RSA SecurID method uses a security token to create authentication codes on a regular basis depending on a starting seed value. It has been made clear that this strategy is not a defense against online phishing or session hijacking. For the protection of e-banking systems, a low-cost hardware token-based PIN/TAN system was proposed. To provide user, server, and transaction authentication, this hardware takes the shape of a physical USB token that must be plugged into an untrusted computer.

When opposed to more conventional approaches like using cookies, the use of tokens has various advantages.

- (1) Tokens have no state. The token is complete and has all of the data required for authentication. Being able to avoid having to store the session state on your server is great for scalability.
- (2) You can create tokens anywhere. Because token generation and verification are independent processes, you have the choice of managing token signing on a different server or even through a different business, like Auth0.
- (3) Granular access control. You may quickly establish user roles, permissions, and resources that the user can access within the token payload.

2.3 Connection-Based

A technique called MP-Auth uses a trusted personal device to carry out cryptographic calculations. It requires the user and server to share a long-lasting password that has already been established. Password information is entered into the personal device rather than the untrusted terminal to secure the password. Because cryptographic calculations carried out on the personal device are sent to the computer, which then passes them to the server, MP-Auth requires a connection between the computer and the personal device for this to happen. The Phool-proof phishing protection system combines public key cryptography, a username/password combination, and an SSL connection and is based on mobile

devices. A connection must be made in order for the strategy to function between the reliable cell phone.

Camera-Based

A camera-based authentication method that constantly monitors the user's interactions with an untrusted computer by watching the information shown on the computer's screen was previously proposed. The goal of the method is to determine whether the data on the computer's screen has been altered. This strategy can need quite a bit of computational work to perform the necessary monitoring. There are two visual cryptographic shares created from the secret authentication message. By superimposing one share on the mobile device's screen with the second share that is shown on the computer's screen, the user can visually obtain the secret using the camera on the mobile device. A commercial transaction authentication system is Conto.

2.4 QR Code-Based

QR codes can be read by mobile devices equipped with a camera and the appropriate scanner. The Mobile Media API, or MMAPI, must be implemented by the mobile device in order to take pictures in Java ME. Snap2pass and its extension Snap 2pay are QR code-based methods that require a camera-equipped smartphone to be connected and active. The server sends the browser a QR code that contains a cryptographic challenge in order to log the user into a website. The user must utilize their smartphone's camera to take a photo of the QR code. The mobile device executes the required cryptographic calculations and then immediately sends a cryptographic answer back to the server. The server then logs the user in via the browser after obtaining a legitimate response. Due to the tremendous development of smart devices being used by the general public, QR codes are becoming more and more popular all over the place. In comparison to the 1D QR code, the 2D QR code is undoubtedly superior and can contain a lot more encoded information.

People use smartphones to perform authentications and the most common type of optimal method. Nowadays, a variety of QR code formats are becoming more and more common, including logo QR codes, encrypted QR codes, iQR codes, and so on. The following generation is embracing QR codes more and more because they provide simpler authentication than traditional methods.

3 Existing System

Traditional 2FA systems use authenticators, which generate codes, similar to OTP; however, this may be more effective than using a one-layered protection. Take Gmail for instance, whenever there is a login attempt from a new device or browser to a

Gmail account by a user who has opted in for 2FA, an OTP is generated in addition to the login details and sent to the owner of the account via a call or text for them to verify the login.

4 Proposed System

In the proposed scheme, the user can easily and efficiently log in to the system. We analyze the security and usability of the proposed scheme and show the resistance of the proposed scheme to hacking of login credentials, shoulder surfing, and accidental login. The adversary to obtain the user's password by watching over the user's shoulder as he enters his password can perform the shoulder surfing attack. After entering the login credentials, the user is prompted to scan a QR code, which is available only for 30 s. The scanner is present only on the user's phone, which the user then uses to scan the QR code that is generated.

4.1 Basics of QR Code

A QR code is a two-vector barcode. More than one vector barcode may be maintained [12]. Therefore, it is advised that QR codes require more advanced reading equipment. The ISO/IEC 18804 Industrial Specification defines QR codes. Denso Wave Japanese Corporation 1994 did create and patent it. The primary goal of developing this technology is to assist users in quickly encoding and decoding their own information. Fig. 1 represents the QR code.

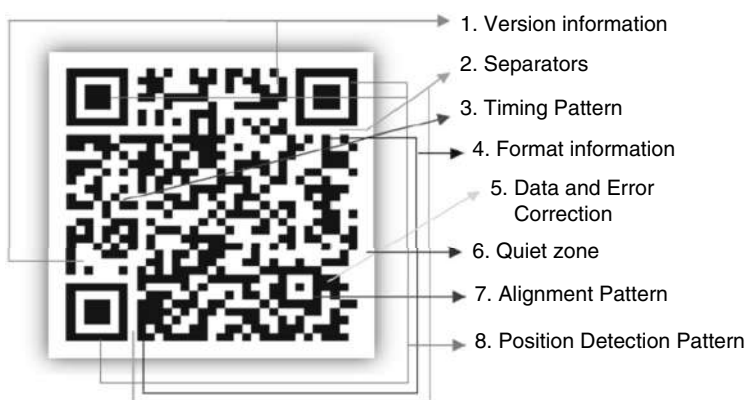


Fig. 1 QR code

Advantages of Applying QR Code

- Open-source techniques
- Free applications
- Simple running procedure
- User easy process

To utilize a QR code, no complicated software or specialized user knowledge is required, unlike the activation of a QR code that requires a smartphone. It is also being utilized to make delivery services offered by informatics libraries available, in addition to QR code scanners. Additionally, it is important to share information and the most recent technological expertise with all users. In addition to offering instant access to the library’s other important assets and saving users’ time, it also offers limited storage space that may hold a significant amount of material.

Figure 2 illustrates the step-by-step process involved in the QR code-based authentication system. This section will present a detailed flow diagram, outlining the sequence of actions performed by the various components.

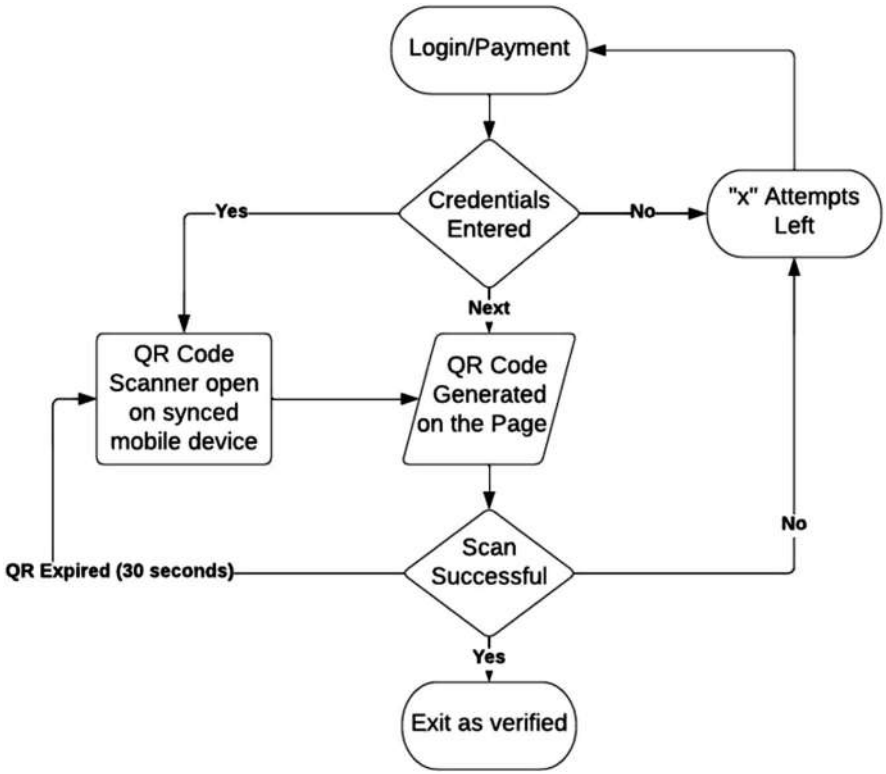


Fig. 2 Flow chart

The authentication process starts with the user launching the mobile application and initiating the QR code scanning functionality. The mobile application accesses the device's camera and prompts the user to scan the QR code displayed on the service provider's interface. Upon scanning the QR code, the mobile application extracts the encoded data and securely transmits it to the authentication server. The authentication server receives the QR code data and performs the necessary decryption to extract the authentication details, such as the user ID and the requested service. The authentication server then validates the received information, checking if the user exists in the system and if the requested service is authorized for the user. If the validation is successful, the authentication server generates a unique session identifier and associates it with the user's session. Next, the authentication server sends a response back to the mobile application, indicating the success or failure of the authentication process. In the case of a successful authentication, the mobile application grants access to the requested service, allowing the user to proceed. If the authentication fails, appropriate error messages are displayed to the user. Throughout the entire process, various security measures, such as encryption and digital signatures, are employed to protect the integrity and confidentiality of the data transmitted between the mobile application and the authentication server.

5 Conclusion

The development of an electronic verification system using the QR system has successfully achieved the goal of enhancing financial authenticity. By extracting data from the official book's content and encrypting it to ensure anonymity, the system has effectively constructed official books and generated unique QR codes for each of them. This process is facilitated through a reliable service provider that offers QR code services. Prior to transmission, the QR codes are encoded and securely stored in a dedicated database.

In this study, an additional layer of security was implemented by encrypting the book data before transforming it into QR codes. This measure serves to prevent fraud and impersonation, while also safeguarding the confidential information used in QR code encoding. Consequently, the approach adopted in this research offers a high level of safety and security.

Moving forward, the developed electronic verification system presents numerous opportunities for further exploration and improvement. One potential avenue for future work is to integrate the system with existing financial platforms or databases to streamline the verification process and enhance its efficiency. Additionally, conducting rigorous testing and evaluation of the system's performance in real-world scenarios would contribute to its robustness and reliability. Furthermore, exploring the possibility of expanding the system's application beyond financial authenticity, such as in areas like product verification or document authentication, could broaden its potential impact.

Overall, the successful implementation of the electronic verification system using QR codes represents a significant step toward ensuring trust and integrity in financial transactions. By continuing to refine and expand upon this technology, we can forge a path toward a more secure and reliable digital ecosystem.

References

1. Shanmugam Shobha, M., & Rekha B Venkatapur, (2020). Enhancement of Signature Schemes for Heightening Security in Blockchain. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* © 2019 IJSRCSEIT, 9(4), (645–656).
2. Karthiyayini, J., (2020). Robot Assisted Emergency and Rescue System with Wireless Sensors. *International Journal of Research and Scientific Innovation*, 12(4), (28-32).
3. Vandana, C.P., & Abir Bhartacharjee, & Yashwant Pandit, & Rakshith, P., & Peshal Parajuli, (2020). Enhanced Cloud Security based on DNA Cryptography. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 5(7), (2477–2486).
4. Nivetha, M. B., & Sivaramakrishnan, M. S. (2014). A comparative analysis of cryptography algorithms. *Wireless Commun*, 6(8), 304–307.
5. Coskun, V., Ok, K., & Ozdenizci, B. (2011). *Near field communication (NFC): From theory to practice*. John Wiley & Sons.
6. Vibha, T. G., and S. Sivaramakrishnan. “A Survey of Deep Learning Region Proposal and Background Recognition Techniques for Moving Object Detection.” *Computer Networks and Inventive Communication Technologies: Proceedings of Fifth ICCNCT 2022*. Singapore: Springer Nature Singapore, 2022. 147–164.
7. Prashanth, P., Anandhi, R. J. (2019). A Review on Data Science Approach Towards Decision-Making. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(4), 692–697
8. Dey, Somdip, et al. “SmartNoshWaste: Using blockchain, machine learning, cloud computing and QR code to reduce food waste in decentralized web 3.0 enabled smart cities.” *Smart Cities* 5.1 (2022): 162–176.
9. Devi, B. K., Vijayakumar, V., Suseela, G., Kavim, B. P., Sivaramakrishnan, S., & Rodrigues, J. J. (2022). An improved security framework in health care using hybrid computing. *Malaysian Journal of Computer Science*, 50-61.
10. Sivaramakrishnan, S., and T. Kesavamurthy. “Identifying Cluster Head and Data Transmission Through Them for Efficient Communication in Wireless Sensor Network.” *Journal of Computational and Theoretical Nanoscience* 14.8 (2017): 4014–4020.
11. Sivaramakrishnan, S., et al. “Augmentation of Terahertz Communication in 6G and Its Dependency for Future State-of-the-Art Technology.” *Challenges and Risks Involved in Deploying 6G and NextGen Networks*. IGI Global, 2022. 91–105.
12. Huang, P. C., Li, Y. H., Chang, C. C., & Liu, Y. (2019). Efficient QR code authentication mechanism based on Sudoku. *Multimedia Tools and Applications*, 78, 26023–26045.

Smart Irrigation Watering System Using IoT



K. Krishna Reddy, G. Faazil, K. Ajith, C. Pavani, J. Sai Tharun,
D. Dhanush Gowdu, and T. Sravani

Abstract With the ‘smart irrigation system’, you can automate your whole irrigation system in addition to its other features. This Internet of Things-based irrigation system is being built with the help of an ESP8266 Node MCU Module and a DHT11 Sensor. In addition to automatically adjusting watering schedules to match soil moisture levels, this system will also send data on the land’s condition to the Thing Speak server. A water pump will be part of the system, and it will be used to sprinkle water on the ground in a way that is dependent on the moisture level, temperature and humidity level of the land. Such a crop will need a soil moisture of roughly 50–55%, and we have previously built a system that is quite similar to an ‘automatic plant irrigation system’. This system delivers notifications on mobile devices, but not on the Internet of Things cloud. Therefore, when the moisture content of the soil decreases to less than 50%, in order to sprinkle the water, the motor pump will switch on automatically, and it will do so until the moisture content reaches 55%, at which time it will shut off. Data from the sensors will be sent to the Thing Speak server at regular intervals, allowing it to be monitored from anywhere in the world. In addition to this, we are integrating machine learning into this process in order to forecast the future scope by making use of historical data and the random forest method.

Keywords Node MCU · Thing Speak · DHT11 · Irrigation · Humidity · Soil moisture · Temperature

K. K. Reddy (✉) · G. Faazil · K. Ajith · C. Pavani · J. S. Tharun · D. D. Gowdu · T. Sravani
Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering
and Technology, Chittoor, India
e-mail: krishnareddy206@mtieat.org

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,
Springer Proceedings in Mathematics & Statistics 441,
https://doi.org/10.1007/978-3-031-51338-1_39

497

1 Introduction

The evolving paradigm for smart agriculture might be an observation system that operates in real time. It keeps an eye on the temperature, humidity, soil moisture and pH levels, among other aspects of the soil. A farmer may reduce his or her water use by as much as 51% just by adopting the modern irrigation practices of today. This approach relies on two watering tactics such as above sprinklers and flood type feeding systems. The Internet of Things (IoT) makes it possible to remotely control many business processes within a sector from any location at any time. In order to overcome these problems, new approaches have been implemented within the field of irrigation techniques. These new techniques include the application of small quantities of water to the components of a plant's foundation zone. The plant soil wetness stress may be avoided by supplying the required amount of water resources often or generally everyday so that the wetness state of the soil can maintain a well diagram. This ensures that the plant soil can continue to function normally. The typical methods, such as mechanical devices or surface irrigation, need almost 50% of the available water sources. It is possible to give plants with very precise amounts of water in large quantities. It is possible for the irrigation system to produce continuous federations due to the dry rows that are in between the plants. This kind of technology allows for the application of fertilizers, which results in a reduction in the amount of money required for the process. In comparison to systems that use overhead mechanical devices, more contemporary solutions significantly lessen the amount of soil and wind erosion that occurs. The properties of the soil might provide an outline of the form of the dripping nature that exists inside the root zone of a plant that is provided with moisture. It gives an art movement way of life in which a person gets to handle his electrical gadgets by using a smartphone, and it also offers an inexpensive use of energy. These benefits may be enjoyed simultaneously. It applied to all aspects of business, such as intelligent agriculture and intelligent parking, making use of IoT for the monitoring of sensor data and random forest algorithm for achieving the highest possible accuracy in crop predicting.

The majority of irrigation systems in India are still managed manually using conventional techniques since agriculture is the country's primary source of revenue and also the primary source of livelihood for 70% of the population. The existing method of irrigation does not make effective use of water and energy, which are both scarce and very precious resources in the field of irrigation. Because of the contributions of technology to today's modern civilization, which has been transformed into a digital world, we are now leaving behind an age in which technology is researched in order to better our way of life. As a result, the 'smart watering system' was developed so that people's lives might be made easier and more convenient. The development of a standardized method for regulating irrigation might be of assistance to millions of people.

The science of the artificial application of water to soil, with consideration given to the amount of moisture already present in the soil, is what is meant by the term 'smart watering system'. It is now possible, thanks to the development of

open-source Arduino boards and the introduction of a moisture sensor, to design devices that can monitor the moisture content of the soil and, based on that information, irrigate the crops or the landscape when it is necessary to do so. The suggested system takes use of the microcontroller ATMEGA328P on the Arduino Uno platform and the Internet of Things. This allows farmers to remotely monitor the state of the water level in the soil by knowing the sensor data. As a result, this makes the farmers' job much simpler, and they can focus on other agricultural operations.

2 Literature Review

- A drop in agricultural production is predicted by the Climate Change Agricultural Yield Assumptions research in ref. [6]. Lower growth rates are also expected in several developing regions, such as Southeast Asia (−5%) and India (−5%).
- There are many potential explanations for the geographical variation in farm losses, including problems with infrastructure and marketing, bad timing of harvests, extreme weather events, and a lack of foresight in selecting crops well suited to the local climate. Here is a look at the contrast that follows:

Kohonen Self-Organizing Map (SOM) and Back Propagation Network (BPN) are examples of unsupervised and supervised learning techniques used by Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, and Priya R. L. soil type predictions. The process begins by training a dataset using these learning networks and then labeling the dataset as organic, inorganic, or real estate. The system compares the accuracy of several network learning techniques and returns the one with the highest score to the user. When soil data is input into the system, it analyzes the data, predicts crop yield, and recommends the optimal amount of fertilizer to use.

India is one of the world's 13 driest nations but utilizes only a small portion its potential water supply. As a result, the country faces the risk of becoming too hot to inhabit. This paper proposes a precision agriculture irrigation system focusing on the hardware design, specifications, and code management of the system, tailored for modern agricultural cultivation and management. The goal is to mitigate the economic impact caused by India's insufficient water resources, while conforming to web standards and integrating mobile device technologies.

The current situation, which includes falling water tables, drying rivers and tanks, and an unpredictable climate, presents an urgent need for efficient water usage. Pratyusha and Chaitanya Suman [1] implemented temperature and moisture sensors in strategic areas to monitor crops, helping combat these issues and ensure agricultural success. A microcontroller-based gateway can be programmed with an algorithm that utilizes threshold values for temperature and soil moisture to regulate water usage [2]. Research in agriculture shows that crop yield is gradually declining over time. However, technology plays a significant role in enhancing overall productivity and reducing labor requirements. Several research

efforts have been aimed at improving conditions for farmers by providing systems that technology to boost agricultural productivity [3]. The need for automated water control is addressed through the solutions that enable remote switching and monitoring of irrigation systems via smartphones. This system provides users with critical information such as moisture levels, temperature, and humidity on their smartphones [4]. Researchers studying wireless sensor networks (WSNs) analyzed soil factors like temperature and humidity. Sensors were buried beneath the soil and connected with relay nodes using an efficient communication protocol. This protocol has a low duty cycle, increasing the lifespan of the Soil monitoring system. A microcontroller and sensors were used to collect hourly data samples, buffer the data, transfer it, and regulate water provision [5]. This prototype measures soil moisture, with acceptable ranges depending on the type of crop or soil. If the soil moisture content falls outside the acceptable range, the irrigation system will automatically activate or deactivate. Each time the irrigation system starts or stops, it pumps water to the plants and sends a message to the user via a GSM module. The message provides an update on the water pump's status and the soil's moisture content.

3 Conclusion

- In this work, we provide an original strategy for smart agriculture that makes use of two developing technologies: the Internet of Things and machine learning. The accuracy of the outcome may be helped along with its pursuit of precision by making use of both live and historical data.
- The accuracy of the system may also be improved by comparing several machine learning methods. This technology will be utilized to lessen the obstacles that the farmers have, and it will boost both the amount and the quality of the work that the farmers do.

Future Scope

- The system is expandable, and thus the following features may be added in the future: soil moisture sensors, environmental sensors and PH sensors are advised to increase precision in crop forecasting.
- Locations' market needs may be taken into consideration, and neighbouring farmers' crops can be recommended while providing guidance on the appropriate crop.

References

1. K. Prathyusha, M. Chaitanya Suman, "Design of Embedded System for the Automation of Drip Irrigation". IJAIEEM (2319–4847), vol 1, Issue 2, October 2012.
2. Sirsath N. S, Dhole P. S, Mohire N. P, Naik S. C & Ratnaparkhi N.S, "SMART AGRICULTURE USING Cloud Network and Mobile Devices".

3. Prathyusha K, G. Sowmya Bala, Dr. K. Sreenivasa Ravi, "A real time irrigation control system for precision agriculture using WSN in Indian agricultural sectors" International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol. 3, No. 4, August 20, 2013.
4. Wang N, Zhang N P, Wang M H. Wireless sensors in agriculture and food industry-Recent development and future perspective[J]. Computers and Electronics in Agriculture, 2006.
5. Akyildig, I.F., 2005. A Survey on Sensor Networks [J].IEEE Communications Magazine, 2002(8):725–734.
6. Liu Hang, Liao Guiping, Yang Fan. Application of wireless sensor network in agriculture producing [J]. Agricultural Network Information, 2008.

IoT-Based Transmission Line Multiple Fault Detection and Indication to Electricity Board



K. Jeevan Reddy, K. Aruna, K. Manoranjitha, M. Bhavani Sankar,
C. Ravindra, and G. Sireesha

Abstract This project addresses a novel approach to detecting numerous faults as well as automatic switching based on the Arduino display that is superimposed over the faults. The passage of electricity via wires is an important consideration for our detecting method. The maximum amount of current that may be carried by each cable is specified. The current abruptly rises if there is a failure in a short circuit. In the event that the circuit is open, there will be no flow of current. In order to determine the amount of current, current transformers are used; the output current of these transformers is then sent to an I-to-V converter unit so that the information may be read in terms of voltage. After that, the voltage is sent to the ADC pin of the Arduino, which converts it into digital form and takes the necessary corrective action if any fault situation, either SC or OC, happens. This error message will appear on both the LCD display and the LEDs. When a fault occurs at this moment, a buzzer will indicate it, and the relays will go into disconnected mode. The state of the loads will become OFF. Once the fault has been cleared, the relays will switch to an automatically connected mode. When a defect occurs at the appropriate moment, the authority will get a notice or a text message.

Keywords IoT · PIC · NODEMCU · ADC pin · Arduino

1 Introduction

At the moment, the human life cycle and the process of living day to day are both significantly impacted most significantly by electricity.

Transmission line faults may lead to a variety of problems, including mechanical failure, high internal and external strains, lightning strikes and many more. In transmission lines, common types of faults include three-phase line faults to other

K. J. Reddy (✉) · K. Aruna · K. Manoranjitha · M. B. Sankar · C. Ravindra · G. Sireesha
Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering
and Technology, Chittoor, India

lines, three-phase line faults to ground and double-phase line faults to ground. Generation, transmission and distribution are the three primary categories that are used to classify the power system. The transmission system is regarded to be the most important component of the power system. The present situation is characterized by a great deal of both supply and demand for transmission networks. Because of the malfunction in the transmission line, there was a disturbance in the delivery of electricity to the customer. There are a few advantages to having an accurate defect detection system, such as reducing the amount of time needed to conserve power, reducing the amount of money spent on maintenance and preserving the life of the electrical equipment. At this point in India's history, if a system does not identify a defect in the system in real time, then it is a very dangerous sort of system to have in a transmission network.

Generation, transmission and distribution are the three subsystems that make up an electrical power system. Because it links the supply with the demand, the transmission network is widely regarded as being one of the most important components of the power system. When compared to losses in other components of the power system, the transmission and distribution network is regarded as having an exceptionally high percentage.

At the moment, the infrastructure that handles electric power is very susceptible to a wide variety of natural and man-made physical events, each one of which has the potential to have a negative impact on the overall performance and stability of the grid. Because of the problems in the transmission network, the customer is unable to get the electricity that they need. If there is a problem with the transmission line, it will not often be discovered until the problem has become serious. However, over time, these seemingly insignificant faults can cause significant damage to the transformer, which in turn can threaten human life. It is also capable of starting a fire.

At the moment in India, we are not told of the defect in real time when a failure happens since there is not yet a technology that can do so. The absence of a real-time system should be cause for concern since it may result in disruptions to the underlying linked equipment and pose a risk to the people who are in the immediate vicinity. Transmission lines are often subjected to routine maintenance and inspections on a fairly regular basis in order to reduce the likelihood of occurrences like this to the greatest degree possible. Because of this, there is a higher need for human labour. The reality of the matter is that the actual purpose of this is not fulfilled since there are numerous instances in which line failure may be caused by rain or the falling of trees, both of which cannot be foreseen. As a result, the process of problem diagnosis and clearing on the transmission network has to be quite quick.

In order to circumvent these challenges, it has been recommended that this project implement an Internet of Things-based transmission line multiple fault detection and indication to EB. When the predetermined threshold is exceeded, the microcontroller immediately commences the sending of a message to the local lineman as well as the control station, in which the precise distance from pole to pole is stated. The realization of the real-time system is facilitated as a result of this.

Because of this, the reaction time for the technical personnel to correct these faults will be shortened, which will assist save transformers from damage and catastrophes. In order to sufficiently and precisely signal and determine where the problem has occurred, a fault detection and localization system that is smart and based on GSM technology is used. A current transformer, a voltage transformer, a PIC 16F877 Microcontroller, an RS-232 connection and a GSM modem are all components that make up this system.

An impedance-based algorithm is used by the system to automatically identify problems, then analyse and categorize them and finally compute the fault's distance from the control room. All of these steps take place automatically. Finally, the information regarding the fault is communicated to the control room by means of IoT technology.

The network of the electric power system is made up of many different sections, and transmission lines play a significant part in the process of moving electricity between different points. The infrastructure that controls the flow of electric power is very susceptible to a wide variety of natural and man-made physical events, any of which may have a negative impact on the overall performance and stability of the grid.

In addition, there is an imminent demand to retrofit the ageing transmission line infrastructure with a high-performance data communication network that is capable of supporting future operational requirements such as real-time monitoring and control, which is essential for the integration of smart grids. Circuit indications have been the primary method that the majority of electric power transmission companies have used to identify problematic portions of their transmission lines.

2 Literature Survey

[1] This paper present a new algorithm for fault detection, classification and location of overhead transmission line using Wavelet Transform (WT) based Discrete Wavelet Transform (DWT) is proposed. The different system faults such as LG, LLG and LLLG in transmission line should be detect, classify and locate rapidly. The proposed method is based on the voltage and current signal information from the power model in MATLAB to generate the transient voltage and current signal in both time and frequency domain. DWT using “db6” as a mother wavelet is used to capture transient current signals and extract the high frequency detail coefficient for detecting and classifying the fault disturbance. The classification process is based on ground threshold value The location of faults is carried out by obtaining the fault information from source terminal end to remote terminal end along with the total transmission line length. The proposed method is tested on the MATLAB/SIMULINK environment with the Simulink model. In simulation process the proposed algorithm achieved the fault detection, classification and locate all eleven types of possible fault in transmission line and the result are compared with the AR and MED method. [2] In the modern scenario, due to increase in

demand and supply across the world, the field of power system is taking its turn and advancements are being introduced day to day. Protection against the transmission line fault is an important issue in electrical power system, because most of the faults occur in transmission line. This paper deals with the classification and detection of fault in a three phase transmission line using Discrete Wavelet Transform (DWT). Different types of faults are created in a two-bus system at different locations using MATLAB and fault currents in all the phases are acquired. Then, DWT is applied on the faulty signals and detailed coefficients are obtained, which in turn are used in the algorithm to classify the fault in transmission line. [3] There are numerous distinct elements that make up the electric power system. One of these is the transmission system, in which power is delivered to consumers via transmission lines from generating plants and substations. Both approaches might experience various malfunctions, which are known as “Faults.” When the insulation of the system fails at any point, a fault is simply described as a collection of unfavourable but unavoidable happenings that can momentarily upset the stable condition of the power system. A short circuit, or fault, is also said to have happened if a conducting object makes touch with a bare power conductor. The reasons of faults are numerous and include lightning, wind damage, trees encroaching on transmission lines, and collisions between cars or aeroplanes. Poles or transmission towers, birds cutting power wires, or acts of vandalism. The purpose of this study was to examine the causes and consequences of defects in overhead transmission lines. We’ll talk about a few of the numerous defect sources and fault detection techniques. The equipment in the power system is severely damaged as a result of these problems. In India, it is typical for the supply systems to have faults that are LG (Line to Ground), LL (Line to Line), or 3L (Three Lines). These faults in the three-phase supply system might impact the power system. [4] This project addresses the need for a Power grids suffer frequent faults, highlighting the need for improved transmission line protection. This paper proposes an Arduino based fault detection system for enhanced reliability and safety. The system utilizes current transformers to sense current variations and potentially impedance measurement for fault location. It also integrates fire sensors and sound sensors for comprehensive fault detection. Upon fault detection, the Arduino triggers alarms, displays status on LCD, and can send SMS alerts. This system offers rapid fault identification and response, reducing downtime and maintenance costs Furthermore, the Arduino’s low cost makes this a feasible solution for improving power transmission line reliability. [5] Transmission lines are prone to a wide variety of faults due to transmission lines conditions. Diagnosing the fault is difficult and entire cable should be replaced. This project is intended to detect the location of fault in transmission line cable lines from the base station in km using a microcontroller. In case of fault, the voltage across series resistors changes accordingly, which is then fed to an ADC to develop precise digital data to a programmed. It further displays fault location

in distance. Using GPS, location can be tracked. The fault occurring distance, phase, and time is displayed on a 16X2 LCD interfaced with the microcontroller. IOT is used to display the information over Internet using the Wi-Fi module. A webpage is created using HTML coding and the information about occurrence of fault is displayed in a webpage. [6] This research work proposes a novel approach for fault detection in transmission lines using IoT (Internet of Things) technology. The proposed system consists of sensors deployed along the transmission lines to monitor the line parameters such as voltage, current, and temperature. These sensors capture data, which is then sent to a cloud-based platform. The system is designed to detect and locate faults such as line breakages, short circuits, and overloads in real-time. The suggested method has a number of benefits, including increased problem detection precision, less downtime, and increased system dependability. The results of the study demonstrate the effectiveness of the proposed system in fault detection and localization, making it a promising solution for the power transmission industry. [7] Fault location Algorithm is an important task in power system engineer because accurate faults location algorithms on power transmission lines need to be detected and located rapidly, classified correctly and cleared as fast as possible to restore the supply back to the system, so as to avoid possible damages to people, property and environment. In this paper review of impedance-based fault location algorithm in electrical power transmission line, which include, the one-ended impedance algorithm such as (Takagi method, Modified Takagi method, Erikson method, Novosel method) and the two-ended impedance based algorithm such as (synchronized two-ended method and unsynchronized two-ended method). Fault data availability forms a basic for choosing the most suitable fault location algorithm. The result of this paper shows that the simple reactance method is the simplest and low cost of the entire impedance - based fault location algorithm. The fault resistance and the load current make the algorithm to deteriorate in accuracy. The Takagi method was developed to correct these lapses because of its robustness to load and sensitivity. Source impedance parameters were used by the Modified Takagi and Erikson methods for elimination of any source of errors. The mutual coupling in double-circuit transmission lines and an uncertain value of zero sequence line impedance are responsible for the inaccuracy of one-ended impedance-based technique. This problem was overcome by two ended fault location algorithm that used voltage and current measurements captured from both ends of transmission line and are attractive for tracking down the specific position where the fault is to be located. Availability of data is one of the most important factors in selecting the best technique for fault location. Additional equipment needed to be installed for improving the accuracy of fault location algorithm is very useful. This paper will form a basis for choosing an appropriate fault location technique for electrical power transmission network.



Fig. 1 Hardware configuration with operational status

3 Simulation

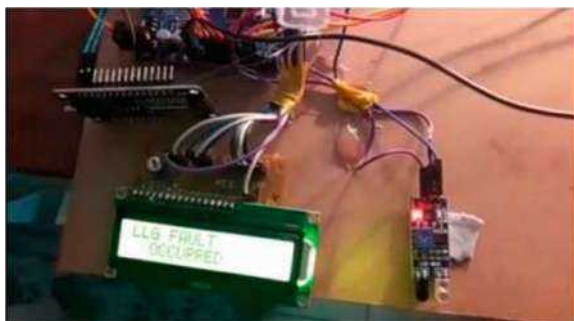
Hardware System and Results

The model has been developed, and the generated data have been thoroughly analyzed. The system consists of several physical components, including a transformer, Arduino, GPS module, gearbox lines, and other essential hardware components, as depicted in Fig. 1. The transformer regulates the voltage levels for the system, ensuring that the components receive the appropriate power supply. The Arduino acts as the central control unit, processing data from the sensors and executing commands based on the programmed algorithms. The GPS module provides real-time location tracking, which is crucial for applications like navigation or asset monitoring. The gearbox lines are responsible for mechanical functions, controlling movements or transmission of power within the system.

Each of these components interacts with one another, forming a cohesive hardware setup that ensures smooth operation. The data collected from the sensors are processed by the Arduino, and based on this information, the system makes decisions such as activating motors, adjusting positions, or transmitting data to a remote server for further analysis. The results are monitored and recorded to evaluate the system's performance under different conditions.

4 Results

The findings have been validated, and they demonstrate that if there is a defect in the transmission line, the Arduino programmer will pick up on it, the GPS will track its position and the LCD will reveal what kind of fault has been identified. This is shown in Fig. 2. When the fault occurred, the relay was able to detect it, and as a result, the load on the line that was affected was turned off. If there is a ground fault in the system, the load in its whole will become disconnected. As soon as the error was fixed, the system returned to its usual functioning.

Fig. 2 Result

5 Conclusion

The planned project will take on all of the weaknesses of the current infrastructure. We have created a GSM-based transmission line monitoring and indication system for this project. This system provides information about the transmission lines to the control room through text message. The distribution system takes up the majority of the focus of the system design that was ultimately implemented. It gives the means to identify issues such as the squandering of energy and the theft of electrical power.

The system performs ongoing monitoring of a wide variety of the system's parameters. Additionally, it helps to detect the fault at the appropriate time, which prevents the unauthorized consumption of electrical power. Through the use of hyper terminal, automatic monitoring, analysing and recording are carried out on the computer screen. The project incorporates both GSM communication technology and microcontroller technology into its continuous monitoring system. Additionally, it represents the architecture of the hardware as well as the flow of the software.

Because of the successful deployment of the system, a significant quantity of power will be saved, and as a result, electricity will be made accessible to a bigger number of customers in a nation with a very high population density such as India.

References

1. A. D. B. "Transmission Line Fault Detection Using Wavelet Transform". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 2, no. 10, Oct. 2014, pp. 313842, <https://doi.org/10.17762/ijritcc.v2i10.3364>.
2. Fault Detection and Classification in Three Phase Transmission Line Using Signal Processing (Dr. D. Kavitha, May 2018)
3. Chandrashekar. P "Transmission Line Fault Detection & Indication through GSM" ISSN (Online): 2347-2812, Vol 2, Issue 5, 2014.
4. Hui Hwang Goh, Syyi Sim, Asad Shaykh, Md, Humayun Kabir, Chin Wan Ling, Qing Shi Chua, Kai Chen Goh "Transmission Line Fault Detection" ISSN: 2502-4752, Vol 8, <https://doi.org/10.11591/ijeecs.v8.i1.pp199-205>, 2017.

5. Arti Sanganwar, Kapil Chalkhure, Shivani Jijankar, Anand Dhore, Prof. Vikramsingh R. Parihar "Transmission Line Multiple Fault Detection" ISSN: 2393-8374, (Online): 2394-0697, Vol 4, Issue 10, 2017.
6. Navaneetha Krishna R, Niranjan L, Shyamsundar. N, Dr. C. Venkatesan "IoT Based Transmission Line Fault Monitoring System" IJRAR August 2020, Vol 7, Issue 3, E-ISSN 2348-1269, P-ISSN 2349-5138, 2020.
7. Swagata Das, Surya Santoso, Anish Gaikwad, Mahendra Patel "Impedance Based Fault Location in Transmission Networks." <https://doi.org/10.1109/ACCESS.2014.2323353>, Vol 2, 2014.

Design of Off-Board Electric Vehicle Charger Using PV Array Through MATLAB-Simulink



N. V. Kishore Kumar, K. Nithya Sri, S. Akram, O. Praveen Kumar, V. Tharun, and G. Jaswanth

Abstract The development of electric vehicles (EV) during the last 10 years has made a major contribution to the growth of the automotive sector. The technique that is used to charge the batteries of electric cars will have a major impact on their development. When the battery of an electric vehicle is charged from the grid, there is an increase in the load demand for that grid. As a consequence of this fact, the authors of this research propose an off-board method for charging the batteries of electric vehicles that is founded on photovoltaic (PV) arrays. The battery of the electric car has to be constantly charged, and this is true regardless of the quantity of available sunshine. The EV battery may be charged by the system that has been designed both during the daytime hours and during the hours in which there is no sunlight due to the SEPIC converter and the three-phase bidirectional DC–DC converter. Both of these converters are capable of helping the system. When there is an abundance of sunshine, both the electric vehicle's battery as well as the backup battery are concurrently charged, and when there is an absence of sunlight, the EV battery is charged with the assistance of the backup battery. In order to do simulations of the proposed charging system, the MATLAB software, which comes packaged with a programme called Simulink, is used.

Keywords Electric vehicle · DC–DC converter · SEPIC converter · Electric grid · PV array · Li battery

1 Introduction

Concerns about the environment are caused by the ever-increasing impact of greenhouse gases produced by traditional internal combustion engines. Because of this, the market for pollution-free electric vehicles (EVs) has exploded in the

N. V. K. Kumar (✉) · K. N. Sri · S. Akram · O. P. Kumar · V. Tharun · G. Jaswanth
Department of Electrical & Electronics Engineering, Mother Theresa Institute of Engineering & Technology, Chittoor, India

automotive sector. However, charging electric vehicle batteries through the utility grid contributes to an increase in the load demand on the grid. This, in turn, results in higher electricity bills for owners of electric vehicles, which compels them to seek out alternative sources of energy. It is possible to utilize renewable energy sources (RESs), which do not produce pollution and do not run out, to charge the battery of an electric vehicle [1–4]. Therefore, an electric vehicle powered by Renewable Energy Sources (RES) can be representing as ‘green transportation’. Solar Energy is one of infinity potential energy resource in nature that can be used to energy conversion into various potential energy transformation, so that it may be used to power electric vehicle [5–8].

As a result, the electricity from the PV array is what is utilized to charge the EV battery in the system that is being suggested, and this is done with the assistance of various power converter topologies. Because of its high power density, high efficiency, low weight, and small size lithium-ion batteries are often utilized in electric vehicles. In addition to this, these batteries are capable of receiving a charge quickly, have a long lifespan and have a low rate of self-discharge. They also have a minimal potential for explosion in the event that they are overcharged or improperly connected [9]. These batteries need for very accurate voltage management when they are being charged. In order to charge an electric vehicle’s battery, a variety of power electronic converters with a voltage controller are used. Because the PV array only produces power on an as-needed basis, you will need power converters in order to keep the electric vehicle’s battery charged [10].

Electric vehicle loads (such as the motor, lights, power windows and doors, radios, amplifiers and mobile phone chargers) can be connected to photovoltaic arrays, ultra capacitors, super capacitors, fuel cells and batteries via multiport converters (MPCs). For this reason, hybrid electric cars often employ multiport converters for their onboard chargers [11]. As a result of the fact that all of the sources are housed inside the EV itself, MPCs have the disadvantage of making electric vehicles heavier, more expensive and more difficult to maintain. In these converter-based electric vehicle battery charging systems, the complexity of the controller implementation also increases. By the year 2050, according to many forecasts of the future, all electric vehicle fleets will be powered by renewable energy sources (RES). In fact, the rising acceptance and usage of electric vehicles are the result of advances in battery technology as well as the growth of battery charging facilities in an effort to supply the energy requirements of these vehicles. As a result, the basic infrastructure of the charging system is an important component in the marketing of electric vehicles [12]. However, the most significant drawback of charging infrastructures for electric vehicles is that their operation is not kind to the environment because they are powered solely by the electrical grid. In fact, renewable resources are a time-limited and divisible source of energy, but the charging of electric vehicles may be managed; therefore, it logically depends on the combination of renewable resources and EVs [13].

In light of this, it is vital to strike a balance between the generation of power and the charging of electric vehicles in order to ensure and maintain the secure continual functioning of the grid. It is generally agreed upon that the unpredictability

of the output of RES is one of the most significant challenges that must be addressed before the energy grid can continue to function effectively in the foreseeable future. Traditionally, load fluctuation control has not shown to be a successful method for balancing the grid, carrying out operational strategy, or controlling electricity under a variety of load operating states. This is because load fluctuation control focuses on maintaining a constant load [14]. In general, load scheduling as a possible solution may be achieved by managing the progression of RES generation; this has been recommended as a solution since scheduling additional electricity production is essential for the running of power systems [15]. In addition, EVs have shown that they are capable of assisting the primary grid in the maintenance of a certain balance between demand and supply; this raises the possibility for the penetration of RES. In fact, a wide variety of published research publications, such as those in, have addressed this topic [16]. In addition, the production of PV might result in an increase in the number of electric cars on the road. This is due to the fact that the energy needs of these automobiles do not generate a significant increase in the overall load. However, the integration of EVs and PVs into the grid, either individually or in combination, requires adequate planning. If this is not done, the consistency of the system may be jeopardized. When it comes to the operation of electricity grids, the unpredictability of time is the most important factor in photovoltaic (PV) generation. The issue with electric vehicles is that they have the potential to disrupt the demand side and cause the grid to become overloaded. When this happens, the power quality and the stability of the system may both suffer. The authors concluded that the integration of EVs and PVs into the grid has to be planned and managed. For instance, the use of a scheduled load strategy may enhance the number of PVs and EVs that are able to penetrate the system [17].

The use of renewable energy sources in electric vehicle charging infrastructure is undeniably advantageous for a number of reasons, including the following:

- (1) It lessens the impact that electric vehicles have on the electricity grid.
- (2) It solves the voltage regulation issues that are associated with the electricity grid.
- (3) It brings down the cost of utility supply.
- (4) It increases the amount of energy that can be stored by increasing the amount of renewable energy produced.
- (5) RES such as photovoltaic (PV), wind turbine (WT), fuel cell (FC), hydro-electricity, biogas, and other RES are excellent candidates for use as possible charging and powering sources for electric vehicles.

The requirements for charging electric vehicles are satisfied by these systems in conjunction with energy storage systems (ESS) and the associated electrical equipment, which are all brought together via the necessary connection techniques. EVs are able to be fuelled using a greater percentage of RES due to the utilization of local power sources, which may include RES [18]. At the moment, renewable energy sources (RES) such as photovoltaic, wind turbine and biomass energy may significantly increase the amount of electricity produced by the power system. In general, they may be readily used to charge electric vehicles owing to the high

Energy density of their batteries, the cheap cost of their construction, the simplicity with which they can be implemented and the enhanced efficiency with which they produce power. Solar energy is gaining widespread adoption as a result of the vast quantities of solar energy that are available on the surface of the earth, as well as its non-polluting and noiseless characteristics [19]. The photovoltaic system in the stand-alone mode is often unreliable since the generation of power by PV is affected by both the weather and the circumstances of the surroundings. The photovoltaic modules, on the other hand, are distinguished by their simple construction, compact dimensions, low weight, and sturdiness during transit and installation. In addition, the installation of the photovoltaic system only takes a short amount of time, it is simple to combine the PV system with other types of power generation and it may be utilized in private and/or public settings. The amount of variation in the power output of PVs may be greatly reduced by the installation of a device that stores energy at the charging station [20]. PV solar panels are used to charge electric vehicles for a variety of reasons, including the following:

- (1) The price of solar photovoltaic panels has dropped consistently over the last several decades, and it is now less than 1 euro per watt produced.
- (2) Photovoltaic panels may be put on roofs as well as in parking spots next to electric vehicle parking lots.
- (3) Because a portion of the electricity used to charge electric vehicles comes from photovoltaic panels, the peak demand caused by the grid-side charging of EVs is minimized.
- (4) PV charging systems store the energy generated by PV panels using a battery storage system (BSS), which assists in the management of daytime and seasonal climatic fluctuations that have an effect on solar output.
- (5) It is more cost-effective to charge electric cars using a system that draws energy from the sun rather than using a system that draws electricity from the grid.

Taking into consideration the existing configuration of the power grid as well as the rapidly growing load from electric vehicles, which necessitates an adjustment to the primary energy structure of the grid side and significantly raises the cost of transforming the power grid, there is the potential for the charging of EVs from PV panels to make them more environmentally and economically viable by reducing the overall cost of charging facilities [21]. This could be accomplished by increasing the efficiency of charging facilities and lowering the price of PV panels. In this thesis, an energy storage system (ESS) is integrated with a traditional alternating current (AC) grid and a solar-powered charging station for electric vehicles. The energy storage system is designed to either store any excess power that is produced by photovoltaic cells or to provide energy to charge electric vehicles in the event that PV production is unavailable, which often occurs during periods of high demand. Because of this, it is possible to charge electric cars quickly even when there is not enough energy to completely satisfy the demand for charging them [22].

The use of this bidirectional converter comes with a number of extra advantages, including those listed above. The system in is an off-board EV battery charging system [23]. When the EV is in a standstill condition, the system in charges the

EV battery using the electricity from the PV array using a bidirectional DC–DC converter [24]. Nevertheless, in order to power the DC load found inside the EV when it is in a running condition, the electric vehicle's battery will be depleted. Only when there is enough light from the sun can the electric vehicle's battery be charged. This is a significant drawback. The recommended charger for charging the battery of an electric vehicle makes use of a PV array that is paired with a SEPIC converter, a bidirectional DC–DC converter, and a backup battery bank in order to get over this disadvantage and charge the EV battery in a continuous manner without any interruptions. The significance of the thesis is in achieving its objectives.

2 Literature Survey

This idea thus proposes the use of an off-board charger, in which the electric vehicle's battery is housed inside the vehicle unit, while the photovoltaic array and backup battery bank are housed within the charging station or parking station, respectively. In the research that has been done on off-board charging systems, numerous converter topologies have been presented. It is the versatility of the SEPIC converter to function in both boost and buck modes that gives it the advantage over other converter topologies and makes it the most popular [5–9]. Additionally, it benefits from having the same input and output voltage polarity, low input current ripple and low electromagnetic interference (EMI). On the other hand, at times of poor solar irradiation or when the sun is not shining, it is necessary to have an extra storage battery bank in order to charge the electric vehicle's battery. The amount of solar irradiation will determine whether this backup battery bank should be discharged while it is being charged or charged while it is being discharged. Because of this, a bidirectional converter that allows electricity flow in either direction is necessary. Non-isolated and isolated converters are the two categories that may be used to describe the bidirectional converters. Isolation is provided by the transformer inside the isolated converters, which results in an increase in the cost, weight and size of the converter [7–9]. The weight and size of an electric vehicle are the primary issues, and as a result, non-isolated bidirectional converters are the components that work best for this application.

The bidirectional interleaved DC–DC converter, also known as the BIDC, is the favoured topology among the many non-isolated bidirectional converter topologies. This is owing to the BIDC's benefits, which include better efficiency in discontinuous conduction mode and low inductance value, as well as decreased ripple current as a result of the multiphase interleaving approach. Employing a zero voltage resonant soft switching technique helps reduce the inductor current parasitic ringing effect, in addition to lowering the turnoff losses that are caused by the snubber capacitor that is placed across the switches.

3 Operation of the Proposed System

The proposed PV–EV battery charger consists of a PV array, a SEPIC converter, a half-bridge BIDC, an EV battery, a backup battery bank and a controller as shown in Fig. 1. The controller is used to generate the gate pulses to the SEPIC converter for obtaining the constant output voltage at the DC link. The gate pulses to the switches of BIDC are also generated to operate BIDC in boost mode to charge the backup battery from PV array and in buck mode to charge EV battery from the backup battery. Also, the controller generates the gate pulses to the auxiliary switches Sa, Sb and Sc. During high solar irradiation, all the auxiliary switches are ON to interface DC link with PV array through the SEPIC converter, DC link with the backup battery through BIDC and DC link with EV battery [10]. When solar irradiation is low, switch Sa is turned OFF isolating the PV array and SEPIC converter from the DC link. Whereas the switch Sc is turned OFF to disconnect BIDC and backup battery from the DC link, when the solar power is insufficient to charge backup battery. The proposed system operates in three modes, namely mode 1, mode 2 and mode 3 as explained in this section.

- Mode 1

During peak sunshine hours, when the generated PV array power is higher, all the auxiliary switches are ON to charge both EV battery and backup battery simultaneously from PV array through SEPIC converter and BIDC, respectively. In this mode, BIDC operates in forward direction boosting the DC link voltage to charge backup battery.

- Mode 2

During low solar irradiation conditions and non-sunshine hours, PV array power is insufficient to charge EV battery. Hence, the PV array is disconnected from the DC link by turning OFF the switch Sa and switches Sb and Sc are ON connecting EV battery to the backup battery through BIDC. In this mode, BIDC operates in reverse direction stepping down the backup battery voltage to charge EV battery.

- Mode 3

When PV array power generated is sufficient to charge only EV battery, switches Sa and Sb are ON and switch Sc is OFF to disconnect the BIDC and backup battery bank from the DC link.

4 Design of Controllers

Figure 1 shows the controller of the proposed charger generates gate pulses to the switches present in the SEPIC converter, BIDC and also to the three auxiliary switches. Waveforms to turn ON and turn OFF the auxiliary switches is shown in

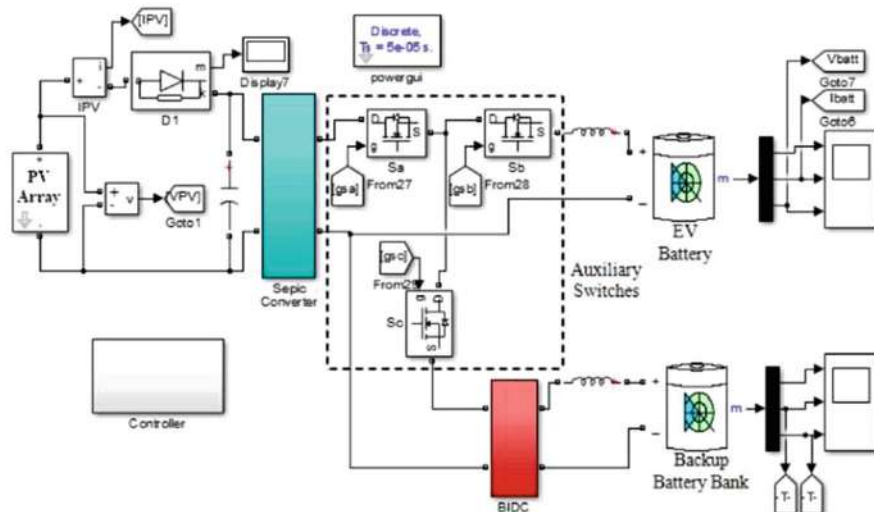


Fig. 1 Simulation model of the proposed charger

Fig. 4. Controller senses the PV array voltage and current and computes the PV array power. If the PV array power is greater than EV battery-rated power, PR, then the controller generates the gate pulses to turn ON all the auxiliary switches to charge both EV battery and backup battery bank simultaneously from the PV array [23]. If the PV array power is lesser than EV battery-rated power but higher than the minimum required power, PM, the switch Sc is turned OFF to disconnect the backup battery from the charging system and switches, Sa and Sb are turned ON to charge the EV battery alone from the PV array. If the PV array power is lesser than the minimum required power, PM, then the switch Sa is turned OFF to isolate the PV array and SEPIC converter from the charging system [25]. The switches Sb and Sc are turned ON, enabling the backup battery to charge EV battery. The PI voltage controller is used in the proposed charging system to generate gate pulses to the MOSFET in the SEPIC converter to maintain a constant voltage at the DC link irrespective of variations in the PV array voltage. BIDC comprises of three legs with two switches in each leg. Gate pulses have to be provided to the two switches in the same leg with the phase shift of 180° from each other. The controller in the proposed system generates six gate pulses to the BIDC depending on the PV array power [26]. If PV array power exceeds PR, gate pulses are generated to the switches of BIDC to operate it in boost mode, stepping up the DC link voltage to charge the backup battery bank. In this mode, the gate pulses are generated to the switches of leg 1 with 0° phase and to the leg 2 switches with 120° phase shift from that of leg 1 switches and to the leg 3 switches with 240° phase shift from that of leg 1 switches [27]. If the PV array power is less than PM, the gate pulses are generated accordingly to operate BIDC in buck mode, producing a step-down voltage at the DC link sufficient to charge the EV battery by the backup battery. In this mode, the

gate pulses are fed to the leg 3 switches with 0° phase and gate pulses to the leg 2 and leg 1 switches are 120° and 240° phase shifted with respect to that of leg 3 switches, respectively.

5 Simulation Studies and Results

The Simulink component of MATLAB is used to simulate the proposed system for analytical purposes Fig. 4. The PV array's classical equation is used in the modelling process [28, 29]. The Sim Power Systems Blockset in the Simulink library is used to model the SEPIC and BIDC converter's power MOSFETs, inductors, and capacitors. The controller is constructed using components from the Simulink library, including a PWM generator, a pulse generator, logic gates, a comparator, a multiplier, and a PI controller. Combining the PV array concept with the.

In addition to the various battery models that were already in the Simulink library, a virtual connection through performs many operations and SEPIC converter and a BIDC were developed by our team. These were incorporated into the construction of the proposed charging system, which can be seen in Fig. 1. Figure 3a, b shows, respectively, the created simulation model of the SEPIC converter and the BIDC. Both models were simulated using the same data And were presented in the form of subsystems. It was feasible to analyses the dynamic response of the recommendation system with the assistance of the created simulation model by applying it to irradiation levels of 850, 100 and 500 W/m² for the PV array in modes 1, 2 and 3, respectively. This allowed for the evaluation of the dynamic response. Figure 3 illustrates the results of the simulation, which show the waveforms of the PV array's voltage and current as well as the gate pulses that are transmitted to the auxiliary switches. Additionally, the waveforms of the gate pulses can also be seen in this figure. The waveforms of the irradiation are shown using a scale of one for every thousand watts per square metre in Fig. 4. When operating in this mode, both the primary battery that powers the electric vehicle and the secondary battery that serves as a backup will be charged simultaneously. When the irradiance is low, about 100 W/m², the gate pulses of the auxiliary switches, Vgsb and Vgsc, are high while the gate pulse, Vgsa, is low. This is due to the fact that the power from the PV array is not sufficient to charge the electric vehicle's battery under these conditions. Therefore, in order to charge the EV battery while operating in this mode, the backup battery bank must first be depleted through the use of BIDC. During the time that the irradiance is set to 500 W/m², the auxiliary switches Sa and Sb are activated, but switch Sc is deactivated. As a result of this action, the backup battery is cut off from the system. Because the electricity supplied by the PV array is only sufficient for charging the EV battery, the secondary battery will not be charged while this mode is active. Instead, the secondary battery will be disconnected. The fact that the electric vehicle's battery is always being charged in any one of the three modes is shown by Fig. 4. This causes the gate pulses that are sent to the switch Sb

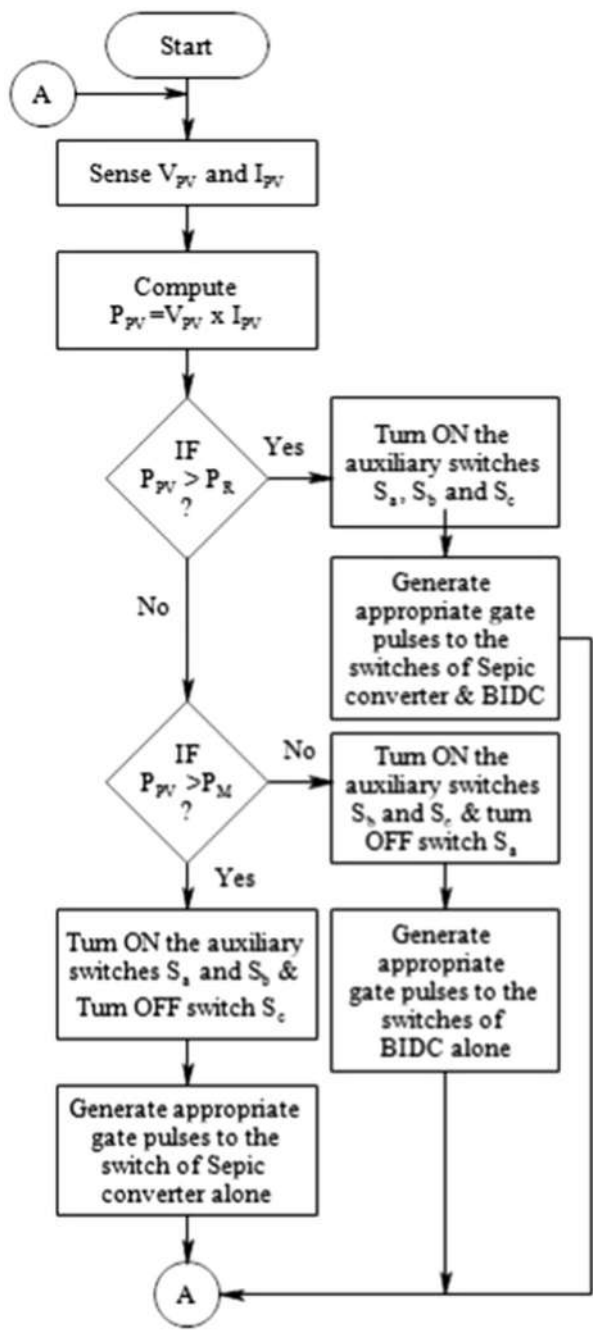


Fig. 2 Flowchart of gate pulses generation for the auxiliary switches

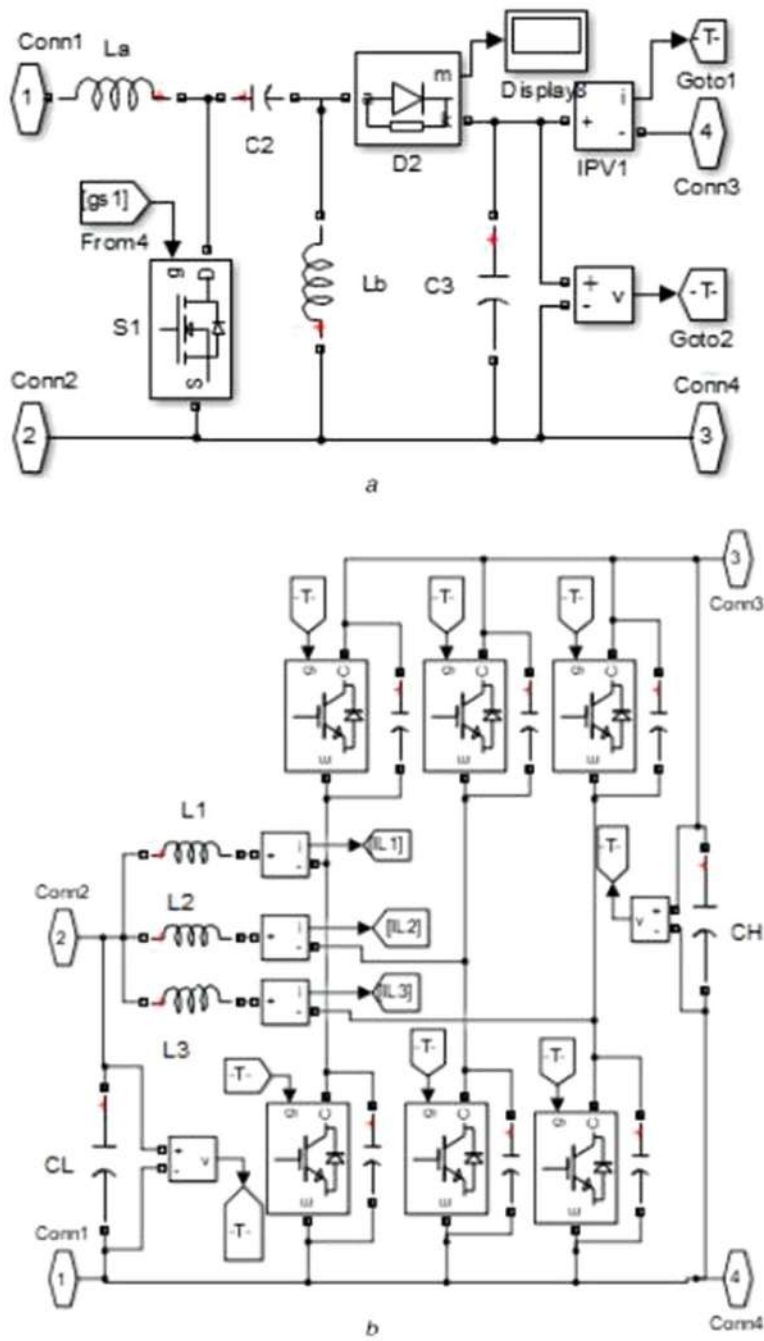


Fig. 3 (a) SEPIC converter simulation model and (b) BIDC converter simulation model

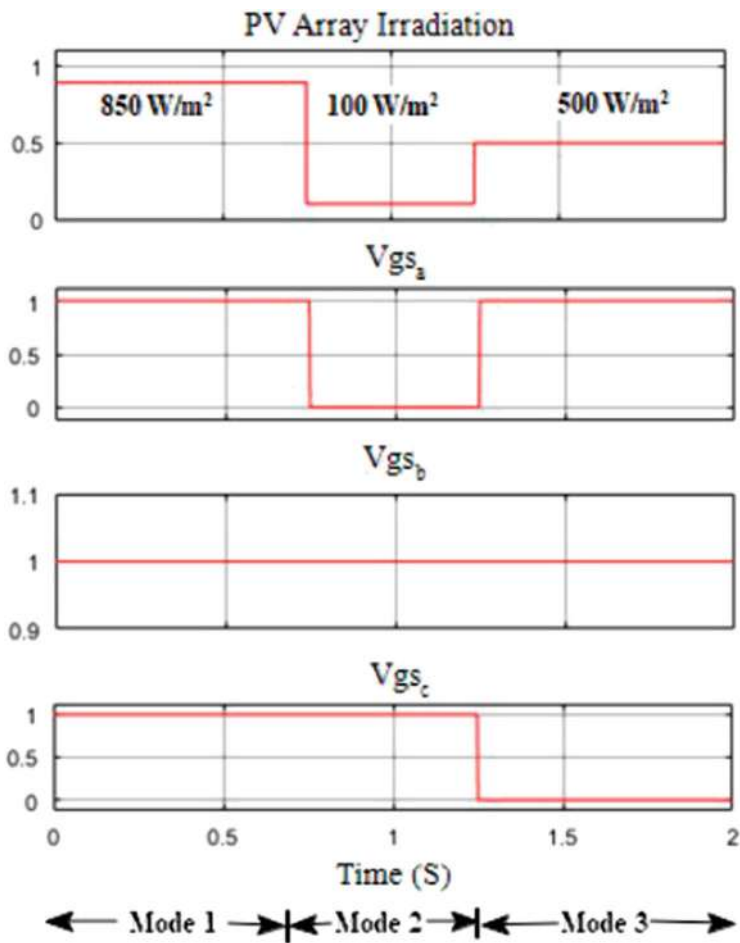


Fig. 4 Irradiation waveforms of the PV arrays and gate pulses to the auxiliary switches

to remain permanently high. The battery of the electric car is unplugged from the charging system after it has been determined that it has reached its full capacity.

Electric vehicle’s battery from being trickle charged, switch OFF Switch, Sb on the charging system. Fig. 4 displays the simulated dynamic waveforms for the PV array, DC link, electric vehicle battery and backup battery, together with the corresponding irradiation levels. In mode 1, the SEPIC converter reduces the 33.3 V PV array voltage (VPV) to the 28 V DC link voltage (Vdc). Figure 5a, b shows this to be the case. As illustrated in Fig. 5c, the EV battery is being charged in this fashion since its state of charge (SOC) is increasing as its negative current decreases. In this mode, the BIBC acts as a forward-biased boost converter, boosting the DC link voltage, Vdc, from 28 to 60.6 V when the backup battery is charged and its state of charge rises (Fig. 5d). In mode 2, when solar radiation is scarce, the PV

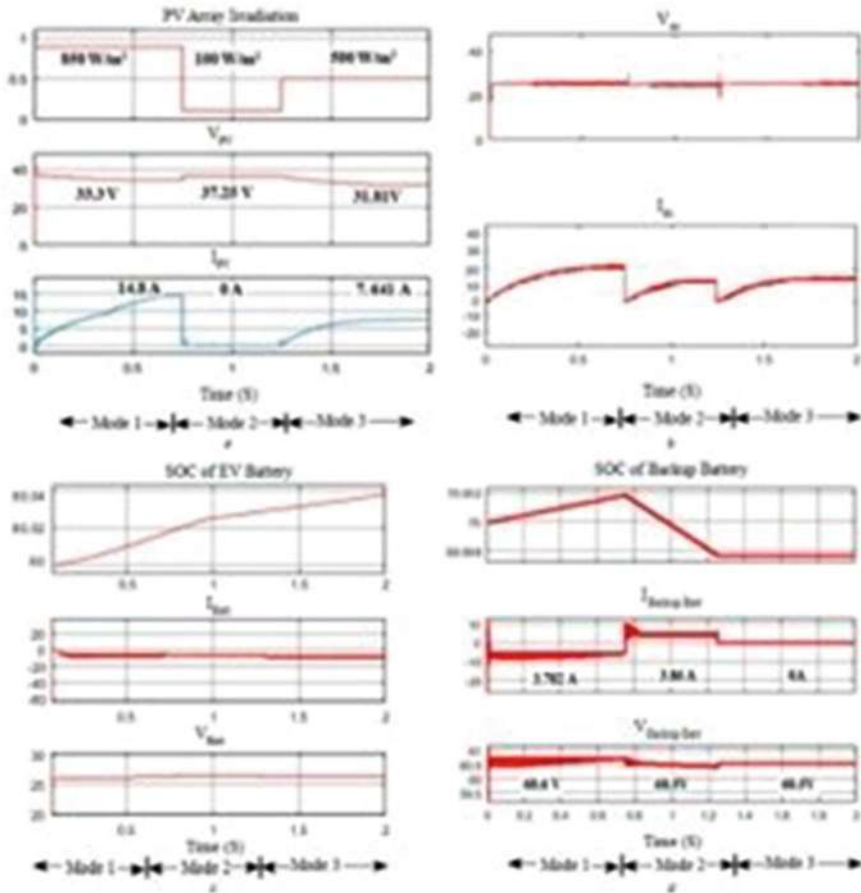


Fig. 5 (a) PV array voltage (VPV) and current (IPV) waveforms; (b) Vdc and Idc, the voltage and current of the DC connection; (c) state of charge (SOC), current (Ibatt) and voltage (Vbatt) of electric vehicle batteries; (d) state of charge (SOC), current (Ibackup Batt) and voltage (Vbackup Batt) of backup batteries

array is turned off, so its voltage increases to its open-circuit value of 37.25 V and its current decreases to 0 A ($V_{PV} = I_{PV}$). Figure 5a shows the voltage and current waveforms for a PV array, which is a visual representation of this concept. During this time, the BIDC operates in reverse buck mode to reduce the backup battery voltage to 27.32 V, the level at which the electric vehicle's battery can be charged. Figure 5c shows this phenomenon. The positive current and decreasing percentage of the backup battery's remaining capacity in Fig. 5d indicate that the battery is being drained in this mode. Mode 3 charges the EV battery by reducing the voltage of the PV array from 31.81 V to a DC link voltage, V_{dc} , of 27.6 V, as shown in Fig. 5a, b. The voltage of the backup battery drops from 60.6 V to 55.2 V at the end of this mode. The EV battery is still being charged even in this mode, as shown by the

fact that the state of charge is rising while the current is falling. In mode 3, as shown in Fig. 5d, the current is reduced to zero while the voltage of the backup battery remains at its previous value of 55.2 V. This results from the fact that the backup battery is no longer receiving any power from the charger. In all three modes of operation, as shown in Fig. 5c, the EV battery's state of charge increases while its current decreases. This means that the PV array or the backup battery is constantly recharging the EV's battery.

The waveforms of the alternating currents that pass through the inductors of a BIDC in its various operating modes are shown in Fig. 5d. The backup battery is being discharged in mode 2 because the inductor current is flowing in the opposite direction compared to mode 1, which has no inductor current at all.

If you observe mode 3, it indicates that the BIDC is no longer attached to the charger. In order to verify the results of the simulations, a hardware prototype is created and evaluated, and the findings of these activities are presented in the following section.

6 Conclusion

An off-board EV battery charging method supplied by a PV array is proposed in this document. In this research, we test the system's flexibility to maintain a constant charge on the EV battery under different irradiation conditions. The system is designed and simulated in the Simulink environment of MATLAB software. Each of the three modes of operation for the proposed charging system has a dedicated hardware prototype that has been tested in the lab, and the results are provided. Both the experimental inspection and the dynamic reaction of the system are given by applying the RCP methodology in the OPAL-RT real-time simulator OP4500. The congruence between simulated and experimental results demonstrates the effectiveness of the proposed charger.

References

1. Santhosh, T.K., Govindaraju, C.: 'Dual input dual output power converter with one-step ahead control for hybrid electric vehicle applications', *IET Electr. Syst. Transp.*, 2017, 7 (3), pp. 190–200
2. Shukla, A., Verma, K., Kumar, R.: 'Voltage-dependent modelling of fast charging electric vehicle load considering battery characteristics', *IET Electr. Syst. Transp.*, 2018, 8 (4), pp. 221–230
3. Wirasingha, S.G., Emadi, A.: 'Pihef: plug-in hybrid electric factor', *IEEE Trans. Veh. Technol.*, 2011, 60, pp. 1279–1284
4. Kirthiga, S., Jothi Swaroopan, N.M.: 'Highly reliable inverter topology with a novel soft computing technique to eliminate leakage current in grid-connected transformerless photovoltaic systems', *Comput. Electr. Eng.*, 2018, 68, pp. 192–203

5. Badawy, M.O., Sozer, Y.: 'Power flow management of a grid tied PV-battery system for electric vehicles charging', *IEEE Trans. Ind. Appl.*, 2017, 53, pp. 1347–1357
6. Van Der Meer, D., Chandra Mouli, G.R., Morales-Espana Mouli, G., et al.: 'Energy management system with PV power forecast to optimally charge EVs at the workplace', *IEEE Trans. Ind. Inf.*, 2018, 14, pp. 311–320
7. Xavier, L.S., Cupertino, A.F., Pereira, H.A.: 'Ancillary services provided by photovoltaic inverters: single and three phase control strategies', *Comput. Electr. Eng.*, 2018, 70, pp. 102–121
8. Krithiga, S., Ammasai Gounden, N.: 'Investigations of an improved PV system topology using multilevel boost converter and line commutated inverter with solutions to grid issues', *Simul. Model. Pract. Theory*, 2014, 42, pp. 147–159
9. Sujitha, N., Krithiga, S.: 'RES based EV battery charging system: a review', *Renew. Sustain. Energy Rev.*, 2017, 75, pp. 978–988
10. Farzin, H., Fotuhi-Firuzabad, M., Moeini-Aghaie, M.: 'A practical scheme to involve degradation cost of lithium-ion batteries in vehicle-to-grid applications', *IEEE Trans. Sustain. Energy*, 2016, 7, pp. 1730–1738
11. Zubair, R., Ibrahim, A., Subhas, M.: 'Multiinput DC–DC converters in renewable energy applications—an overview', *Renew. Sustain. Energy Rev.*, 2015, 41, pp. 521–539
12. Duong, T., Sajib, C., Yuanfeng, L., et al.: 'Optimized multiport dc/dc converter for vehicle drive trains: topology and design optimization', *Appl. Sci.*, 2018, 1351, pp. 1–17
13. Santhosh, T.K., Natarajan, K., Govindaraju, C.: 'Synthesis and implementation of a multiport dc/dc converter for hybrid electric vehicles', *J. Power Electron.*, 2015, 15 (5), pp. 1178–1189
14. Hongfei, W., Peng, X., Haibing, H., et al.: 'Multiport converters based on integration of full-bridge and bidirectional dc–dc topologies for renewable generation systems', *IEEE Trans. Ind. Electron.*, 2014, 61, pp. 856–869
15. Shi, C., Khaligh, A.: 'A two-stage three-phase integrated charger for electric vehicles with dual cascaded control strategy', *IEEE J. Emerging Sel. Topics Power Electron.*, 2018, 6 (2), pp. 898–909
16. Chiang, S.J., Shieh, H., Chen, M.: 'Modeling and control of PV charger system with SEPIC converter', *IEEE Trans. Ind. Electron.*, 2009, 56 (11), pp. 4344–4353
17. Falin, J.: 'Designing DC/DC converters based on SEPIC topology', *Analog Appl. J.*, 2008, 4Q, pp. 18–23. Available at <https://e2echina.ti.com/cfs-file/key/telligent-evolution-components/attachments/13-112-00-00-00-58-20/Designing-DC-DC-converters-based-onSEPICtopology.pdf>
18. Banaei, M.R., Sani, S.G.: 'Analysis and implementation of a new SEPIC based single switch buck–boost DC–DC converter with continuous input current', *IEEE Trans. Power Electron.*, 2018, 33, (12), pp. 10317–10325
19. Singh, A.K., Pathak, M.K.: 'Single-stage ZETA-SEPIC-based multifunctional integrated converter for plug-in electric vehicles', *IET Electr. Syst. Transp.*, 2018, 8 (2), pp. 101–111
20. Du, Y., Zhou, X., Bai, S., et al.: 'Review of non-isolated bi-directional DCDC converters for plug-in hybrid electric vehicle charge station application at municipal parking decks'. 2010 Twenty-Fifth Annual IEEE Applied Power Electronics Conf. Exposition, Palm Springs, CA, USA., 2010, pp. 1145–1151
21. Kwon, M., Oh, S., Choi, S.: 'High gain soft-switching bidirectional DC–DC converter for eco-friendly vehicles', *IEEE Trans. Power Electron.*, 2014, 29, pp. 1659–1666
22. Mirzaei, A., Jusoh, A., Salam, Z., et al.: 'Analysis and design of a high efficiency bidirectional DC–DC converter for battery and ultracapacitor applications', *Simul. Model. Pract. Theory*, 2011, 19, pp. 1651–1667
23. Han, J.T., Lim, C.-S., Cho, J.-H., et al.: 'A high efficiency non-isolated bidirectional DCDC converter with zero-voltage-transition'. 2013—39th Annual Conf. IEEE Industrial Electronics Society, 2013, pp. 198–203
24. Zhang, J., Lai, J.-S., Kim, R.-Y., et al.: 'High-power density design of a soft switching high-power bidirectional DC–DC converter', *IEEE Trans. Power Electron.*, 2007, 22, pp. 1145–1153

25. Paul, A., Subramanian, K., Sujitha, N.: 'PV-based off-board electric vehicle battery charger using BIDC', *Turk. J. Electr. Eng. Comput. Sci.*, 2019, 27, (4), pp. 2850–2865
26. Sree, L., Umamaheswari, M.G.: 'A Hankel matrix reduced order SEPIC model for simplified voltage control optimization and MPPT', *Sol. Energy*, 2018, 170, pp. 280–292.
27. Zhang, J.: 'Bidirectional DC-DC power converter design optimization, modeling and control'. Dissertation, the faculty of the Virginia Polytechnic Institute and State University, Virginia Polytechnic Institute and State University, 2008
28. Gounden, N.G.A., Krithiga, S.: 'Power electronic configuration for the operation of PV system in combined grid-connected and stand-alone modes', *IET Power Electron.*, 2014, 7, pp. 640–647
29. Arul Daniel, S., Ammasai Gounden, N.: 'A novel hybrid isolated generating system based on PV fed inverter assisted wind driven induction generators', *IEEE Trans. Energy Convers.*, 2004, 19, (2), pp. 416–422.

Human Stress Detection in Sleep Mode Compared with Non-sleep Mode Using Machine Learning Algorithms



S. A. Sajidha, M. Sanjay, Pillaram Manoj, A. Sheik Abdullah, R. Priyadarshini, V. M. Nisha, and Aakif Mairaj

Abstract Stress is a widespread issue that affects people everywhere. Several things, including difficulty at job, money troubles, marital problems, and health issues, can contribute to it. Anxiety, depression, and cardiovascular illnesses are just a few of the physical and mental health issues that stress can cause. In order to properly manage stress, it is crucial to understand what causes it in people. Researchers and clinicians can encourage the development of new approaches to deal with the problematic effects of a sustained stress response by identifying and managing stress in its early phases. We propose to use sleep dataset in order to predict the stress levels as compared to non-sleep mode data that has attributes such as snoring range, respiration rate, body temperature, limb movement rate, blood oxygen levels, eye movement, number of hours slept, and heart rate with stress levels classified from 0 (low/normal) to 4 (extreme/high). We compare the performance of Logistic regression, XGB classifier, Random Forest classifier, Naive Bayes, Decision tree, LightGBM, CatBoost (Categorical Boosting), Gradient Boosting Regressor, SVM, ANN, and KNN models to predict the stress levels for sleep mode data. We found that only SVM and Naïve bayes gave a maximum accuracy of 85% stress prediction accuracy using non-sleep mode data whereas many models such as Logistic regression, Naive Bayes, LightGBM, CatBoost, SVM, and KNN predicted with 100% accuracy for sleep mode data.

S. A. Sajidha (✉) · M. Sanjay · P. Manoj · A. Sheik Abdullah · R. Priyadarshini · V. M. Nisha
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

e-mail: sajidha.sa@vit.ac.in; sanjay.m2020a@vitstudent.ac.in;
pillaram.manoj2020@vitstudent.ac.in; sheikabdullah.a@vit.ac.in; priyadarshini.r@vit.ac.in;
priyadarshini.r@vit.ac.in

A. Mairaj
School of Sciences, Indiana University Kokomo, Kokomo, IN, USA
e-mail: amairaj@iu.edu

Keywords Stress detection · Human stress · Stress levels · Pressure · Sleep mode · Non-sleep mode · Machine learning algorithms · Deep learning algorithm · Accuracy · User interface

1 Introduction

Stress is a difficult and severe psychological condition that occurs when the body experiences stress against any kind of pressure. This need may come from a wide range of sources, including employment, family, and personal connections. Chronic and long-term stress can have harmful impacts on physical and mental health, such as headaches, sleep difficulties, and an increased risk of cardiovascular disease. To avoid long-term harmful consequences, it is important to recognize and control stress. In order to begin pressure reduction mediations, nonintrusive pressure detecting tools that continuously monitor symptoms of anxiety with little impact on professionals' daily activities could be used. These applications could not provide more effective and affordable mediations in challenging work environments, much alone more beneficial circumstances in which workers would be more willing to handle their responsibilities.

The physical response to pressure is stress (Fig. 1). The sensation is squeezed. Pressure can be brought on by a variety of events or conditions in life. Each one has their own unique technique of dealing with pressure. Their ability to adjust to it depends on their ancestry, early life events, character, social situation, and financial situation. However, too much pressure can lead to unfavorable outcomes that make people feel depressed. Stress may be a mental health issue that shortens one in four people's lives, according to the World Health Organization (WHO).

In addition to causing mental health problems, human pressure also leads to socioeconomic problems, a lack of trust in the workplace, strained professional relationships, melancholy, and in some extreme situations, the long-term compulsion to commit suicide. They could end up in a very long period of constant anxiety and terror as a result. Over time, it may negatively impact both physical and mental health. People are therefore uneasy despite their thriving. The strain may be profound, physical, or even mental in origin.

2 Literature Survey

2.1 *For Sleep Mode Dataset*

Supratak et al. [1] makes a study employed deep learning algorithms to categorize various sleep phases even though it was not primarily focused on stress prediction. According to the study, the convolutional neural network algorithm showed an accuracy of 84.6% when classifying sleep phases, which might be helpful for identifying changes in sleep patterns caused by stress.



Fig. 1 Human stress

Can et al. [2] sought to design a system for continuous stress detection in real-life scenarios using wearable sensors. To estimate stress levels from sensor data, they employed machine learning methods such as support vector machine, Random Forest, and k-nearest neighbors' Random Forest method, which was shown to be the most accurate in forecasting stress levels, with an accuracy of 83.3%.

2.2 For Non-sleep Mode Datasets

Baheti et al. [3] in their paper propose a machine learning-based stress detection system. In order to detect stress, the system gathers physiological information such as heart rate, blood pressure, and skin conductance level and uses machine learning algorithms. The effectiveness of various machine learning algorithms, including Naive Bayes, k-nearest neighbors, support vector machines, and Random Forest for stress detection, is compared in the study.

Shelke et al. [4] conveyed that the use of machine learning algorithms to identify people's degrees of stress is explored in their research. The writers talk about how crucial it is to recognize stress levels precisely and promptly in order to avoid a variety of physical and mental health problems. They get to the conclusion that SVM and DNN are more effective at detecting stress levels than Random Forest.

Li et al. [5] concentrated on using artificial intelligence (AI) models called deep neural networks to identify stress. Most likely, the researchers discussed their research on the application of deep neural networks for stress detection. For applications like stress detection that call for the analysis of varied and multidimensional data, deep neural networks are a form of machine learning model that is well-suited.

Ahuja et al. [6] probably describe their investigation into the use of machine learning algorithms for identifying mental stress in university students. They might have discussed their technique, accuracy results, and findings in the publication. Different performance indicators, such as precision, recall, F1-score, or area under the receiver operating characteristic (ROC) curve, could be used to assess the accuracy.

Zainudin et al. [7] most likely cover the stress detection using machine learning and deep learning approaches in the publication titled “Stress Detection Using Machine Learning and Deep Learning” published in 2021. It is possible that the paper discussed the researchers’ approach, conclusions, and research algorithms. The precise algorithms used, such as decision trees, support vector machines, neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), could be described.

Ersoy et al. [8] most likely survey research focusing on stress detection in everyday life scenarios using smartphones and wearable sensors. The authors may have undertaken a survey of existing literature and research papers on this topic, providing a complete overview of the field’s current state. The authors may have presented an overview of these sensors’ capabilities and limitations in recording physiological or behavioral stress signals.

Archana et al. [9] most likely describe a research study on stress detection using machine learning techniques. A description of the specific methods used, such as support vector machines, decision trees, Random Forests, k-nearest neighbors, logistic regression, or other relevant algorithms, should be included.

Padma et al. [10] most likely focus on stress detection using machine learning and image processing techniques. The authors may have proposed a novel method for accurately detecting stress levels that combines machine learning and image processing. This could describe the particular algorithms used, such as decision trees, convolutional neural networks (CNNs), and support vector machines (SVMs) algorithms.

3 Dataset Description

Snoring range: This feature measures the user’s snoring range while sleeping. It is most usually measured with a microphone or other comparable equipment.

Respiration rate: This option counts the number of breaths the user takes per minute while sleeping. A respiratory belt or a pulse oximeter can be used to measure it.

Body temperature: This setting measures the user's body temperature when sleeping. It can be measured with a thermistor or a thermal camera.

Limb movement rate: This parameter measures the frequency and extent of the user's limb movements while sleeping. It can be measured with an instrument called an accelerometer.

Blood oxygen levels: This parameter checks the user's blood oxygen saturation when sleeping. A pulse oximeter is often used to measure it.

Eye movement: This characteristic monitors the frequency and amount of the user's eye movements while sleeping. It can be measured with an instrument called an electrooculogram (EOG).

Number of hours slept: This parameter calculates the total amount of time the user slept throughout the measurement period. It might be derived from other measurements or self-reported by the user.

Heart rate: This setting measures the user's heart rate when sleeping. It can be measured with an instrument called an electrocardiogram (ECG).

Stress levels: This parameter assesses the user's perceived stress levels when sleeping and is graded from 0 (low/normal) to 4 (extreme/high). The user most likely self-reported it.

4 Methodology

Data collection: We found a dataset in Kaggle that is obtained from Smart yoga pillow, and our data is structured data with a class column.

Data preprocessing: The data are preprocessed to reduce noise to segment the data into distinct time periods, such as epochs.

Feature extraction: Using a variety of signal processing techniques, pertinent features, such as frequency content, amplitude, and variability, are retrieved from the preprocessed data.

Development of a machine learning model: Using the extracted features and a labeled dataset that contains examples of both stressed and non-stressed states, a machine learning model is created.

Model training and validation: To assess the machine learning model's accuracy and generalization performance, a subset of the labeled dataset is used for training and the remaining subset for validation.

Stress detection: By applying the trained model to fresh, unexplored data and predicting the stress state based on the extracted features, it is possible to identify stress in real-world data.

5 Proposed Architecture

Figure 2 shows the architecture of the proposed methodology for our approach. The collected dataset from Kaggle is obtained from Smart yoga pillow and it is a structured data. To determine the amount of stress, different machine learning techniques may be used to preprocess and feature extracted data. Logistic regression, Naive Bayes, XGBoost, Random Forest classifier, Decision tree, LightGBM, CatBoost, SVM, and gradient boosting regressor are the algorithms we used to predict the stress levels. Deep learning techniques like artificial neural networks (ANN) may also be used to forecast stress levels. A comparison between sleeping and non-sleeping datasets may be done in order to assess the efficiency of stress detection using machine learning and deep learning models. In conclusion, a user interface may be created to let people track their levels of stress using machine learning algorithms. Individuals may receive specific comments and suggestions using this to help them better manage their stress levels and general health.

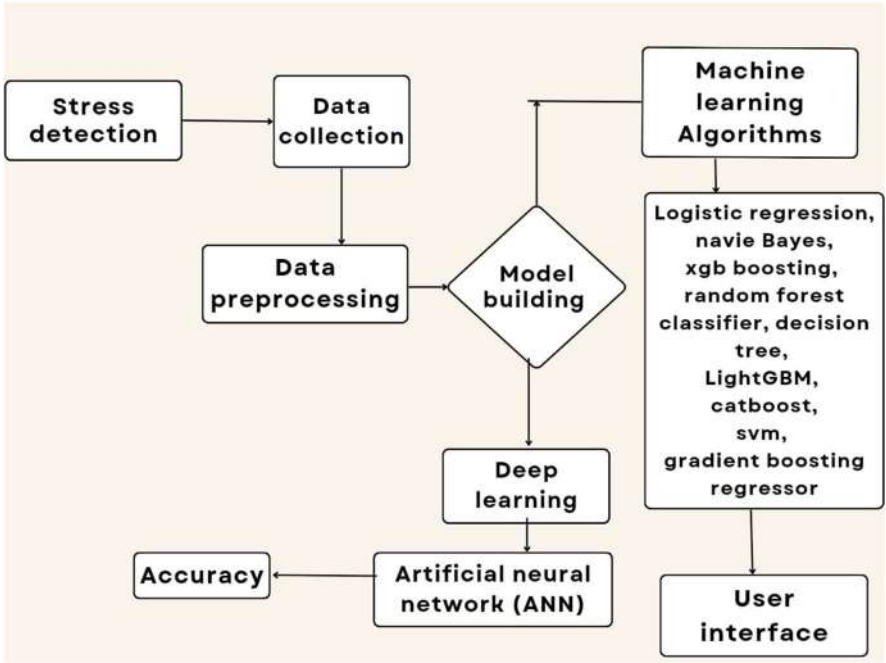


Fig. 2 Proposed methodology of human stress

6 Results and Discussion

6.1 Logistic Regression

The goal of logistic regression is to forecast stress levels using physiological cues including heart rate, respiration rate, blood oxygen level, body temperature, and eye movements. The logistic function is used to convert the linear relationship between the independent and dependent variables into a sigmoidal curve that predicts the likelihood of the result. The accuracy for logistic regression algorithm is 100%.

6.2 XGB Classifier

The XGB classifier is a binary classification implementation of the XGBoost algorithm. To reduce overfitting and increase generalization performance, it employs an ensemble of decision trees and regularization algorithms. The accuracy for XGB classifier algorithm is 98%.

6.3 Random Forest Classifier

Random Forest classifier is a binary classification implementation of the Random Forest technique. To reduce overfitting and increase generalization performance, it employs an ensemble of decision trees and random sampling. The accuracy for Random Forest classifier regression algorithm is 98%.

6.4 Naive Bayes

The Naive Bayes classifier is a binary classification implementation of the Naive Bayes algorithm. Given the class variable, it assumes that the features are independent of one another and computes the posterior probability of each class given the observed features. The Naive Bayes classifier predicts the chance of an individual being in the high-stress group based on a dataset that contains physiological signal data and stress level evaluations. The accuracy for Naive Bayes algorithm is 100%.

6.5 *Decision Tree*

Decision tree algorithm is a machine learning strategy for categorization analysis. It builds a decision tree based on the values of input variables and predicts the value of a target variable using this model. The algorithm is used to predict whether or not a person would fall into the high-stress group. The accuracy for decision tree regression algorithm is 98.4%.

6.6 *KNN*

The k-nearest neighbors' algorithm is a nonparametric, supervised learning classifier that uses distance to create classifications or predictions about an individual data point's grouping. While it can be used for either regression or classification problems, it is most commonly used as a classification algorithm based on the assumption that similar points can be found nearby. The accuracy for KNN algorithm is 100%.

Hyperparameters

Searching for hyperparameters in KNN and DT models. The K 5, 10, and 15 hyperparameter settings produced the same accuracy values in KNN testing. K is therefore decided to be 5 as the value. The accuracy value from the DT hyperparameter test with max_depth values of 20, 40, and 60 is the same. Therefore, a value of 40 is chosen for max_depth.

6.7 *LightGBM*

LightGBM is a framework for gradient boosting that use tree-based learning techniques. Gradient-based one-side sampling (GOSS), a cutting-edge method used by LightGBM, is used to filter out the data instances and get the split value, while XGBoost computes the optimal split using a presorted method and a Histogram-based approach. The accuracy for LightGBM algorithm is 100%.

6.8 *CatBoost*

CatBoost (categorical boosting) is a machine learning technique that is specifically developed to operate with categorical information and solve some of the limitations of classic gradient boosting methods. CatBoost can handle both numerical and categorical data and recognizes which features are categorical automatically. The accuracy for CatBoost algorithm is 100%.

6.9 Gradient Boosting Regressor

The gradient boosting regressor (GBR), a machine learning technique, is a member of the ensemble boosting method family. In an iterative process, it turns a number of poor learners into a single strong learner. It can handle both linear and nonlinear correlations between the data and the target variable. GBR is an effective technique for regression tasks. The accuracy for gradient boosting regressor algorithm is 99.65%.

6.10 SVM

The support vector machine (SVM) is a well-known machine learning technique for classification and regression analysis. It is a supervised learning algorithm that analyses data and recognizes patterns before categorizing or grouping fresh data points. SVM has several uses in a variety of industries, including banking, bioinformatics, image classification, and text classification. It is an effective machine learning tool due to its excellent accuracy and capacity for handling complex data. The accuracy for SVM algorithm is 100%.

6.11 ANN

Artificial neural network (ANN) are a subset of machine learning models that take their cues from the design and operation of biological neural networks seen in the human brain. An input layer, one or more hidden layers, and an output layer are the three layers that make up an artificial neural network (ANN). It has been used successfully in a variety of applications, including picture and audio recognition, natural language processing, and predictive modeling in fields such as finance, healthcare, and marketing.

Figure 3 shows the training and validation loss and MSE (mean square error). The epoch is 100 for ANN and the above plot is for training, validation loss and training, and validation MAE for ANN algorithm. The accuracy for ANN algorithm is 73%.

6.12 User Interface

Figure 4 shows the user interface for predicting stress levels by the help of features such as snoring rate, respiration rate, and other features that will predict the level of stress and recommend actions to reduce it.

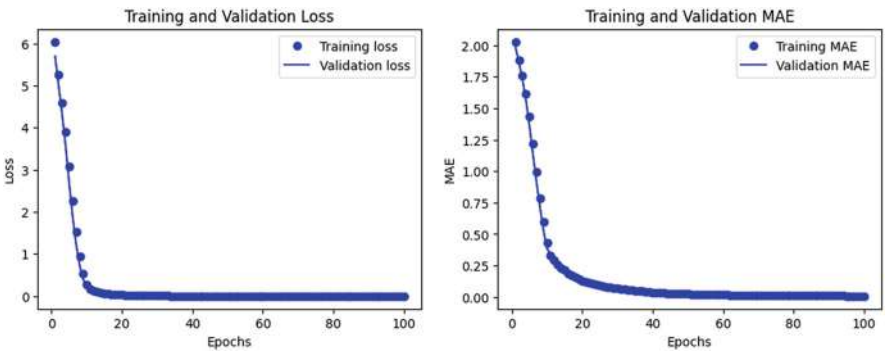


Fig. 3 Training and validation loss and MSE (mean square error)

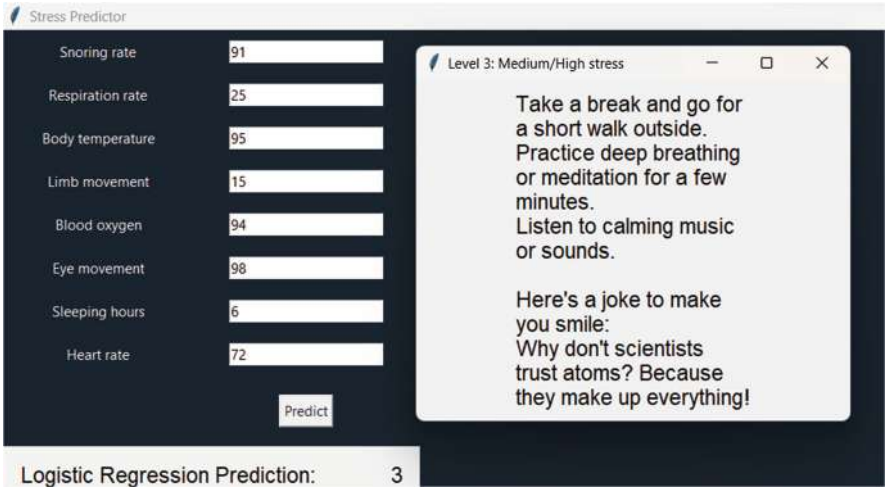


Fig. 4 User interface for predicting stress levels using logistic regression algorithm

7 Result Comparison

It is showing the performance of different machine learning algorithms on two datasets, namely “sleeping” and “non-sleeping.”

For the “sleeping” dataset, the algorithms used were CNN, KNN, logistic regression, Random Forest, PCA + LDA, and PCA+ SVM. The best accuracy reported in the paper was achieved by logistic regression with an accuracy of 92.15%. This is shown in Table 1 and Fig. 5.

For the “non-sleeping” dataset, the algorithms used were Random Forest, Naive Bayes, SVM, and KNN. The best accuracy reported in the paper was achieved by SVM with an accuracy 85.71%. This is shown in Table 2 and Fig. 6.

Table 1 Accuracy of ML for sleeping dataset

Paper no.	Algorithms	Accuracy (%)
Supratak et al. [1]	KNN	84.31
Supratak et al. [1]	Logistic regression	92.15
Supratak et al. [1]	Random Forest	90.19
Supratak et al. [1]	PCA + LDA	82.35
Supratak et al. [1]	PCA + SVM	86.27
Supratak et al. [1]	CNN	84.60

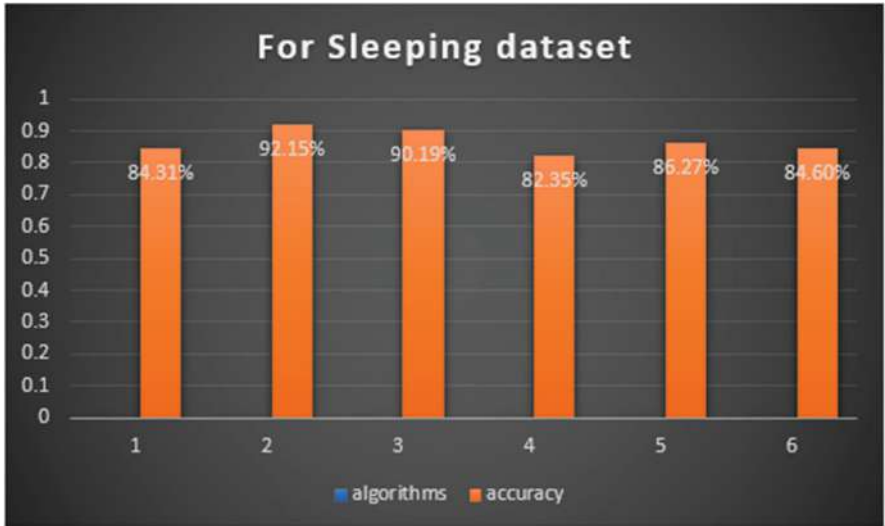


Fig. 5 Research papers accuracy for sleeping dataset using ML and deep learning algorithms

Table 2 Accuracy of ML for non-sleeping dataset

Paper no.	Algorithms	Accuracy (%)
Can et al. [2]	Random Forest	83.33
Can et al. [2]	Naive Bayes	71.42
Can et al. [2]	SVM	85.71
Can et al. [2]	KNN	55.55

Table 3 shows the accuracy of sleeping dataset for the algorithms we tested in this work.

Table 4 and Fig. 7 show the performance of our dataset namely “sleeping” on different machine learning algorithms with the algorithms in literature survey. Logistic regression, Naive Bayes, LightGBM, CatBoost (categorical boosting), SVM, and KNN have the highest accuracies of 100%, indicating that it may be the most suitable algorithm for the sleeping dataset.

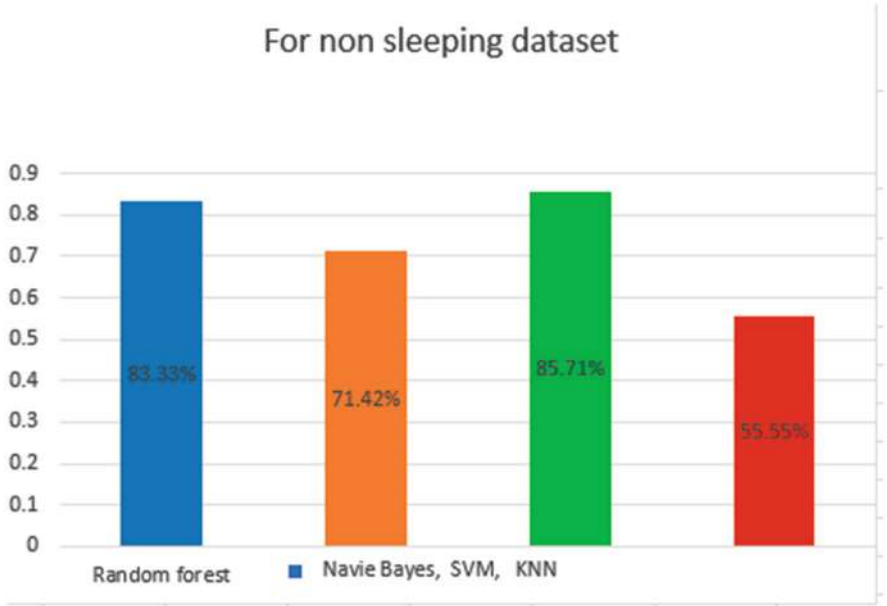


Fig. 6 Research papers accuracy for non-sleeping dataset using machine learning algorithms

Table 3 Our models with accuracy sleeping data

S. no.	Algorithms	Accuracy (%)
1	Logistic regression	100
2	XGB classifier	98
3	Random Forest classifier	98
4	Naive Bayes	100
5	Decision tree	98.4
6	LightGBM	100
7	CatBoost (categorical boosting)	100
8	Gradient boosting regressor	99.65
9	SVM	100
10	ANN	73
11	KNN	100

Table 4 Comparison of accuracy with literature

Algorithms	Logistic regression	Random Forest	Naive Bayes	KNN	SVM
Sleeping	92.15%	90.19%	–	84.31%	86.27%
Non-sleeping	–	83.33%	71.42%	55.55%	85.71%
Our accuracy	100	98	100	100	100

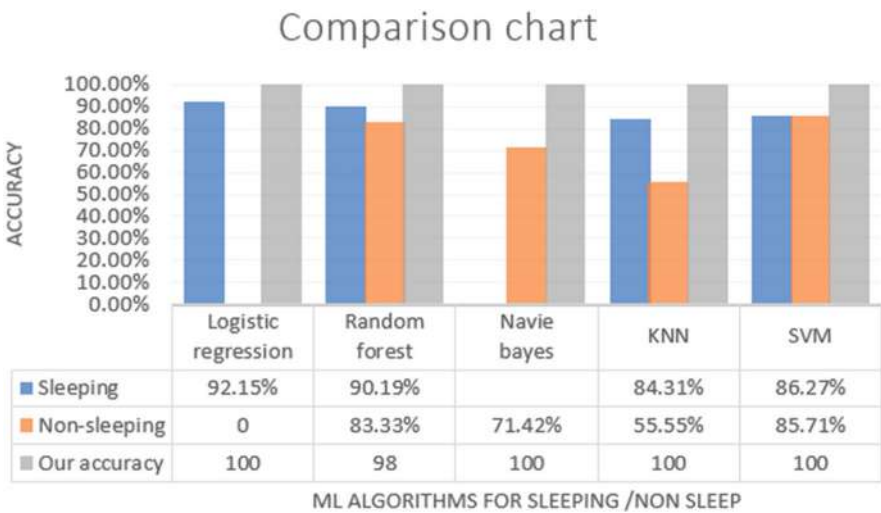


Fig. 7 Comparison of our accuracy with literature

8 Conclusion

The study used several methods to predict human stress, including Logistic regression, XGB classifier, Random Forest classifier, Naive Bayes, Decision tree, LightGBM, CatBoost, Gradient boosting regressor, KNN, SVM, and ANN. When the accuracy of each algorithm was measured, machine learning algorithms such as SVM, Logistic regression, CatBoost, LightGBM, KNN, and Naive Bayes were found to have the best accuracy (100%) when compared to deep learning algorithm ANN.

We are getting the inputs from the user (manually) or using any instruments to measure and based on the inputs we are predicting stress level of the person and each level will increase by 25% and there will be four levels. Our research shows that machine learning systems can effectively anticipate stress levels. Our study also found that parameters such as limb movement, snorting rate, respiratory rate, body temperature, spo2 level, eye movement, sleeping hours, and heart rate could accurately predict human stress levels. According to the findings, machine learning algorithms using sleep mode can be used to predict stress levels.

Future Work

Future work will include more classification models and the effectiveness of various stress detection techniques such as physiological signals, speech analysis, or facial expression identification. This would enable to recognize the advantages and disadvantages of each technique and choose the best strategy for specific situations.

References

1. Akara Supratak, Hao Dong, Chao Wu, and Yike Guo, "DeepSleepNet: a Model for Automatic Sleep Stage Scoring based on Raw Single-Channel EEG.
2. Yekta Said Can, Niaz Chalabianloo, Deniz Ekiz, and Cem Ersoy "Continuous Stress Detection Using Wearable Sensors in Real Life: An Algorithmic Programming Contest Case Study".
3. Reshma Radheshamjee Baheti, and Supriya Kinariwala, "Stress Analysis and Detection Using Machine Learning Techniques" International Journal of Engineering and Advanced Technology (IJEAT), Volume 9 Issue 1, October 2019.
4. "Using Machine Learning to Detect Stress" IRE Journals, Volume 4 Issue 10, ISSN: 2456–8880, by Sayali Shelke, Shubhangi Kor, Sahil Bavaskar, and Kirti Rajadnya, APR 2021.
5. "Deep neural networks for stress detection" by Zhandong Liu & Russell Li.
6. "Mental Stress Detection in University Students Using Machine Learning Algorithms" by Ravinder Ahuja and Alisha Bang, Procedia Computer Science, volume 152, pages 349–353.
7. "Stress Detection Using Deep and Machine Learning" Zainudin, Z., et al. 2021 Journal of Physics: Conf. Series, 1997 012019.
8. "Monitoring stress with a wrist device using context," M. Gjoreski, M. Lutrek, M. Gams, and H. Gjoreski, J. Biomed. Inform, vol. 73, pp. 159-170, September 2017.
9. "Stress detection Using Machine Learning Algorithms by B. M. Devaraju, V. R. Archana.
10. E. Padma, Talapaneni Praveen and Shaik Karimula "Stress detection using machine learning and image processing", ISSN NO:0377-9254 Vol-13, Issue 03, march 2022.

Medical Diagnosis Prediction Using Deep Learning



Sri Karthik Avala, Simran Bohra, and S. A. Sajidha

Abstract This chapter provides a comprehensive review of deep learning techniques employed for generating accurate and reliable radiology reports from chest X-ray images. The primary focus of this review is to highlight the advancements and potential applications of deep learning models in enhancing medical diagnosis prediction. A significant contribution of this review is the identification of a Bidirectional GRU model with attention as a superior performer compared to a basic model. This advanced model demonstrates remarkable proficiency in accurately predicting both short and long sentences, surpassing the baseline model's accuracy. Future research can concentrate on developing techniques capable of handling larger and more intricate datasets while simultaneously enhancing accuracy and reliability. Personalized medicine emerges as a promising area for exploration, leveraging patient-specific data to tailor treatment plans according to individual needs. The ethical considerations associated with the development and application of medical diagnosis prediction algorithms are emphasized. Future research endeavors should address concerns such as patient privacy and data analysis bias. Additionally, exploring the integration of prediction algorithms with telemedicine and remote monitoring technologies holds potential for enhancing patient care and outcomes.

Keywords Medical diagnosis · Deep learning techniques · Electronic health records (EHRs) · Natural language processing · Automatic generation of medical reports · Medical documentation · Healthcare providers · Encoder–decoder · Attention mechanism · Region of interest · Computational power

S. K. Avala (✉) · S. Bohra

Mtech Integrated, Business analytics (Scope department), VIT Chennai University, Chennai, India
e-mail: srikarthik.a2020@vitstudent.ac.in; simran.bohra2020@vitstudent.ac.in

S. A. Sajidha

School of Computing Science and Engineering, VIT Chennai University, Chennai, India
e-mail: sajidha.sa@vit.ac.in

Motivation The motivation behind the chapter is driven by the potential of medical diagnosis prediction to revolutionize healthcare and improve patient outcomes. With the availability of large and complex medical data, machine learning and artificial intelligence techniques offer the possibility of developing accurate and reliable algorithms. The authors aim to contribute to this field by exploring the effectiveness of the Bidirectional GRU with attention model compared to a basic model. By achieving better predictions, including for both short and long sentences, the research has the potential to enhance personalized medicine, early disease detection, and overall patient care. Addressing ethical considerations and integrating prediction algorithms with other healthcare technologies further adds value to the study's motivations.

1 Introduction

Deep learning has rapidly and effectively developed in research areas such as computer vision (CV) and natural language processing (NLP), and it has found a significant application area in healthcare. The last 5 years have witnessed significant advancement in computer vision research, particularly in the area of disease prediction from disease diagnosis. In the health industry, doctors examine patients based on their medical reports, making it more crucial that they have good experience. In nations with little resources, the shortage of medical experts is even more important. This problem combines the two deep learning subfields of computer vision and natural language processing. Using one or more medical photos of the patient as input, a text report that closely resembles one made by a radiologist is given as an output. There are several presumptions and many medical examination procedures, such as computed tomography (CT) and magnetic resonance imaging (MRI), but X-ray was opted for this study. The dataset was taken from Indiana University, two chest X-rays (frontal and lateral) are included in the dataset, and the reports are in XML format. We will make use of various cutting-edge methods for this task, including LSTM [1], GRU, and others. We will pair GRU [2] with the attention model because it focuses on the region of interest. Hence, it enhances report text prediction. We will be testing the accuracy using Bleu score. The value of Bleu score can range anywhere from 0 to 1, with 1 indicating that the actual report and the predicted report are identical.

2 Related Works

“Generating Radiology Reports Using Deep Learning [3]” by Michael B. Gotway and Jie Cao proposes a deep learning algorithm that can automatically generate radiology reports by extracting key information from radiology images. The algorithm consists of a convolutional neural network for image classification and

feature extraction, and a recurrent neural network for report generation. The chapter demonstrates the effectiveness of this approach on a dataset of chest radiographs, achieving high accuracy and consistency in comparison with radiologists' reports. This chapter contributes to improving the efficiency and accuracy of radiology reporting using deep learning techniques.

The paper "Automated Report Generation for Chest X-Ray Images using Convolutional Neural Networks [4]" by Siddiqui et al. proposes a deep learning algorithm for generating automated reports for chest X-ray images. The algorithm consists of a convolutional neural network for image classification and feature extraction, and a natural language processing model for report generation. The authors demonstrate the effectiveness of their approach by comparing the automated reports generated by their algorithm with reports generated by radiologists, and find that their algorithm achieves high accuracy and consistency. This paper highlights the potential of deep learning algorithms for improving the efficiency and accuracy of radiology reporting, but further research is needed to validate the generalizability of this approach.

"Automatic Report Generation for Abdominal CT Scans Using Deep Learning [5]" by Huang et al. proposes a deep learning algorithm for generating automated reports for abdominal CT scans. The algorithm consists of a convolutional neural network for image feature extraction and a natural language processing model for report generation. The authors demonstrate the effectiveness of their approach by comparing the automated reports generated by their algorithm with reports generated by radiologists, and find that their algorithm achieves high accuracy and consistency.

"Medical Image Report Generation Using Attention-based Encoder-Decoder Architecture [6]" by Prasad et al. proposes a deep learning algorithm for generating medical image reports using an attention-based encoder-decoder architecture. The authors demonstrate the effectiveness of their approach on a dataset of chest X-ray images and find that their algorithm outperforms several baseline models in terms of report quality and accuracy.

"Radiology Report Generation Using Graph-based Attention Networks [7]" by Qian et al. proposes a deep learning algorithm for radiology report generation using graph-based attention networks. The authors demonstrate the effectiveness of their approach on a dataset of chest X-ray images and find that their algorithm achieves high accuracy and consistency in comparison with radiologists' reports.

"Generating Radiology Reports from Clinical Images Using Sequence to Sequence Models [8]" by Selvaraj et al. proposes a deep learning algorithm for generating radiology reports from clinical images using sequence to sequence models. The authors demonstrate the effectiveness of their approach on a dataset of chest X-ray images and find that their algorithm outperforms several baseline models in terms of report quality and accuracy.

"Deep Learning-based Chest X-ray Report Generation: A Review [9]" by Dey et al. provides a comprehensive review of the recent advancements in deep learning-based chest X-ray report generation. The authors review various deep learning algorithms and their applications to generate chest X-ray reports. The

paper highlights the potential of these approaches for improving the efficiency and accuracy of radiology reporting.

3 Problem Statement

The manual generation of medical reports for X-ray images poses significant challenges in terms of time, expertise, and resource allocation. Skilled radiologists are often overwhelmed by the volume of X-ray images requiring analysis and report creation, leading to potential delays in diagnosis and treatment. Therefore, there is a critical need to develop an automated system that can accurately interpret X-ray images and generate comprehensive medical reports efficiently.

The objective of medical diagnosis prediction is to use the available dataset and information about a patient's symptoms, medical history, and other relevant factors to predict their likelihood of having a particular medical condition or disease. The goal is to provide doctors with a tool that can help them make informed decisions about patient care and treatment options. Medical diagnosis prediction can be particularly useful in cases where a condition is difficult to diagnose, or when the patient is presenting with atypical symptoms. Also, human beings cannot be as accurate as deep learning models. Deep learning models will give more accurate results in very less time. By leveraging machine learning algorithms and other predictive models, medical diagnosis prediction can improve the accuracy of diagnoses, potentially leading to earlier detection and treatment of diseases, as well as better patient outcomes.

4 Analysis of Datasets

4.1 Dataset Description

The data for this project is taken from Indiana University hospital network. The data is of two parts.

1. X-rays of the Chest

<http://academictorrents.com/details/5a3a439df24931f410fac269b87b050203d9467d>

For this task we will use publicly available Indian University (IU) datasets. This dataset contains 7471 images, which are images of chest X-ray (frontal, lateral), report of the corresponding image in XML file (Fig. 1).

2. Reports

<https://academictorrents.com/detail/66450ba52ba3f83fbf82ef9c91f2bde0e845aba9>

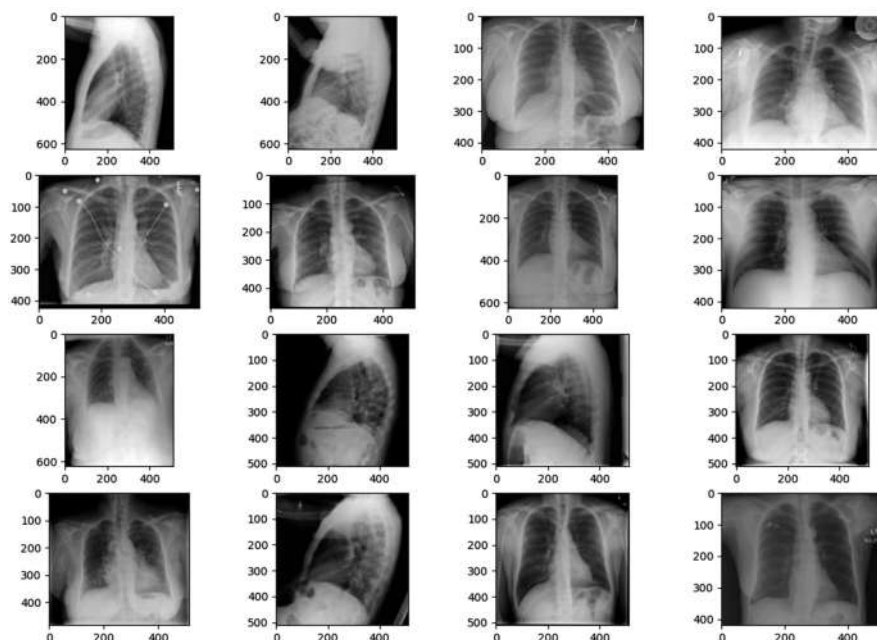


Fig. 1 The dataset contains a set of chest X-ray images

The second segment of the dataset comprises patient reports in XML format. Each XML file represents the report for a specific patient. To identify the images associated with these reports, we need to verify the XML tag `<parent Images id = "image-id">`. The "id" attribute within this tag contains the image names of the corresponding X-ray images (Fig. 2).

XML file has four important attributes: indication, impression, findings, and comparison.

- **Comparison:** This contains details about a sequential follow-up procedure.
- **Indication:** This section includes all the clinical information about the patient.
- **Findings:** It contains data related to X-ray images and tells which part of lung is affected.
- **Impression:** It is generated from indications and findings. It basically tells the medical report result.

4.2 Data Preparation

4.2.1 Data Collection

The dataset is comprised of pairs of X-ray images and corresponding XML files containing medical reports. These XML files contain various attributes such as

```

<?xml version="1.0"?>
<record id="99">
  <source>Chest</source>
  <id>99</id>
  <license>open-access</license>
  <licenseurl>http://creativecommons.org/licenses/by-nc-nd/4.0/</licenseurl>
  <cclicense>by-nc-nd</cclicense>
  <articleid>
    <articleid>2013-06-01</articleid>
    <articleid>99</articleid>
    <publisher>Indiana University</publisher>
    <title>Indiana University Chest X-ray Collection</title>
    <note>The data are drawn from multiple hospital systems.</note>
    <specialty>pulmonary diseases</specialty>
    <subset>Chest</subset>
  <medlinecitation>Indiana University</medlinecitation>
  <article PubMedId>Electronic</article PubMedId>
  <Journal>
    <JournalIssue>
      <Year>2013</Year>
      <Month>06</Month>
      <Day>01</Day>
    </JournalIssue>
    <Journal>
      <ArticleTitle>Indiana University Chest X-ray Collection</ArticleTitle>
    </Journal>
  </Journal>
  <abstract>
    <abstractText Label="COMPARISON">None.</abstractText>
    <abstractText Label="INDICATION">1000-year-old with increasing dyspnea.</abstractText>
    <abstractText Label="FINDINGS">Normal heart and mediastinum. Clear lungs. Trachea is midline. No pneumothorax. No pleural effusion. Radiopaque foreign body overlying left chest.</abstractText>
    <abstractText Label="IMPRESSION">No acute abnormality.</abstractText>
  </abstract>
  <affiliation>Indiana University</affiliation>
  <AuthorList Completion="Y">
    <Author>
      <Name>
        <LastName>Ballou</LastName>
        <ForeName>Marc</ForeName>
        <Initials>MD</Initials>
      </Name>
      <Author>
        <Name>
          <LastName>Roosman</LastName>
          <ForeName>Marc</ForeName>
          <Initials>MD</Initials>
        </Name>
      </Author>
    </AuthorList>
  </AuthorList>
  <Language>eng</Language>
  <PublicationTypeList>
    <PublicationType>Radiology Report</PublicationType>
  </PublicationTypeList>
  <JournalDate>
    <Year>2013</Year>
    <Month>06</Month>
    <Day>01</Day>
  </JournalDate>

```

Fig. 2 Sample of XML report

patient information, image ID, comparison, impression, findings, indication, and more. We will extract the “impressions” attribute from the XML files and consider it as the report. Additionally, we extracted the image ID to retrieve the X-rays that correspond to each report. After collecting data, we have 3955 rows in data-frame.

4.2.2 Data Preprocess

As per our observation, we say that dataset have 3955 rows and 4 columns, and columns are indication, impression, comparison, and images of –X-rays extract from XML files. We see that indication feature have 93 None value, impression feature has 34 None value, and findings feature have 530 None value. As per the observation, note that there are a total of 104 None values in images features. So, we removed total 104 rows in our datasets and replaced with the None values to “no indication” and “no impression.”

During this stage, the text data undergoes preprocessing to eliminate undesirable tags, texts, punctuation, and numbers. Words such as “aren’t” and “shouldn’t” are expanded and converted to lowercase. Stop words are removed, and numerical values are converted into their word equivalents. Special characters, symbols, parentheses, and brackets are eliminated. Words such as “xx” and “xxx” are removed. Extra spaces are also removed.

4.3 Exploratory Data Analytics

In this section we will use different methods of EDA to analyze and visualize the features.

4.3.1 EDA on Text Feature (Fig. 3)



Word-Cloud of findings Feature

4.3.2 EDA on Image Feature (Fig. 4)

Upon analyzing the plot of image features, it becomes evident that a significant number of patients possess two chest X-ray images. The maximum count of images is 5, while the minimum count is 1. Out of the total patients, 3208 have two X-ray images, 446 patients have only one X-ray image, 181 patients have three X-ray images, 15 patients have four X-ray images, and a solitary patient has five X-ray images.

To ensure consistency in the image data points, adjustments need to be made for data points with either more than two images or less than two images. Our approach will involve setting two images in every data point. If a data point contains a single X-ray image, we will duplicate it to create a pair.

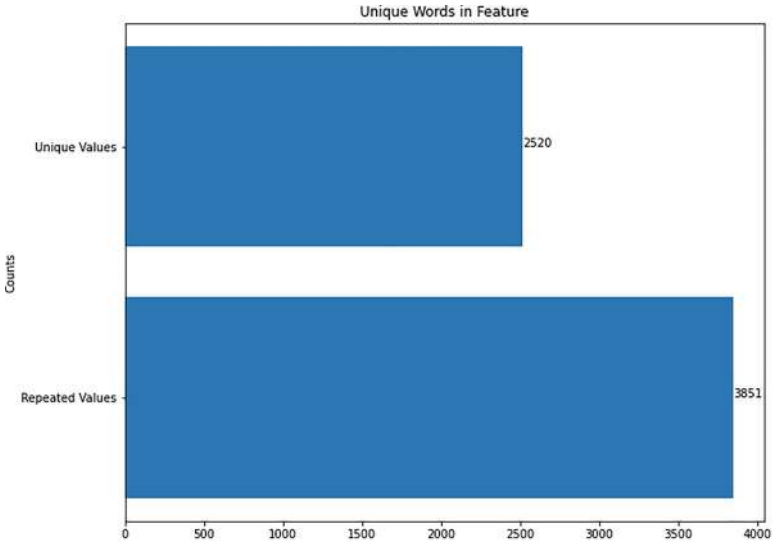


Fig. 3 For the text data analysis, we use findings features

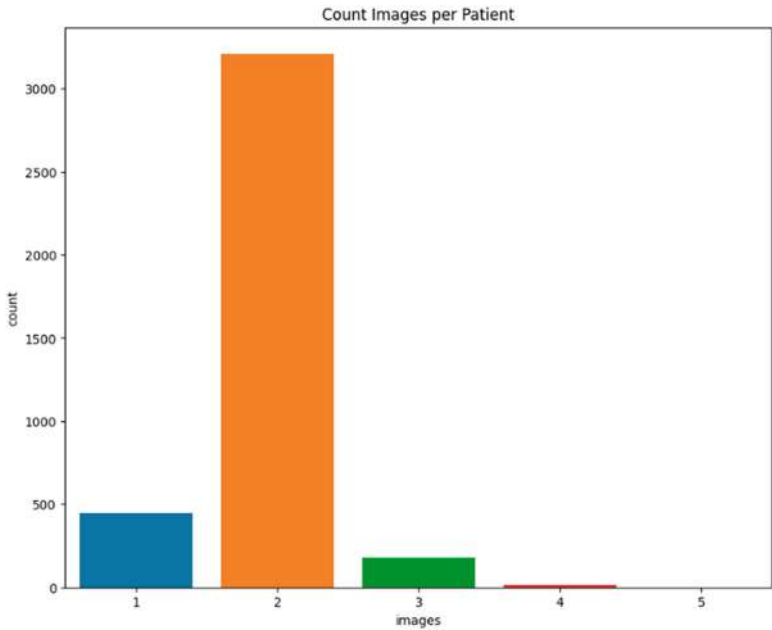
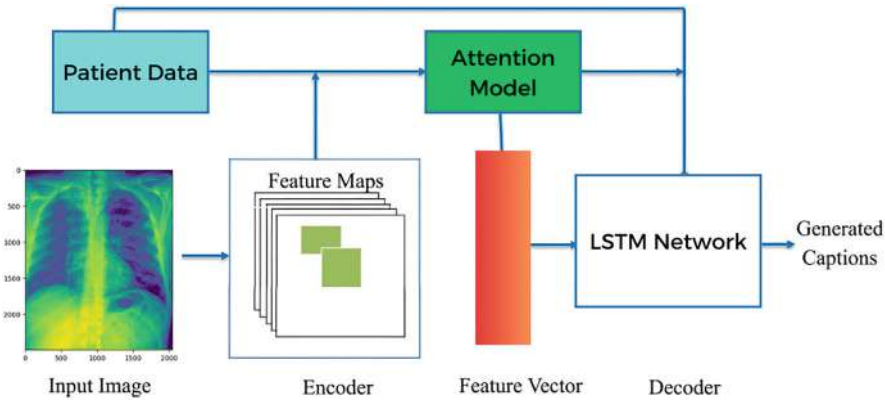


Fig. 4 Analyzing total image per patient

Image count	Edited set
1	1st (frontal or lateral + duplicate)
2	1st (frontal) + 2nd (lateral)
3	1st (frontal) + 3rd (lateral)2nd (frontal) + 3rd (lateral)
4	1st (frontal) + 2nd (lateral)3rd (frontal) + 4th (lateral)
5	1st (frontal) + 2nd (lateral)3rd (frontal) + 4th (lateral)1st (frontal) + 5th (lateral)



5 Methodology

5.1 ML Formulation

We have decided to subdivide the problem into two modules: the first one is the Encoder module, which is used for feature extraction from images: using transfer learning we extract data from X-ray images; the second part is Decoder, which will take the images as input and provide medical reports as output.

So, we will process the image data into textual data, and for this operation we will use Bleu score (Bilingual Evaluation Understudy); Bleu score value lies between 0 and 1, if the actual report and predicted report are same, then the score is 1.

5.2 Basic Model (Encoder–Decoder)

5.2.1 Add Tokens in the Textual Data

Once we have generated new data points from the existing ones, we will proceed to incorporate the <start> and <end> tokens into the text data. Subsequently, we will divide the data into training, testing, and validation sets.

5.2.2 Tokenization

Since machines primarily comprehend numerical values, it is not possible to directly feed text data into deep learning and machine learning models. Therefore, we need to convert the text data into numerical form. To accomplish this, we use a Tokenizer, which transforms the textual information into numerical representations.

The total vocabulary size, or the number of unique words present, is 1312. Additionally, the maximum length of the output sentence is determined to be 60.

5.2.3 Embedding Matrix

For the embedding layer, we used an embedding matrix. The tokenized text data was fed into this layer, which converted each individual token into a 300-dimensional vector. During our analysis, we discovered that the pretrained GloVe-6B-token model encompasses 86.20% (1131 words) of the total 1312 words present in our dataset. This indicates that we would lose information for a total of 181 words. To address this issue, we incorporated the FastText [11] model into our approach. We trained the FastText model using our text data, which employs n -grams to process the text and converts each token into a 300-dimensional vector.

5.2.4 Image Feature

For feature extraction from images and tokenization of text data, we employed the transfer learning technique. Specifically, we utilized the InceptionV3 [10] model, which has been pretrained on the ImageNet dataset and commonly used for image classification tasks. However, we made modifications to the Inception model for our specific task. We removed the last layer of the model and instead utilized average pooling.

To create image tensors for all available images, we applied feature vectorization using InceptionV3. The resulting image tensors were converted into a shape of (1, 2048) and concatenated together for both the Frontal and Lateral images.

5.2.5 Encoder

Once the two concatenated image tensors are passed to the encoder, the shape of the input tensor becomes (1, 4096). Next, we feed the image features into a dense layer to convert the last dimension. This allows us to concatenate the image features with the text data.

5.2.6 Decoder

Our model architecture consists of an embedding layer, an LSTM layer, and a dense layer. The output shape of the dense layer is (batch_size, vocab_size). The LSTM layer stands for “long short-term memory” networks, which is a type of recurrent neural network known for its ability to capture long-range dependencies in sequential data.

5.2.7 Model Training (Fig. 5)

In the training, a “start-of-sequence” token can be used to start the training process and the generated word in the output is used as input on the next step.

In this iterative process, the output of each step is used as the input for the subsequent step until the model reaches the “end-of-sequence” condition. This recursive procedure continues until the desired sequence is generated or the end token is encountered.

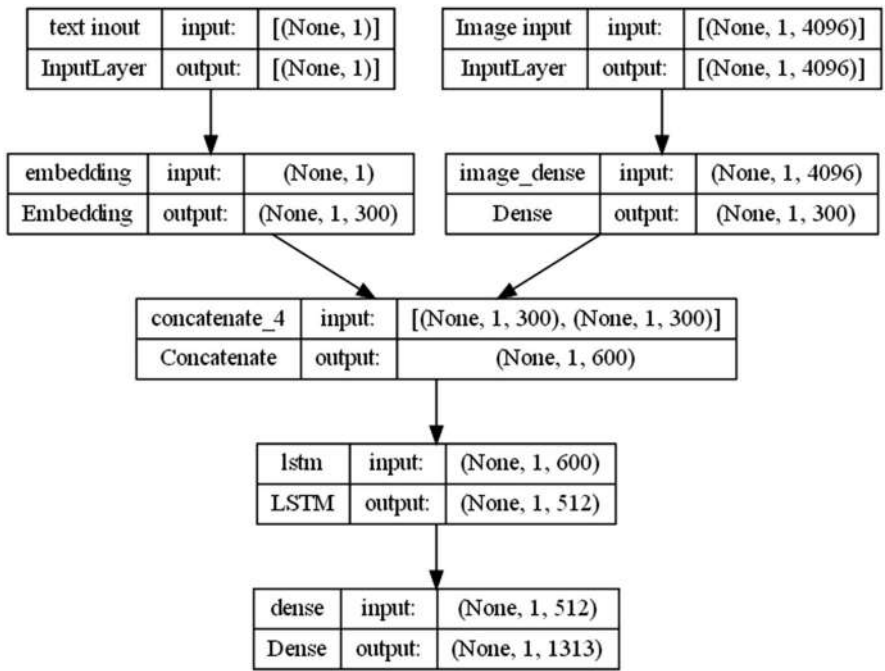


Fig. 5 For the training part, we use the “teacher forcing” method

5.2.8 Model Performance (Fig. 6)

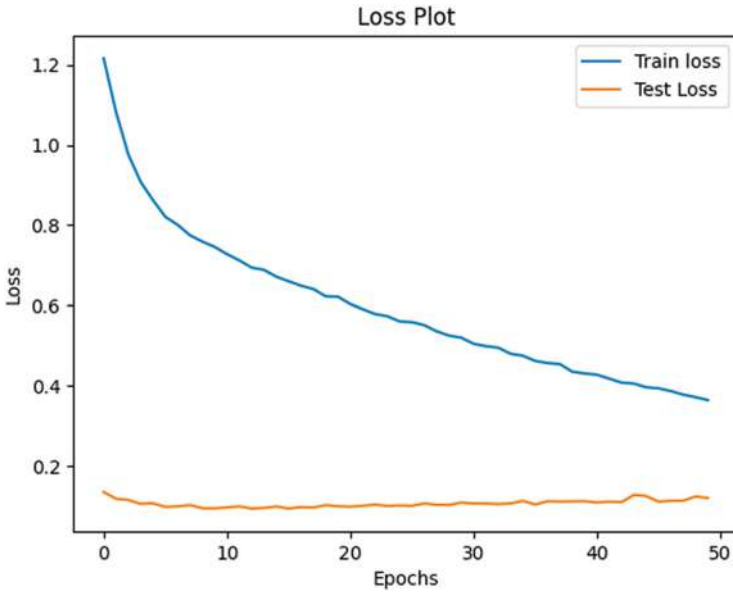


Fig. 6 For the model performance, we have the loss using model history

5.2.9 Model Evaluation

During the evaluation phase, we employed a greedy search approach based on the “teacher forcing” method to generate the output report. At each time step (t), we initialized the word generation process using the <start> token. The predicted word was then fed back as input to the decoder in the subsequent time step ($t + 1$). This iterative process continued until the completion of the output report.

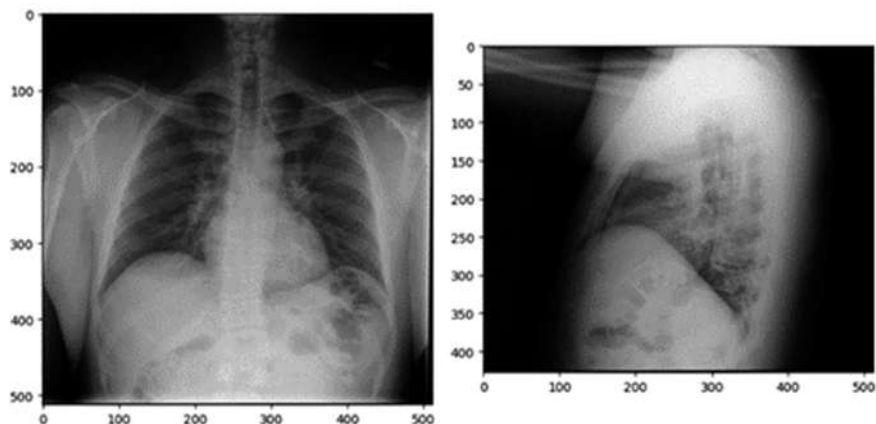
5.2.10 Sample Output

We observe in the above predictions, our model does not perform good in long sentences. Our model performed best in short reports and worst performed in long reports. Some medical reports are meaningless and not useful. Our model performs well in short sentences. This is a basic model, and we used the LSTM model with inception-generated tensors, and we will use attention mechanisms to improve the long sentence dependency. We will also use CheXnet pre-trained model as an encoder with attention model, and we will modify pretrained for our use.

```

*****
Actual <start> low lung volumes otherwise no definite acute findings <end>
Predicted:  no acute cardiopulmonary abnormality <end>
*****
Individual 1-gram: 0.1839 Cumulative 1-gram: 0.1839
Individual 2-gram: 0.0000 Cumulative 2-gram: 0.0000
Individual 3-gram: 0.0000 Cumulative 3-gram: 0.0000
Individual 4-gram: 0.0000 Cumulative 4-gram: 0.0000

```



5.3 Main Model (Encoder–Decoder with Attention)

Attention networks are widely used in deep learning due to their ability to selectively focus on specific parts of the input that are relevant to the task at hand. In computer vision tasks, attention can be used to prioritize certain pixels over others, while in natural language processing tasks such as machine translation, attention can be used to prioritize certain words over others. A research paper can be consulted to learn more about attention mechanisms. For the attention model, the same textual data preprocessing used in the basic model will be utilized, but the maximum output length will be increased to 91 from 60. A pretrained DenseNet121 model with 121 layers of convolutional layers will be used, with pretrained weights provided in the CheXNet [12] file. These weights have been trained on a large dataset of over 100,000 frontal-view X-ray images with 14 diseases, making them suitable for the task at hand. The input layer for the image will concatenate both the frontal and lateral images into a single image, which will then be passed into the image model (CheXNet model) to produce a (None, 9, 9, 1024) shaped tensor. An embedding layer will be used to map each word into a 300-dimensional vector, and a Bi-GRU layer will be utilized to extract high-level information from the attention output. The Bi-GRU layer is able to gather information both forwards and backwards. The attention layer generates a weight vector that is utilized to combine word-level features from each time step, resulting in a sentence-level feature vector obtained by performing element-wise multiplication with the weight vector. Subsequently, the output layer is configured with units corresponding to the size of the vocabulary (Fig. 7).

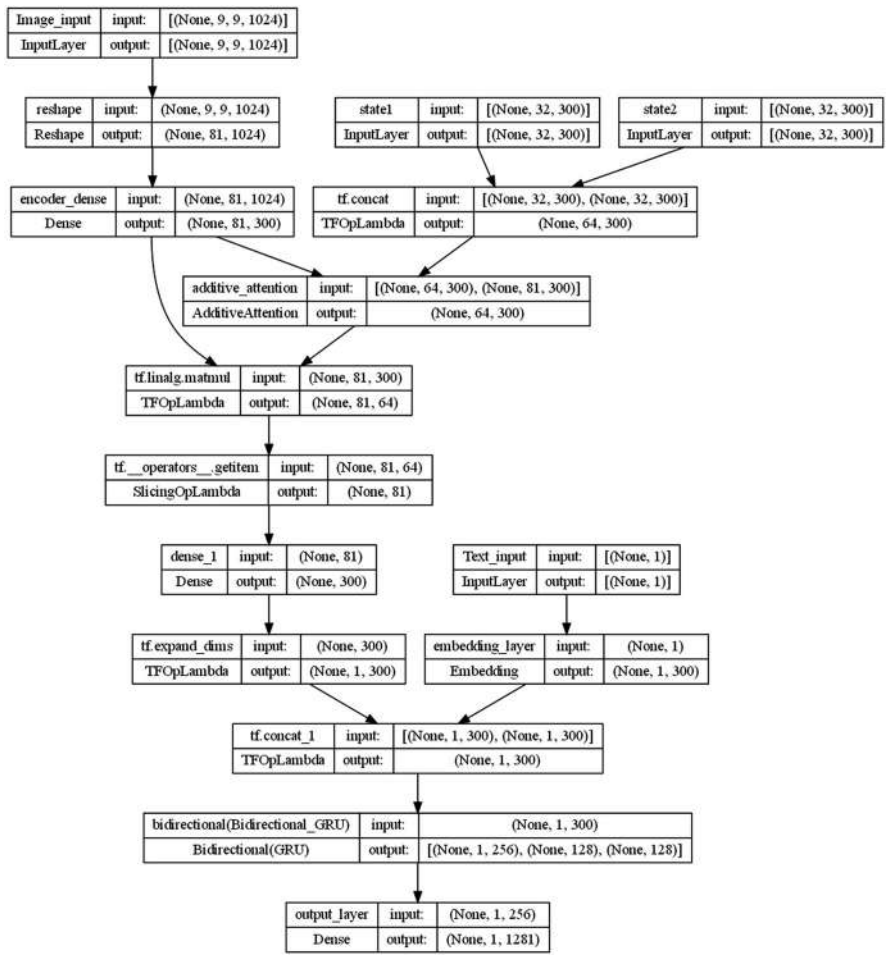


Fig. 7 At each time step, the model predicts a vector of this size, enabling the generation of output predictions

5.3.1 Encoder Model

The encoder class is responsible for transforming the image features from the CheXNet model, which have a shape of $(None, 9, 9, 1024)$, into a lower-dimensional tensor with dimensions $(None, 81, \text{embedding_dim})$. This conversion allows for a more compact representation of the image features while preserving their essential information.

5.3.2 Attention Model

The attention class utilizes the previous hidden state of the decoder model and the encoder output to compute the attention weights and context vector. By leveraging the encoder output and the decoder’s previous hidden state, the attention mechanism determines the importance of different parts of the input sequence, enabling the model to focus on relevant information while generating the output.

5.3.3 Decoder Model

Within the Decoder class, four inputs are passed at every time step. These include the decoder input (which is based on “teacher forcing” method), the encoder output (which is the image tensor produced by the encoder), and the forward and backward hidden states from the previous time step. The Decoder then produces a dense vector with a length equivalent to the vocabulary size, as well as attention weights and the forward and backward hidden states from the GRU.

5.3.4 Model Performance (Figs. 8 and 9)

5.3.5 Model Evaluation

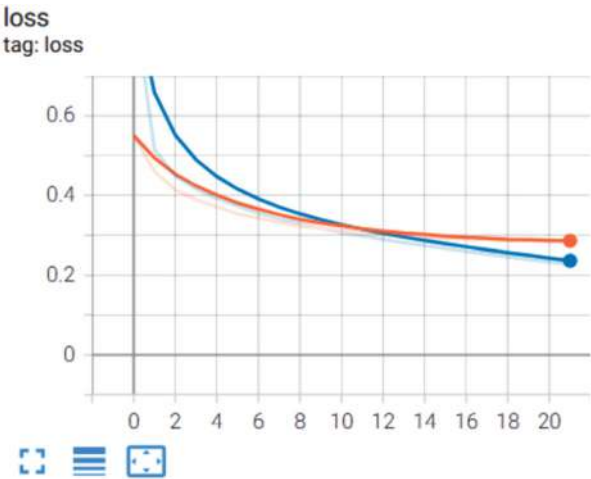


Fig. 8 Train and test loss of the model

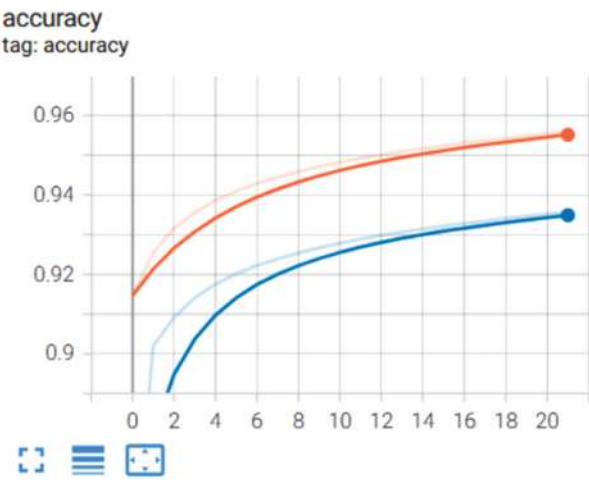


Fig. 9 Train and test accuracy of the model

Greedy Search

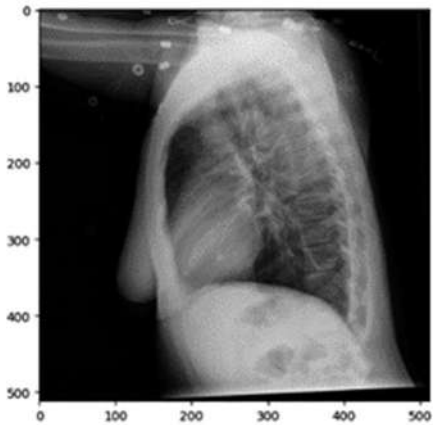
For our evaluation, we will conduct a Greedy Search and select words with the highest probability. Using the arg-max function, we will obtain words with higher probability.

Sample Output

```
*****
Actual <start> no acute cardiopulmonary abnormality <end>
Predicted: no acute cardiopulmonary abnormality <end>
*****
Individual 1-gram: 1.0000 Cumulative 1-gram: 1.0000
Individual 2-gram: 1.0000 Cumulative 2-gram: 1.0000
Individual 3-gram: 1.0000 Cumulative 3-gram: 1.0000
Individual 4-gram: 1.0000 Cumulative 4-gram: 1.0000
```



```
*****
Actual <start> one no acute cardiopulmonary disease <end>
Predicted: no acute cardiopulmonary abnormality <end>
*****
Individual 1-gram: 0.5841 Cumulative 1-gram: 0.5841
Individual 2-gram: 0.5192 Cumulative 2-gram: 0.5507
Individual 3-gram: 0.3894 Cumulative 3-gram: 0.4929
Individual 4-gram: 0.0000 Cumulative 4-gram: 0.0000
```



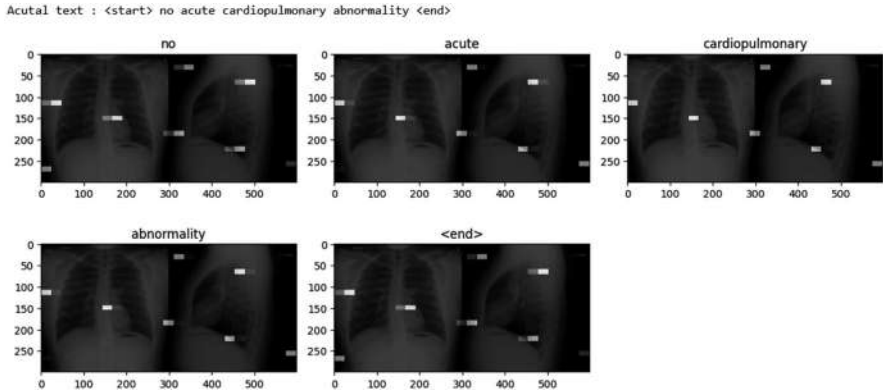


Fig. 10 Attention locations

5.3.6 Attention Plot

We receive attention weight at each time step, allowing the model to select only the pertinent portions of the encoding. Consider some elements to be more significant than others (Fig. 10).

6 Results and Discussions

6.1 Discussion

Medical diagnosis prediction is an important topic that has received considerable attention in the field of healthcare. There are a variety of approaches to predicting medical diagnoses, ranging from traditional statistical models to machine learning and artificial intelligence.

One of the key challenges in medical diagnosis prediction is the large and complex nature of medical data. This can include a range of different types of data, including patient medical histories, lab results, imaging data, and more. To make accurate predictions, it is important to develop algorithms that can effectively process and analyze this data.

Another important consideration in medical diagnosis prediction is the need to ensure that predictions are accurate and reliable. This requires careful validation of models, as well as ongoing monitoring and refinement as new data becomes available.

There are many potential applications for medical diagnosis prediction, including early detection of diseases, personalized treatment planning, and more effective patient care. However, there are also important ethical considerations that must

be taken into account, such as ensuring patient privacy and avoiding bias in data analysis and modeling.

Overall, discussions on medical diagnosis prediction can cover a wide range of topics, including the development of new algorithms and models, validation and testing of these models, ethical considerations, and practical applications in healthcare. As technology continues to advance and more data becomes available, it is likely that this topic will continue to be an important focus of research and development in the healthcare industry.

6.2 Results

Improved patient outcomes: Accurate diagnosis prediction can lead to more effective treatments, earlier interventions, and better patient outcomes.

Reduced healthcare costs: Accurate diagnosis prediction can reduce the need for unnecessary tests and procedures, leading to lower healthcare costs.

Personalized treatment: Accurate diagnosis prediction can help healthcare professionals tailor treatment plans to the specific needs of each patient, resulting in more effective treatment and better outcomes.

Improved public health: Accurate diagnosis prediction can help identify emerging diseases and outbreaks, allowing public health officials to take appropriate measures to prevent the spread of disease.

Better allocation of resources: Accurate diagnosis prediction can help healthcare organizations allocate resources more effectively, ensuring that patients receive the care they need.

Increased efficiency: Accurate diagnosis prediction can help healthcare professionals make more efficient use of their time and resources, allowing them to focus on the most critical cases (Table 1).

6.3 Description of the Existing Solution

In this section, the existing solution, CheXNet, is discussed.

CheXNet is a convolutional neural network (CNN) architecture specifically designed for chest X-ray analysis. It was developed to assist in the detection and classification of various thoracic abnormalities present in chest X-ray images.

The motivation behind creating CheXNet was the need for an accurate and reliable tool to aid radiologists in interpreting chest X-ray images. The interpretation of chest X-rays is a complex task that requires expertise and experience. CheXNet aims to provide an additional level of analysis and support to radiologists by automating the process and reducing human error.

Table 1 Justification

S. no.	Title	Author	Model used	Accuracy
1.	Generating radiology reports using deep learning	Michael B. Gotway and Jie Cao	CNN and LSTM network	BLEU score: 0.31
2.	Automated report Generation for chest X-ray	Shoaib Ahmed Siddiqui, Muhammad Abbas, Muhammad Asim, and Abdul Majid	Convolutional neural network (CNN)	F1 score: 0.82 Overall accuracy: 91.4%
3.	Automatic report Generation for abdominal CT scans using deep learning	Tzu-Yen Huang, Wei-Chih Huang, Chung-Ming Chen, and Shih-Ping Wang	CNN and recurrent neural network	BLEU-4 score: 0.21 ROUGE-L score: 0.31
4.	Medical image report generation using attention-based encoder-decoder architecture	Nishtha Prasad, Deepak Kumar, and Santanu Chaudhury	CNN, RNN, and attention mechanism	BLEU-4 score: 0.30 ROUGE-L score: 0.44
5.	Radiology report generation using graph-based attention networks	Wei Qian, Yanyan Lan, and Jiafeng Guo	Graph based attention network (GAN)	BLEU-4 score: 0.36 ROUGE-L score: 0.54
6.	Generating radiology reports from clinical images using sequence to sequence models	Senthil Kumaran Selvaraj, Omkar M. Parkhi, and Venkatesh Babu R.	Sequence-to-sequence (Seq2Seq) model used in NLP	BLEU-4 score: 0.25 ROUGE-L score: 0.31
7.	Deep learning-based chest X-ray report generation: A review	Pranjal Dey, Sandipan Sarkar, and Heng-sheng Chiang	CNN, RNN, and encoder-decoder model	Comparison study of high and low accuracy models

The architecture of CheXNet is inspired by DenseNet, which is a densely connected CNN. It consists of several convolutional layers, followed by fully connected layers. The key feature of CheXNet is its ability to simultaneously classify multiple abnormalities present in a single chest X-ray image. This multi-label classification approach allows CheXNet to detect and classify a wide range of thoracic abnormalities, including pneumonia, pneumothorax, and more.

Training CheXNet requires a large dataset of labeled chest X-ray images. The network leverages transfer learning by first pretraining on the ImageNet dataset, which is a large-scale dataset covering a broad range of object categories. This initial pretraining helps CheXNet learn general features and patterns from a diverse set of images. After pretraining, CheXNet is fine-tuned on the chest X-ray dataset, adapting the learned features to the specific domain of thoracic abnormalities.

CheXNet has shown promising results in thoracic abnormality detection. In research studies, it has achieved performance comparable to, and in some cases even exceeding, that of experienced radiologists. Its ability to accurately identify abnormalities in chest X-ray images can aid in early diagnosis, triage, and treatment planning, potentially improving patient outcomes (Table 2).

6.4 Improvement

There is always room for improvement in the accuracy and reliability of medical diagnosis prediction algorithms. Future work could focus on developing new techniques that can handle larger and more complex datasets while also improving accuracy and reliability. Another potential area of future work is in the development of personalized medicine. This involves using patient-specific data to develop treatment plans that are tailored to the individual needs of each patient. Machine learning and AI techniques could be used to analyze large amounts of patient data and identify personalized treatment options. The ultimate goal of medical diagnosis prediction is to improve patient outcomes. Future work could focus on identifying ways to use prediction algorithms to improve patient care, such as by identifying patients at risk of developing certain diseases and intervening early to prevent or treat these conditions. As with any technology, there are important ethical considerations that must be taken into account in the development and use of medical diagnosis prediction algorithms. Future work could focus on identifying and addressing potential ethical issues, such as ensuring patient privacy and avoiding bias in data analysis and modeling. Medical diagnosis prediction algorithms can be used in conjunction with other healthcare technologies, such as telemedicine and remote monitoring. Future work could explore how these technologies can be integrated to improve patient care and outcomes.

Table 2 Comparison between both the models

Parameters	CheXNet	Our model (CheXNet + InceptionV3)
Architecture	Here CheXNet alone is used, which is specifically designed for chest X-ray analysis	CheXNet and InceptionV3 have different architectures, with CheXNet having a specific design for chest X-ray analysis and InceptionV3 being a general-purpose CNN. Our model combines the strengths of both architectures to generate X-ray reports, with CheXNet for initial analysis and Inception V3 for further feature extraction and classification
Training data	It uses labeled chest X-ray datasets for training. CheXNet was trained on a large dataset of labeled chest X-ray images with specific thoracic abnormalities.	It also uses labeled chest X-ray datasets for training. CheXNet was trained on a large dataset of labeled chest X-ray images with specific thoracic abnormalities, while Inception V3 was pretrained on the ImageNet dataset, which covers a wide range of general images.
Performance	Using CheXNet alone leads to good performance in X-ray report generation, but does not provide much accuracy in feature extraction and classification	Combining the strengths of CheXNet and Inception V3, it potentially leads to a better performance in X-ray report generation. CheXNet provides accurate detection of thoracic abnormalities, and InceptionV3 further extracts features and classifies images into multiple categories
BLEU score	0.3577200	0.383689
Loss	1.5302	0.283540
Resource efficiency	Using only CheXNet lead to faster inference times and lower computational complexity	The combination of CheXNet and Inception V3 lead to higher computational complexity and longer inference times
Flexibility	It is not as flexible in terms of fine-tuning for different abnormalities or adapting to new datasets	Our model provides flexibility in fine-tuning the model for different thoracic abnormalities or adapting it to new datasets

7 Conclusion

Medical diagnosis prediction is an important area of research and development in the healthcare industry. With the availability of large and complex medical data, machine learning and artificial intelligence techniques are being used to develop accurate and reliable algorithms that can predict medical diagnoses. There are many potential applications for medical diagnosis prediction, including early detection of diseases, personalized treatment planning, and more effective patient care. However, there are also important ethical considerations that must be taken into account to ensure patient privacy and avoid bias in data analysis and modeling. Medical diagnosis prediction has the potential to revolutionize healthcare, improve patient outcomes, and save lives. As research and development continue in this area, it is likely that we will see continued progress and innovation in the field of medical diagnosis prediction.

In conclusion, the model that was built using Bidirectional GRU with attention appears to perform better than the basic model. Specifically, the main model is capable of accurately predicting both short and long sentences, which is an improvement over the basic model. Additionally, the loss has converged to 0.28 with a 93.5% accuracy rate in the training data and a 95.5% accuracy rate in the validation data. These results indicate a strong similarity between the predicted and actual outputs. Moreover, the BLEU score is 0.383689.

References

1. Automated Report Generation for Chest X-Ray Images using Convolutional Neural Networks by Shoaib Ahmed Siddiqui, Muhammad Abbas, Muhammad Asim, and Abdul Majid
2. K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078, 2014.
3. Generating Radiology Reports Using Deep Learning“ by Michael B. Gotway and Jie Cao
4. Automated Report Generation for Chest X-Ray Images using Convolutional Neural Networks by Shoaib Ahmed Siddiqui, Muhammad Abbas, Muhammad Asim, and Abdul Majid
5. Automatic Report Generation for Abdominal CT Scans using Deep Learning by Tzu-Yen Huang, Wei-Chih Huang, Chung-Ming Chen, and Shih-Ping Wang
6. Medical Image Report Generation using Attention-based Encoder-Decoder Architecture” by Nishtha Prasad, Deepak Kumar, and Santanu Chaudhury
7. Radiology Report Generation using Graph-based Attention Networks by Wei Qian, Yanyan Lan, and Jiafeng Guo
8. Generating Radiology Reports from Clinical Images using Sequence to Sequence Models“ by Senthil Kumaran Selvaraj, Omkar M. Parkhi, and Venkatesh Babu R
9. Deep Learning-based Chest X-ray Report Generation: A Review” by Pranjal Dey, Sandipan Sarkar, and Heng-Sheng Chiang.
10. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016, June). Rethinking the inception architecture for computer vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr.2016.308>

11. A. Filonenko, K. Gudkov, A. Lebedev, I. Zagaynov, and N. Orlov, "FaSText: Fast and Small Text Extractor," in *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, Sep. 2019. Accessed: Jun. 15, 2023. [Online]. Available: <https://doi.org/10.1109/icdarw.2019.30064>
12. S. Kora Venu, "An Ensemble-based Approach by Fine-Tuning the Deep Transfer Learning Models to Classify Pneumonia from Chest X-Ray Images," in *Proceedings of the 13th International Conference on Agents and Artificial Intelligence*, 2021. Accessed: Jun. 15, 2023. [Online]. Available: <https://doi.org/10.5220/0010377403900401>

Detecting Hard Landing of Flights: E-Pilots



Revelle Akshara, Nomula Aishwarya, M. Karuna Shree, and N. Lahari

Abstract A common type of landing accident is hard landing. Hard landing, a phase where landing of flights is difficult, which can also make the passengers life to risk. Early detection of this phase results in safe landing. To predict whether it is hard landing or not, an idea has been introduced to obtain best prediction. Hybrid LSTM (long short-term memory) algorithm was used. The study presents a machine learning system that can be used in the cockpit to help the flight crew make decisions about go-arounds based on likelihood of a hard landing. To forecast challenging touchdowns, here the study offers a hybrid approach that uses attributes that model the temporal dependencies of aircraft data as inputs to a semantic network. A machine learning model for cockpit will read data from flight such as tyre elevation, speed, and other values to predict the type of landing. In this LSTM, the features such as Pilot (DH2TD), Actuator (AP2DH), and Physical (AP2TD) together make hybrid LSTM. If predicted hard landing, then it instructs to avoid landing or divert to other route. In the existing method there was a result of 85% sensitivity and 74% specificity, which improve with the hybrid LSTM to a sensitivity of 95% and specificity of 96%.

Keywords LSTM · AP · TD · DH · QAR · EASA · Aircraft · ACARS · KNN · RNN · R-FTG

1 Introduction

There are many well-known precursors of path excursions during landing, according to EASA. They include shaky technique, a difficult touchdown, an unusual mentality or bounce, aeroplane side differences at arrival on the ground, and a limited rolling range at landing. Unpredictable approach is the main precursor. Just 3%

R. Akshara (✉) · N. Aishwarya · M. K. Shree · N. Lahari
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, India

of approaches in commercial airline operations, according to Boeing, satisfied the criteria for an unexpected approach, but 97% of those approaches continued to landing rather than performing a go-around. In their 16-year analysis period, Blajev and Curtis' study found that a go-around decision may have prevented 83% of runway excursion accidents [2]. Despite a clear strategy and training on go-around plans being provided by the majority of airline companies, operational data indicates that the trip team's choice to request a go-around may be influenced by a number of additional factors. The main goal in the European Plan for Aviation Security is the usage of artificial intelligence in aviation industry. The assumption behind the system is that a difficult landing occurrence can be anticipated and predicted.

2 Literature Survey

A tough touchdown (TD) is a phenomena where the aeroplane impacts the ground too forcefully right after touchdown [1]. Systems can be made to make search on risk of aircrafts overruns. There are already models that show chances of overruns, location of ruins, and damage consequences. However, these models work more effectively if were given more records and accurate data [3, 4]. Deep learning approaches are also utilized [8]. ACARS data is used to find and forecast the hazard in aircraft. Here, it shows that technique based on SVM is giving more accuracy than other methods, which gives an identification accuracy of 90.99%. A 30-day look back time with long short-term memory (LSTM) and comparing algorithms such as KNN, RNN, and BPNN gives good prediction performance and low prediction error [7]. The study of 62 features of flight QAR data, which are roughly chosen for making a model of LSTM. SVM, BP, and LR, are employed to compare the F1 score with F1 score of LSTM, which shows that LSTM predicts more effectively and reached high accuracy rate [9, 17]. Prediction model based on SVM is made, which predicted a sensitivity of 0.4678 with original data. However, by adding three factors, sensitivity raised to 0.5274, which shows prediction with SVM gets great results [5]. The purpose of this research was to develop and test predictive models for unstable approach risk misperception in the National Airspace System using machine learning. Once evidence of unstable approaches was identified and extracted from the flight recorder data, a determination was made whether a rejected landing or continuance to landing was made. This model was able to predict the pilot error of continuing an unstable approach to landing with an accuracy of 98% [6]. Airlines can avoid most runway excursions if flight crews choose to execute a go-around maneuver instead of continuing an unstabilized approach to a landing [10, 11]. The adaptability of the proposed R-FTG was also confirmed in the simulations at different bank angle constraints. The numerical simulations with wind effects show that flight trajectories can be generated corresponding to changes in wind direction. When these factors are taken into account, we compare our method to the LSTM version as well as two other often-used versions that have been found [14]. The dataset was utilized to retrain LSTM, SVM, and LR models by incorporating

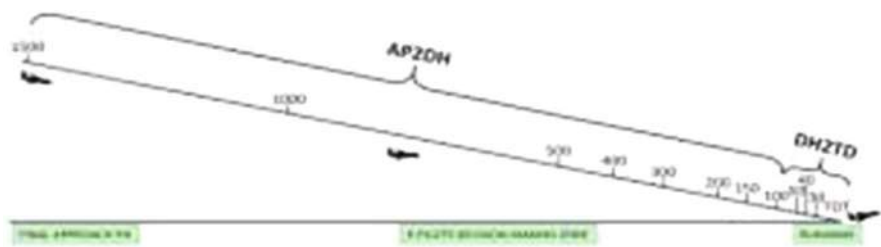


Fig. 1 Altitude phases

the variables and measures suggested by the research [12–13, 15–16]. Following the same approach, we developed an LSTM model that includes a fully connected layer to facilitate classification. Depicts the architecture of proposed system. This model was trained using data from 9 s of testing, specifically from the 2nd to the 10th second prior to touchdown. In the above-mentioned study, there is no indication for the numbers of hyper-parameters, so we manually set the size and discovery rate to 8 and 0.0001, respectively.

3 Proposed System

To forecast whether a landing would be hard or not, the hybrid LSTM algorithm has been proposed. In the proposed research, a machine learning model for the cockpit reads flight data, including tyre elevation, speed, and other values, and predicts the kind of landing. If a hard landing is predicted, the model instructs the pilot to avoid landing or change the landing route. LSTM is trained using several features, including Pilot (DH2TD), Actuator (AP2DH), and Physical (AP2TD). Figure 1 briefly shows the various phases of landing. Three separate LSTM algorithms that have been trained on the aforementioned three features are combined to create a hybrid model.

4 Methodologies

4.1 Algorithms

SVM (Support Vector Machine)

It is a model that examines the data for classification and regression analysis. This classifies the data based on parameters to create a hyperplane. The hyperplane is a correct fit line that classifies the data more accurately.

LR (Logistic Regression)

A machine learning concept that uses only efficient variables from the observations of the different types of data for classification is known as logistic regression. It predicts the probability of target variable.

Hybrid LSTM (Long Short-Term Memory)

In LSTM, data storing is for long period, which makes the prediction process easy. LSTMs gives us learning rates, output, and input biases. In this LSTM, the features such as Pilot (DH2TD), Actuator (AP2DH), and Physical (AP2TD) together make hybrid LSTM.

Mathematical Equations

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

5 Dataset

Source of Dataset <https://github.com/smzn/Turbulence>

Preprocessing of Dataset There are many preprocessing methods available, but we have chosen normalization as the method for the dataset to preprocess. Normalization can help to improve the performance of machine learning algorithms by scaling the input features to a common scale. This can help to reduce the impact of outliers and improve the accuracy of the model. Normalization is used to scale the data of an attribute so that it falls in a smaller range.

Dataset Description

1. Actuator

- 1.1. *Aileron position*: Ailerons are a primary flight control surface that control movement about the longitudinal axis of an aircraft. This movement is referred to as “roll.”
- 1.2. *Elevation position*: The term aviation is defined as the highest point of the landing area. It is measured in feet.
- 1.3. *Flap position*: A flap is a high-lift device used to reduce the stalling speed of an aircraft wing at a given weight. Flaps are usually mounted on the wing trailing edges of a fixed-wing aircraft. Flaps are used to reduce the take-off distance and the landing distance.
- 1.4. *Rudder degree*: The acute angle between the rudder and the fore-and-aft line of a ship or airplane.

2. *Physical*

- 2.1. *Air speed*: An aircraft’s airspeed is the speed at which it travels through the air. It is measured in knots.
- 2.2. *Drift angle*: The heading of an aircraft may be different than its track due to the wind. This difference is called drift angle.
- 2.3. *Speed to target*: The speed that is maintained while flying down final in the landing configuration.

3. *Pilot*

- 3.1. *Side stick*: An aircraft control stick that is located on the side console of the pilot, usually on the right-hand side, or outboard on a two-seat flightdeck.
- 3.2. *Throttle*: A throttle lever, more often referred to as a thrust lever or power lever, is the means by which the pilot controls the amount of fuel provided to the engine with which it is associated.

6 Working Explanation

To run the project, simply double-click on the “run.bat” file and obtain the output as displayed in Fig. 2.

On the output screen, select the “Upload Flight Landing Dataset” option to transfer the dataset and obtain the output as displayed in Fig. 3.

Select the whole dataset folder with three files in the screen above (i.e., Fig. 3), upload it, and then click the “Choose Folder” button to load the dataset will provide the output as shown in Fig. 4.



Fig. 2 Output screen



Fig. 3 Uploading dataset

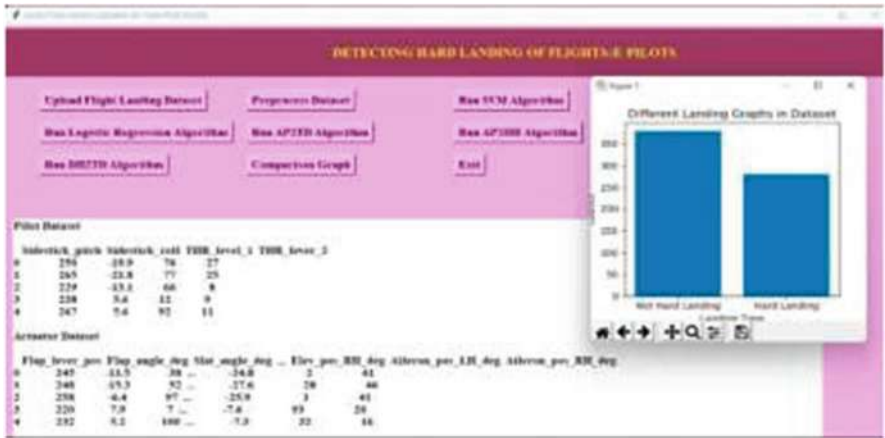


Fig. 4 Physical dataset values

One can scroll down on the text area to view the physical dataset values. The graph shows the type of landing and the counts of landing found in the dataset. The dataset is loaded in the above screen, and we can see some records from the Pilot and Actuator datasets. To normalize, shuffle, and partition the dataset into train and test, click the “Preprocess Dataset” button after closing the previous graph and the results shown in Fig. 5 will result.

In the screen, we can see the total records in the dataset, the total features in all datasets, the total dataset in the pilot dataset, and the size of the training and testing records. Now that the test and training sets of data are prepared, click “Run SVM Algorithm” to train SVM as shown in Fig. 6 to obtain the output.



Fig. 5 Testing and training dataset configuration

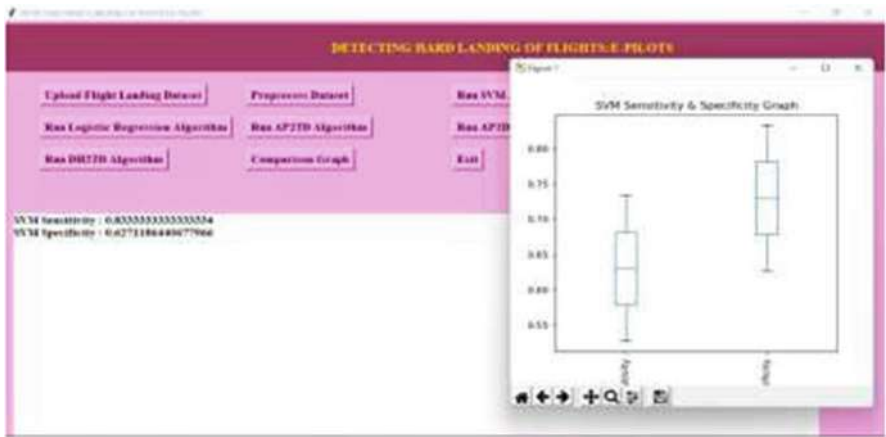


Fig. 6 Results obtained using SVM

Using SVM, we obtained sensitivity of 0.82 and specificity of 0.55 in the screen above. In the box plot, the *x*-axis shows metrics names and the *y*-axis gives values. Close the previous graph, then press the “Run Logistic Regression Algorithm” button to train the logistic regression and obtain the output as shown in Fig. 7.

Using the “Run AP2TD Algorithm” button will train an LSTM on “Physical Features” and produce the output shown in Fig. 8. Using logistic regression, we obtained 0.60% sensitivity values on the screen above in Fig. 7.

Click the “Run AP2DH Algorithm” to train the LSTM on Actuator features and get the output as shown in Fig. 9. In the previous screen using AP2TD physical features, we obtained LSTM sensitivity as 0.92 and specificity as 0.95.

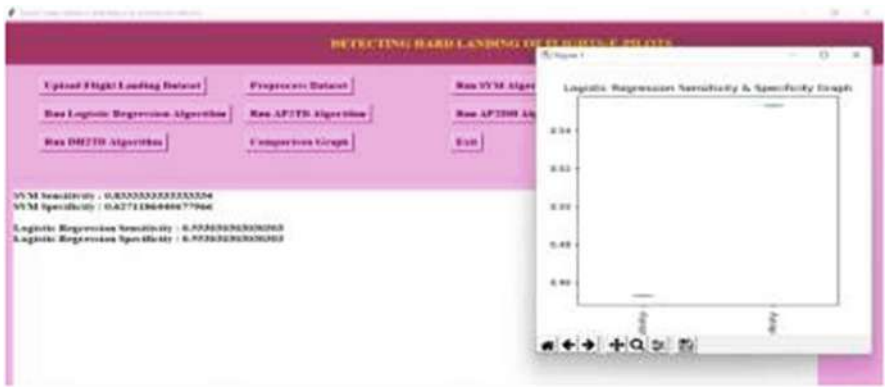


Fig. 7 Results obtained using logistic regression

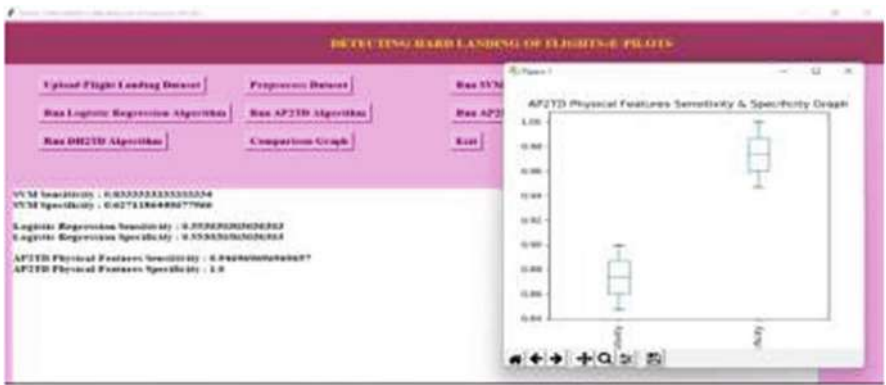


Fig. 8 Results obtained after applying LSTM to AP2TD data

Clicking the “Run DH2TD Algorithm” button will train the LSTM on the Selected features and produce the output as shown in Fig. 10. The AP2DH LSTM on the previous screen has 0.99% sensitivity and 0.98 specificity.

Using DH2TD, we obtained LSTM sensitivity of 0.93 and specificity of 0.92 in the previous screen. Click the “Comparison Graph” button and get Fig. 11 as the comparison graph.

Table 1 shows a detail information about sensitivity and specificity scores of algorithms used here for comparison. The *x*-axis in the plot shows names of algorithms, and the *y*-axis shows sensitivity and specificity levels. Orange bar denotes specificity, while blue bar stands for sensitivity. The proposed AP2TD, AP2DH, and DH2TD models have high sensitivity and specificity values when compared to the LSTM and logistic regression models that are already in use.

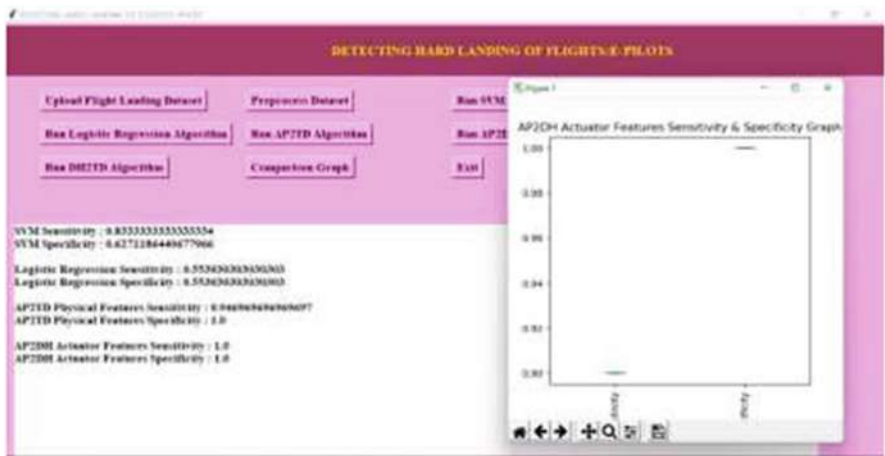


Fig. 9 Results obtained after applying LSTM to AP2DH data

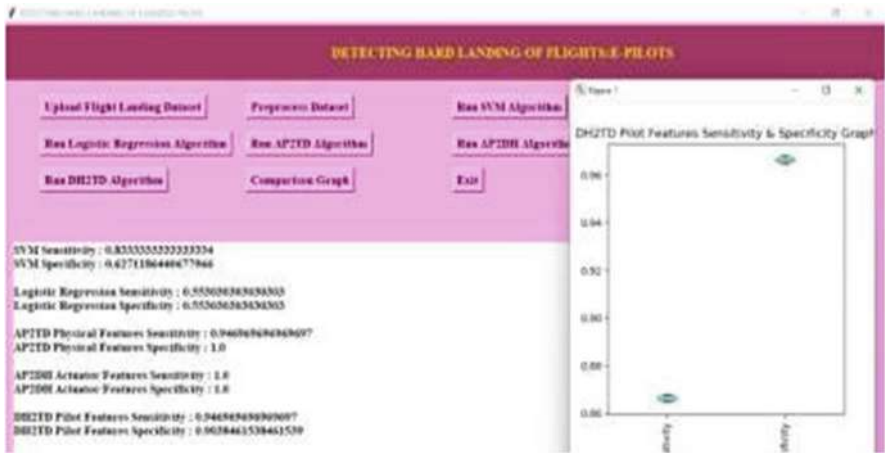


Fig. 10 Results obtained after applying LSTM to DH2TD data

We can examine the sensitivity and specificity values for the hybrid LSTM by integrating all three models in the last section of Fig. 12. Sensitivity was 0.95 and specificity was 0.96 for the hybrid LSTM. These numbers are more in line with the ones in the base paper.

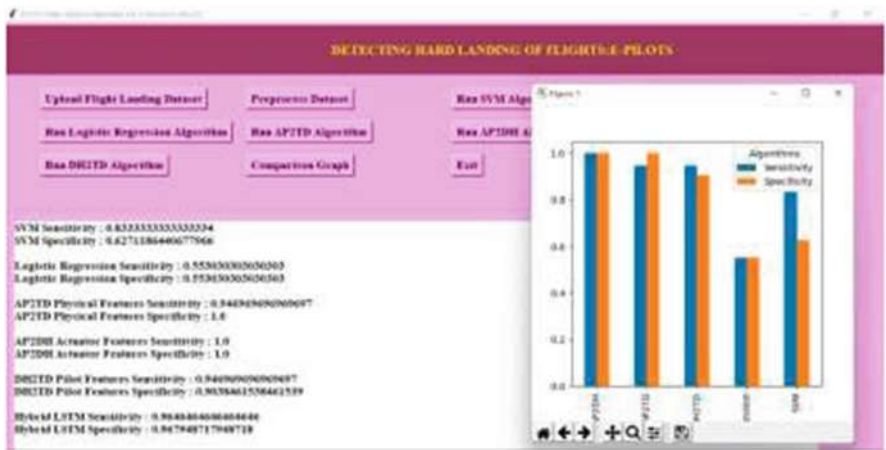


Fig. 11 Comparison graph

Table 1 Comparison of outputs of algorithms used

Algorithms used for comparison	Sensitivity	Specificity
SVM	82	55
Logistic regression	60	51
AP2TD data	92	95
AP2DH data	99	98
DH2TD data	93	92
Hybrid LSTM	95	96

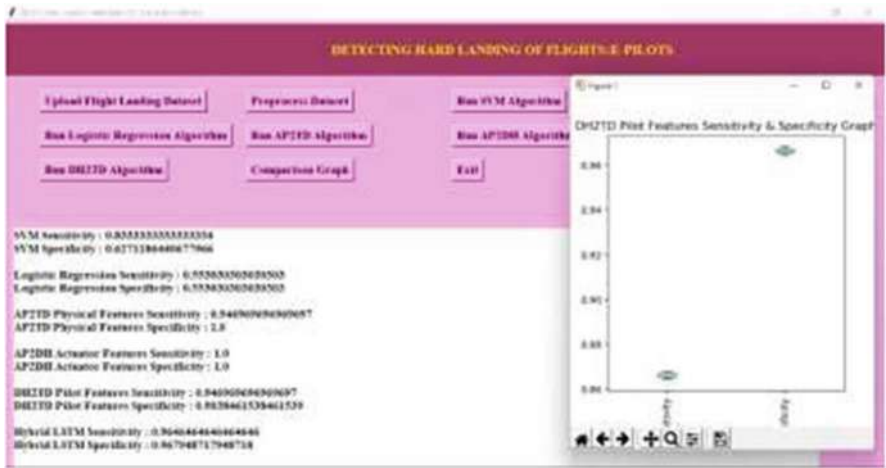


Fig. 12 Sensitivity and specificity for hybrid LSTM

7 Conclusion

The automation features (autopilot, trip manager, and auto-thrust) indicate that these features have little bearing on the likelihood of an HL (hard landing) event and may not even be required to be included in models. Eventually, there are several issues that have not been addressed in this study, which require additional research and future work. It would be necessary to examine such issues, as we foresee doing in future occupations, in an industry as safety-sensitive as aviation. In future, it can include air traffic administration, where information is provided to the air traffic controller to select a safe route and prepare for the most likely situation, roving information, airplane mass, and center of gravity to make the study more better.

References

1. Kirkland, I. D. L., et al. "An improved methodology for assessing risk in aircraft operations at airports, applied to runway overruns." *Safety Science* 42.10 (2004): 891–905.
2. Blajev, Tzvetomir, and W. Curtis. "Go-Around Decision Making and Execution Project: Final Report to Flight Safety Foundation." Flight Safety Foundation (2017).
3. Zhou, Di, et al. "Deep learning-based approach for civil aircraft hazard identification and prediction." *IEEE Access* 8 (2020): 103665–103683.
4. Payan, Alexia P., et al. "Improvement of rotorcraft safety metrics using performance models and data integration." *Journal of Aerospace Information Systems* 14.1 (2017): 26–39.
5. Odisho, Edwin V., Dothang Truong, and Robert E. Joslin. "Applying machine learning to enhance runway safety through runway excursion risk mitigation." *Journal of Aerospace Information Systems* 19.2 (2022): 98–112.
6. Coker, Michael, and Lead Safety Pilot. "Why and when to perform a go-around maneuver." Boeing Edge 2014 (2014): 5–11.
7. Zhang, Haochi, and Tongyu Zhu. "Aircraft hard landing prediction using LSTM neural network." *Proceedings of the 2nd International Symposium on Computer Science and Intelligent Control*. 2018.
8. European Union Aviation Safety Agency. "Artificial Intelligence Roadmap. A Human-centric approach to AI in Aviation." (2020).
9. Hu, Chen, et al. "The study on hard landing prediction model with optimized parameter SVM method." 2016 35th Chinese Control Conference (CCC). IEEE, 2016.
10. Li, Lishuai, et al. "Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring." *Transportation Research Part C: Emerging Technologies* 64 (2016): 45–57.
11. Miwa, Masahiro, et al. "Real-time flight trajectory generation applicable to emergency landing approach." *Transactions of the Japan Society for Aeronautical and Space Sciences* 52.175 (2009): 21–28.
12. Zhou, Di, et al. "Deep learning-based approach for civil aircraft hazard identification and prediction." *IEEE Access* 8 (2020): 103665–103683.
13. Vieira, Fábio Manzoni, et al. "Health monitoring using support vector classification on an auxiliary power unit." 2009 IEEE aerospace conference. IEEE, 2009.
14. Gil, Debora, et al. "E-pilots: A system to predict hard landing during the approach phase of commercial flights." *IEEE Access* 10 (2021): 7489–7503.
15. Cross, Elizabeth, et al. "Prediction of landing gear loads using machine learning techniques." *Structural Health Monitoring* (2012): 1056–1063.

16. Luo, Y. L., and Xiang Zhang. "Aircraft electronic board fault detection based on infrared thermal imaging and integrated SVM." *Meas Control Technol* 12 (2012): 012–012.
17. Wang, Lei, Changxu Wu, and Ruishan Sun. "Pilot operating characteristics analysis of long landing based on flight QAR data." *Engineering Psychology and Cognitive Ergonomics. Applications and Services: 10th International Conference, EPCE 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21–26, 2013, Proceedings, Part II* 10. Springer Berlin Heidelberg, 2013.

Detection of Glaucoma Using MobileNet, XAI, and IML



A. Rakesh, P. Chetan, R. Shailender Raj, E. Ravi Kondale,
and V. Kakulapati 

Abstract Glaucoma causes permanent vision loss. Glaucoma is estimated to affect over 80 million public worldwide in the year 2030. Individuals, physicians, and other health-care professionals realize absolutely nothing regarding the way data inspection and processing work, but contemporary ML (machine learning) models are mostly used for glaucoma prediction. Options for enhancing customer trust along the way are provided by IML (interpretable machine learning) and XAI (explainable artificial intelligence). The chapter discusses the XAI and IML methods for analyzing glaucoma hypotheses and findings. XAI relies heavily on the ANFIS (adaptive neuro-fuzzy inference system) and PDA (pixel density analysis) to provide consistent elucidations for glaucoma forecasts derived from both high-quality audio and image data. To make meaning of data, IML employs a technique called SP-LIME (semi-private selection adjacent interpretable model-agonistic clarification). This is used for the decoding of SNN findings. The implementation consequences demonstrate that XAI and IML techniques provide patients and physicians with convincing and unambiguous recommendations using two specific, readily accessible datasets: understanding scan fundus images and glaucoma case medical information.

Keywords Explainable AI · IML · Pixel density · SP-LIME · Glaucoma · Vision · Loss · Physicians · Fundus · SNN · Pixel density · Patient · Prediction

1 Introduction

Common eye illnesses that cause impairment to the nerve that transmits vision include glaucoma. Injury to the optic nerve leads to this condition, which is another leading cause of disability worldwide. Open-angle eye disease and angle-closure

A. Rakesh · P. Chetan · R. S. Raj · E. R. Kondale · V. Kakulapati (✉)
Sreenidhi Institute of Science and Technology, Hyderabad, India
e-mail: ravikondale@sreenidhi.edu.in

glaucoma are the two primary subtypes. Medication may slow the progression of glaucoma, but it cannot restore the sight lost. It is crucial to recognize the first glaucoma symptoms with regular eye exams. When detected early, acute angle-closure glaucoma may be treated effectively [1].

In glaucoma, the optic nerve undergoes alterations, and the patient also has visual field abnormalities. An important risk factor is elevated IOP (intraocular pressure). A physical assessment of the optic nerve and RNFL (retinal nerve fiber layer) components from fundus pictures, as well as IOP data, is used in the standard diagnostic process for glaucoma by an ophthalmologist. Additional evaluations using OCT (optical coherence tomography) of the optic nerve and RNFL are performing in high-income nations, allowing for evaluations with age-matched normalized data [2].

Artificial intelligence (AI) and machine learning (ML) have made huge commitments to a large number of fields, including hereditary qualities, monetary issues, casual associations, and more. Clinical imaging models and equipment are helpful when looking into problems in the market for medical services. However, the large number of variables and hyper parameters make it difficult to examine and evaluate ML models due to their complexity. AI and ML frameworks need to be fair and easy to understand in order to analyze facts and make decisions that are more understandable.

Incorrect diagnosis of glaucoma is avoidable with make use of AI systems [3], and ML is an essential tool in the fields of life science and medicine. Health-care applications make use of classification, and ML may be put to use for forecasting glaucoma. Forecasting and diagnosing glaucoma will soon rely heavily on AI.

It is possible to identify and cure glaucoma, a leading cause of irreversible blindness, with the use of diagnostic instruments including tannometers, gonioscopy, and scanning laser tomography. When applied to the task of illness classification in glaucoma, deep learning's comprehension of the XAI idea and the use of mobile's pretrained models may eradicate these limitations. Diseased glaucoma has been categorized using pretrained models from ImageNet [4, 5].

This chapter conducted an interpretability evaluation to determine which characteristics were used by the model for glaucoma detection. The model was tested on people who were not diagnosed with glaucoma on the day of imaging but were subsequently found to have the condition. Techniques for interpreting the consequences of ML were used to shed light on the factors contributing to good model accuracy. Finally, the model was verified by checking its predictions against those of experienced medical professionals.

The remaining of this study is organized as literature survey of glaucoma disease discussed in Sect. 2, followed by the description of proposed methodology, models, and techniques in Sect. 3. Whereas Sect. 4 discusses dataset description, and the proposed model and the implementation evaluation have been explained. Finally, Sect. 5 discusses concluding remarks, followed by future enhancement.

2 Literature Survey

The chapter discusses the use of ML algorithms in administration, using a convolutional neural network to test the Shapley Additive Explanations (SHAP) model. The classifier got a normal F1 score of 96.1% on GTEx tests and distinguished 2423 significantly unmistakable highlights, of which 98.6% were perceived by differential enunciation assessment across all tissues. The text also demonstrates the advantages of applying ML to transcriptome information and how SHAP can be utilized to show a profound learning model [6, 7].

Reference [8] discusses the advantages and disadvantages of using ML frameworks to predict personalized treatment outcomes and offer treatment recommendations. It presents options for cutting-edge causal derivation computations and future evaluation moves that could utilize electronic health records (EHRs) and ML [9] to give customized treatment suggestions. It also investigates how exploratory data from randomized controlled trials (RCTs) and pharmacology approaches, as well as pharmacometric and quantitative frameworks, could be used to improve and verify ML computations. To incorporate RCT-based models into these ML models, collaboration across logical areas is needed.

The field of XAI has seen an increase in interest in the production of novel strategies for understanding and unraveling machine learning (ML) models. This evaluation is based on the nearness of ML interpretability, with a writing audit, scientific classification, and links to the changing executions of these methods provided. ML frameworks have been difficult to take on in delicate but imperative enterprises like medical care, so there has been a new ascent in sensible interest in the field. This evaluation is based on the nearness of ML interpretability, with a writing audit, scientific classification, and links to the changing executions of these methods provided [10].

A preamble to interpretable machine learning (iML) with an emphasis on genomic applications was explored [11]. It provides a typology of long-term processes, clear procedures, and major shows. It argues that to fulfill the promise of accurate medicine, iML arrangements are required, but there are still a number of frustrating issues. It examines the limitations of craftsmanship progressions and plans different future assessment decisions. Cross-disciplinary joint effort is expected for continued progress.

The human microbiome has been linked to various illnesses, including asthma, gum sickness, dermatitis, and hurting improvement. To gain a deeper accepting of the microbiome's impact on individual health, an XAI method is used to collect leg skin micro biota testing from two solid female partners. The reasons are made sense of by changes in the absolute flood of significant microorganisms, which drives the projections. Skin hydration, mature, menopause status, and smoking are completely anticipated by the leg skin microbiome. Adjustments in the microbial part of skin dampness might facilitate the advancement of new drugs for sound skin, and those related with might give understanding into the skin's maturing framework. In the association between smoking and menopause, leg microbiome summaries is

basically interchangeable from past disclosures from respiratory group microbiomes and stomach microbiomes [12].

3 Methodology

In this study, two techniques are utilized: (1) fundus picture handling and (2) class name expectation with the assistance of incorporation of clinical information. Overall, managing different classifications and portrayals is not limited to fundus image data.

The MobileNet model family for Tensor Flow is a set of mobile-first computer vision models that use depth-wise distinct convolutions. To accommodate limited resources, these models are tiny, low-latency, and low-power. Each of the layers are evaluated by group norm and ReLU irregularities until the last fully connected layer, which bypasses nonlinearity and subsequently inputs into a Softmax layer for categorization [13, 14].

3.1 XAI and IMLA

For automated decision-making, powerful AI and ML systems are frequently restricted to black-box systems. We utilized a XAI procedure to decipher our expectation discoveries for dependable glaucoma expectation. In conclusion, we used PDA to explain medical information and SP-LIME to evaluate picture attributes.

A desirable AI system would be able to articulate the rationale behind its own actions, highlight its strengths and faults, and predict its future behavior. To reach this goal, cutting-edge methods of interaction between humans and computers are fused with novel approaches to artificial learning to generate models that are easier to explain. The ultimate goal is to provide the user with explanation dialogues that make sense given the format of the model.

In order to take the decision-making process of a sophisticated black-box method, the concept of LIME [15]. By randomly adjusting the samples, LIME creates a whole new dataset. A linear model, which is a local approximation of the black-box model, is trained using the new dataset. The interpretable model is then used to derive the black-box model's local decision behavior.

By providing more context for a model's choice, explainable AI (XAI) methods hope to increase confidence in such judgments. Unfortunately, the explanations provided by the model are not always human-comprehensible, leading to an explainability gap [16]. Better XAI models are the goal of a XAI challenge aimed at the financial services industry, and the same challenges apply to BERT (Bidirectional Encoder Representations from Transformers) [17].

SP-LIME is a worldwide, post-hoc, model-free substitute model that adds usefulness to LIME by choosing explicit circumstances as worldwide choice limits.

The whole SP-LIME information space is addressed by the black-box model and an example of the information dataset. For each example decided to address the element space, it creates a solitary rundown of the commitments made by highlights. Along these lines, clients can assess each of these models all alone while also acquiring an extensive perception of the model from the assortment of clarifications.

SNN and ANFIS's fluffy layer: to improve XAI's logic and interpretability, we focus on forecast methods that use both clinical records and fundus pictures. Despite the fact that certain fundus pictures are most ideal way to foresee glaucoma, clinical information like age and glucose levels ought to be used to make a decision. To make a deal among such data, we use spike mind association (SNN) and flexible neuro-cushioned induction structures (ANFISs). SNN—a black box procedure—works among fundus pictures and perceives thickness about neurons in strong and polluted pictures. To extend importance about results expected through SNN, we use submodular pick-close through interpretable model-agonistic explanation (SP-LIME) and pixel thickness assessment (PDA). SP-LIME and PDA, which make sense about and decipher SNN results, are parts of IML system. In view of neuron thickness values and clinical records, ANFIS decides seriousness about glaucoma. The aforementioned study's illustrative goal is to make sense about how glaucoma risk factors collaborate among each other. XAI techniques also provide a clarification on the seriousness of fundus photography.

3.2 PDA (*Pixel Density Analysis*)

It is used to sort out the amount of pixels in the picture practical. For glaucoma imaging, PDA gives solid translations of glaucoma expectation discoveries. We assess individual pixels for glaucoma pictures' RGB channels.

4 Implementation

The fundus dataset is taken from Kaggle, which is publicly available. After collecting data, preprocessing for removing noisy and duplicate data. Then data is split into training and testing. In this study, pixel thickness assessment and submodular pick-nearby interpretable model-agonistic clarification (SP-LIME) are used to expand the pixel thickness examination (PDA) of the discoveries anticipated by SNN. The SNN findings are explained and evaluated by the IML method's SP-LIME and PDA equipment (training and validation loss shown in Figs. 1 and 2). The explanation given by this work intends to show how glaucoma possibility features associate with each other (shown in Fig. 3) (Fig. 4). Additionally, XAI strategies clarify the significance of fundus images (shown in Fig. 5).

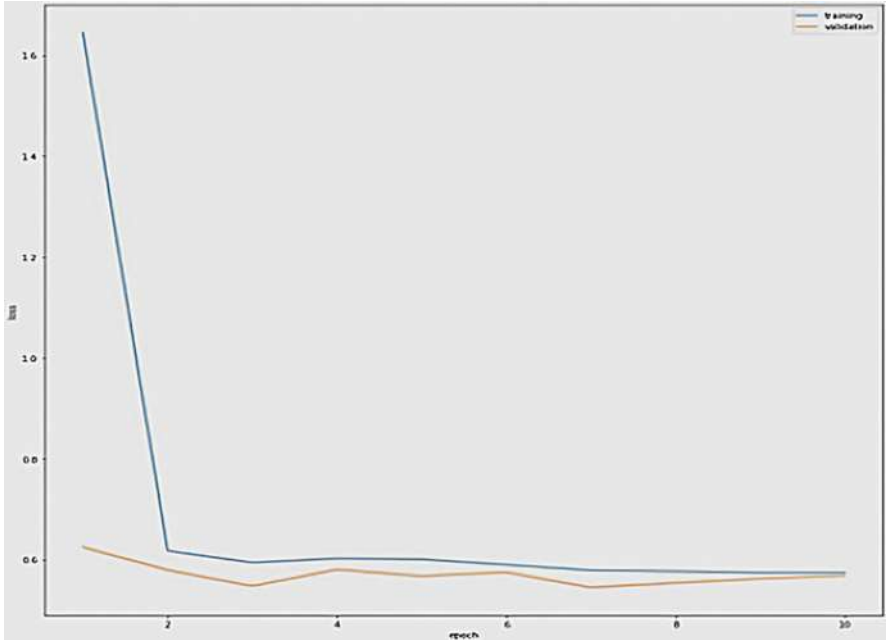


Fig. 1 The training loss of SNN

Mobile Net

Trainable params: 3,239,746
Non-trainable params: 21,888

5 Discussion

Multiple ML strategies, including diverse datasets and techniques, have been discussed for glaucoma assessment; nevertheless, these strategies are extremely challenging to implement. Fundus imaging is the preferred method for glaucoma assessment, despite the rising popularity of other forms of diagnostic imaging technology. Utilizing a novel combination of DenseNet and DarkNet, we can distinguish the difference between normal and glaucomatous fundus images. For glaucoma assumption assessment, we offer XAI and IML to appreciate risk factors in treatment arrangement, working with pictures, and managing clinical information. As opposed to standard assessment, we investigate the usage of different learning and depiction. In our review, XAI and IML ponder how to decipher these discoveries, so clinical experts and patients can perceive how well these models can foresee the names of correlation classes that are related with picture features. XAI

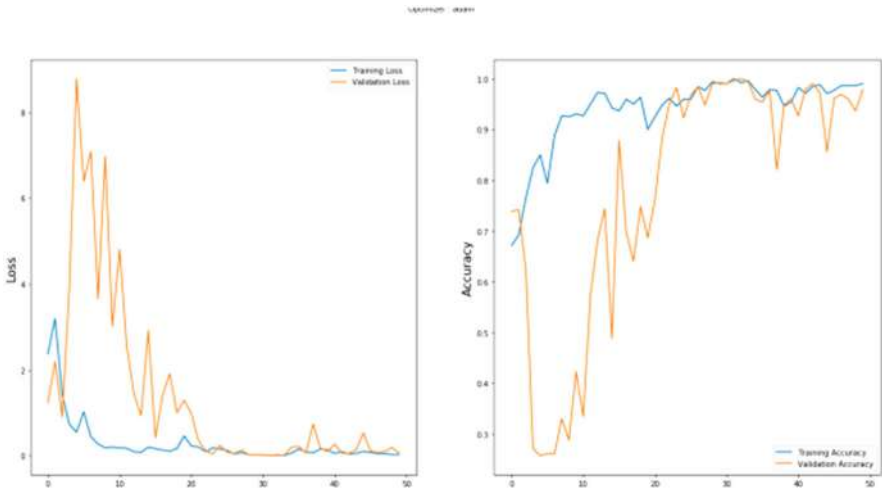


Fig. 2 Training and validation loss

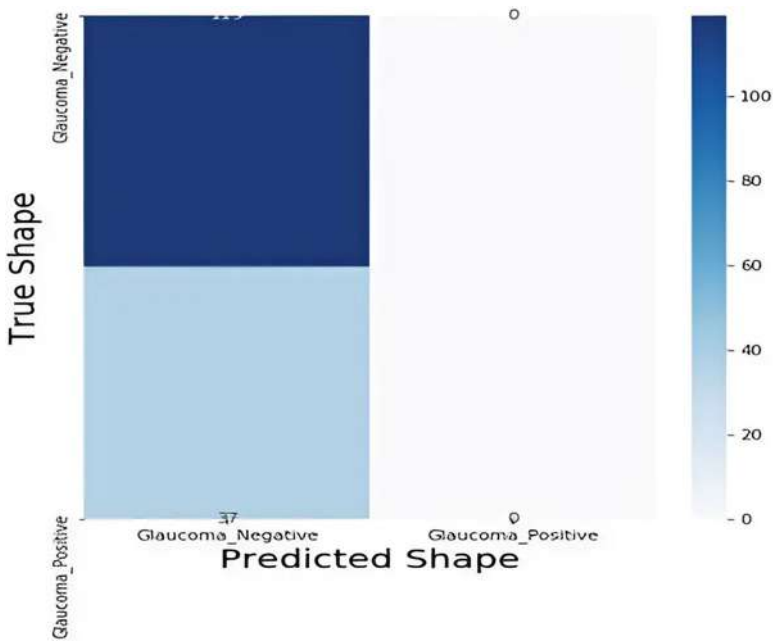


Fig. 3 The confusion matrix of proposed model

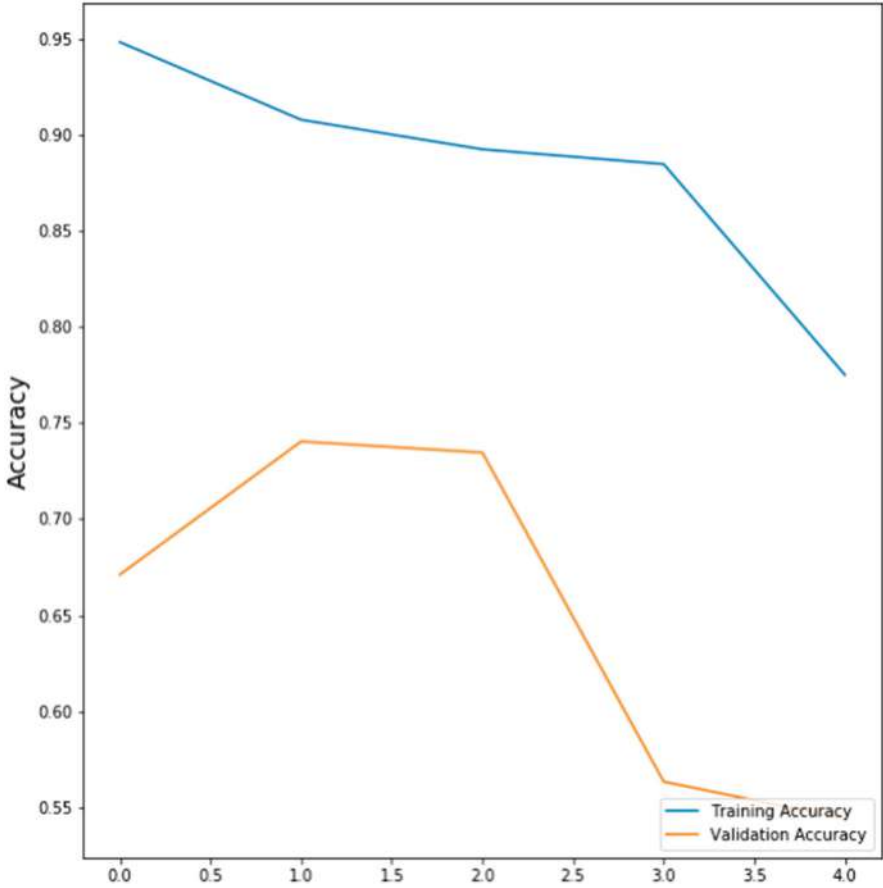


Fig. 4 Training and validation accuracy

gives substantial clarifications to the revelation of pixels in glaucoma pictures that contribute in much the same way to ML procedures.

Benefits

- (1) Doctors, clinical subject matter experts, and patients benefit from the appealing and rational conclusions produced by the XAI and IML models.
- (2) IML provides a reasonable clarification to the discoveries by utilizing the sub-modular pick local interpretable model-agonistic explanation (SP-LIME) technique.

Limitations

- (1) The incapability of the system is because of the consideration of different excess districts that do not have anything to do with glaucoma location.

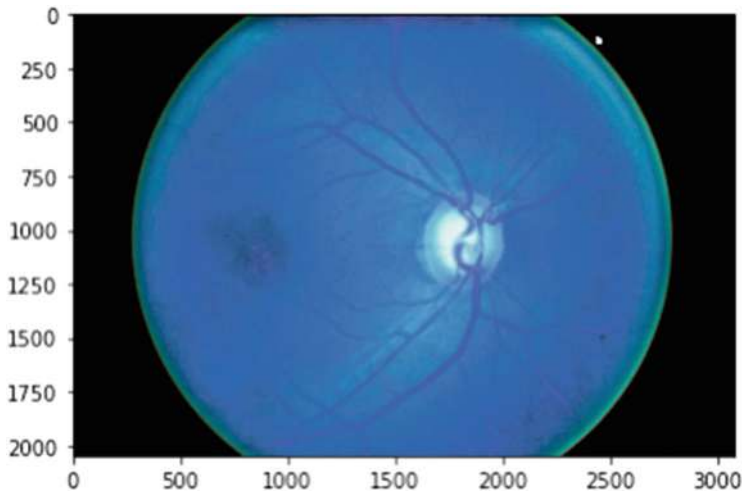


Fig. 5 Prediction result

- (2) However, patients, medical professionals, and physicians are unconvinced by their explainability and interpretability when it comes to data analysis and decision-making.

6 Conclusion

For glaucoma detection and prediction, we utilized XAI and IML, both of which enhanced the consistency and reasonableness of the models. In a similar vein, we looked into how multimodal learning and depiction could be used with clinical data and glaucoma fundus images. SP-LIME looks at how a specific choice is handled and shows the significant components that brought about a particular end. When this method is used, the decision-making process may be easily understood by both individuals and medical professionals. The glaucoma class risk variables (such as glucose level, gender, fundus pixel intensity, and so on) are described by SP-LIME. Essentially, PDA is a method that might be instructed and exhibits the particular return on initial capital investment of fundus pictures influenced by glaucoma. PDA determines the return on initial investment of fundus photographs by examining the major region responsible for glaucoma prediction. Doctors and clinical subject matter experts can trust the data provided by the XAI and IML approaches.

7 Future Scope

In future studies, multimodal gene expression data and natural inspired optimization algorithm will be combined to detect the reduce glaucoma cases. Develop a recommender system that will give accurate information to health-care professionals regarding glaucoma.

References

1. Salam AA, Khalil T, Akram MU, Jameel A, Basit I. Automated detection of glaucoma using structural and non structural features. Springerplus. 2016 Dec; 5(1):1–21.
2. Bragança CP, Torres JM, Soares CPdA, Macedo LO. Detection of Glaucoma on Fundus Images Using Deep Learning on a New Image Set Obtained with a Smartphone and Handheld Ophthalmoscope. *Healthcare*. 2022; 10(12):2345. <https://doi.org/10.3390/healthcare10122345>
3. Asaoka R., Murata A., Iwase A., Araie M. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology*. 2016; 123:1974–1980. <https://doi.org/10.1016/j.ophtha.2016.05.029>
4. Lim, T. C., Chattopadhyay, S., Acharya, U. R. (2012). A survey and comparative study on the instruments for glaucoma detection. *Medical Engineering Physics*, 34(2), 129–139.
5. A. Binder et al., “Morphological and molecular breast cancer profiling through explainable machine learning,” *Nature Mach. Intell.*, vol. 3, no. 4, pp. 355–366, Apr. 2021, <https://doi.org/10.1038/s42256-021-00303-4>.
6. Yap, et al., (2021). Verifying explainability of a deep learning tissue classifier trained on RNA-seq data. *Scientific Reports*. 11. <https://doi.org/10.1038/s41598-021-81773-9>.
7. P. M. Maloca et al., “Unraveling the deep learning gearbox in optical coherence tomography image segmentation towards explainable artificial intelligence,” *Commun. Biol.*, vol. 4, no. 1, p. 170, Dec. 2021, <https://doi.org/10.1038/s42003-021-01697-y>.
8. Bica I, et al., From Real-World Patient Data to Individualized Treatment Effects Using Machine Learning: Current and Future Methods to Address Underlying Challenges. *Clin Pharmacol Ther*. 2021 Jan; 109(1):87-100. <https://doi.org/10.1002/cpt.1907>. Epub 2020 Jun 28. PMID: 32449163.
9. A. Moncada-Torres, M. C. van Maaren, M. P. Hendriks, S. Siesling, and G. Geleijnse, “Explainable machine learning can outperform cox regression predictions and provide insights in breast cancer survival,” *Sci. Rep.*, vol. 11, no. 1, pp. 1–13, Dec. 2021, doi: <https://doi.org/10.1038/s41598-021-86327-7>.
10. Linardatos P, et al. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*. 2021; 23(1):18. doi:<https://doi.org/10.3390/e23010018>
11. Watson, D.S. Interpretable machine learning for genomics. *Hum Genet* 141, 1499–1513 (2022). doi:<https://doi.org/10.1007/s00439-021-02387-9>
12. Carrieri AP, et al. Explainable AI reveals changes in skin microbiome composition linked to phenotypic differences. *Sci Rep*. 2021 Feb 25; 11(1):4565. Doi: <https://doi.org/10.1038/s41598-021-83922-6>. PMID: 33633172; PMCID: PMC7907326.
13. G. Howard, Mobilenets: Open-source models for efficient on-device vision, Jun. 2017. [Online]. Available: <https://ai.googleblog.com/2017/06/mobilenets-opensource-models-for.html>
14. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” Apr. 17, 2017. arXiv: <http://arxiv.org/abs/1704.04861v1>

15. Ribeiro MT, Singh S, Guestrin C. “Why Should I Trust You?”: explaining the predictions of any classifier. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA: ACM. (2016), p. 1135–44. <https://doi.org/10.1145/2939672.2939778>.
16. R. LaLonde, D. Torigian and U. Bagci, “Encoding visual attributes in capsules for explainable medical diagnoses,” in Medical Image Computing and Computer Assisted Intervention—MICCAI 2020.
17. M. Oussalah AI explainability. A bridge between machine vision and Natural Language Processing International Conference on Pattern Recognition (2021).

Attainment Expedients of Markovian Heterogeneous Water Heaters in Queueing Models by Matrix Geometry Method



P. Syamala, R. Ramesh, and M. Seenivasan

Abstract This chapter estimates the attainment expedients of two Markovian heterogeneous water heaters in queueing models by matrix geometry method. We consider the heat conduction as a Queue formation. The two water heaters are considered as service providers (servers). These two water heaters have different rates of heating speed. The speed rate of first water heater is μ_1 and the second is μ_2 . In particularly we consider $\mu_1 > \mu_2$. By our proposed matrix geometry method, we catch mainly the retention rate of customers and more performance measures.

Keywords Heat conduction · Markovian process · Heterogeneous water heaters · Matrix geometry method · Queueing probabilities

1 Introduction

Conduction dominates in various phases such as solid, liquid, and plasma. Conduction of heat is that the transmission of inner energy by minuscule impingements of particles and shifting of electrons inside the mass. The inner energy means the impinging particles included the atoms, electrons, and molecules, disposed the muddled minuscule potential and kinetic energy.

The heat waves passes voluntarily from a torrid object to a frigid object, e.g., heat is transmitted from the burning candle (Fig. 1) to a metal rod in touch with the flame. Devoid of a contrary extrinsic operating power genesis, inside an object or

P. Syamala

Department of Mathematics, AVVM Sri Pushpam College (Affiliated to Bharathidasan University, Tiruchirappalli), Thanjavur, TamilNadu, India

R. Ramesh (✉)

Department of Mathematics, Arignar Anna Government Arts College, Musiri, Savanthilingapuram, Trichy, Tamilnadu, India

M. Seenivasan

Mathematics Wing – DDE, Annamalai University, Chidambaram, Tamil Nadu, India

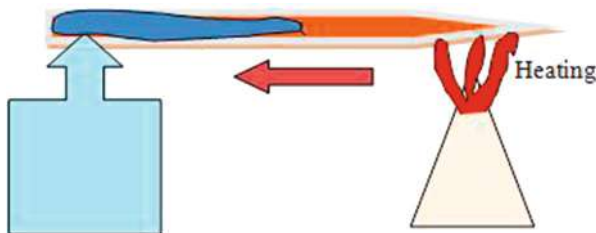


Fig. 1 Heat conduction

betwixt objects, [heat](#) variations decomposition over duration and [tepid equipoise](#) is proposed, heat became more consistent.

In particular, the water heaters work simply by converting the electrical energy into heat through the heating element to raise the temperature of the water to a particular degree.

Recently, Xiong et al. [12] gave some contributions about heat conduction. Wang et al. [11] analyzed about thermal contact resistance.

Queueing theory gets hold an efficient fragment in the present living society. At present myriad, congesting problems in front of ration, medical shops, ATM centers, reservation counters, clinics, making calls in mobile networks, and so on are faced by us. At this moment, queueing models are holding a very impressive position.

The initial concepts [3, 4, 6] of queueing portraits are most effective in many researches. The balking and renegeing concepts [1, 2, 5] play a vital role in various fields.

If two or more servers have different service rates, these are called heterogeneous servers. So many fields, especially the duty paid shops, have the queueing system with heterogeneous severers [7, 8] by the Poisson stream. Recently, Ramesh et al. gave the [9] achievements of heterogeneous servers.

More inventors introduced dissimilar techniques for finding the attainments of queues. Particularly, the matrix geometry method takes place a huge role. Recently, Seenivasan et al. presented [10] some results for heterogeneous sever. Moreover, the proposed technique is also very concrete.

2 Markovian Heterogeneous Water Heaters in Queueing Models

Consider two water heaters (servers) have the heating speed rates (service rates), with parameters μ_1 and μ_2 for heaters 1 and 2 are followed by the exponential distribution. Here, we take $\mu_1 > \mu_2$. The persons are arriving in a Markovian queueing pattern with mean arrival rate λ to get the heat water in a solo queue by the Poisson stream. Suppose an arrived person observes all heaters are engaged,

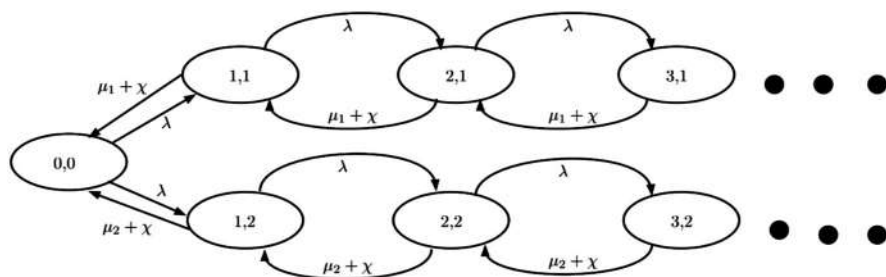


Fig. 2 Transition diagram

the person halts in a dash in the sequence of advent. The waiting person employs the heater that falling blank first. If one water heater is unoccupied, then the person employs that particular heater, and if all heaters are unoccupied, the person selects the faster one. Let the system's size be finite (N). Every person while occurring in the line will stay a particular duration of flow, and the impatient customer leaves the system is called reneging, which follows an exponential stream with parameter χ , otherwise he retains in the same line to get retention. The rate transition diagram is shown in Fig. 2.

Let a Markovian process be $\{(s(t), r(t)) : t \geq 0\}$, here $s(t)$ and $r(t)$ represent the processing state of process at t , respectively.

$r(t) = 1$, first heater state,

$r(t) = 2$, second heater state, and,

$s(t)$ = number of patrons in the queue.

The state space Ω with QBD process is defined as

$$\Omega = (0, 0) \cup (i, j); i = 1, 2 \& j = 1, 2, \dots, n \geq 1$$

Let the generator matrix in the infinite form be

$$Q = \begin{pmatrix} B_{00} & B_{01} & 0 & \dots & \dots & \dots & \dots & \dots \\ B_{10} & A_1 & B_1 & \dots & \dots & \dots & \dots & \dots \\ 0 & C_1 & A_1 & B_1 & \dots & \dots & \dots & \dots \\ 0 & 0 & C_1 & A_1 & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & C_1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

where

$$\begin{aligned}
 B_{00} &= (-2\lambda); B_{01} = (\lambda \ \lambda); \quad B_{10} = \begin{pmatrix} \mu_1 + \chi \\ \mu_2 + \chi \end{pmatrix}; \\
 A_1 &= \begin{pmatrix} -(\lambda + \chi + \mu_1) & 0 \\ 0 & -(\lambda + \chi + \mu_2) \end{pmatrix}; \quad B_1 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}; \\
 E &= \begin{pmatrix} \mu_1 + \chi & 0 \\ 0 & \mu_2 + \chi \end{pmatrix}
 \end{aligned}$$

We define, $\pi_{ij} = \{r = i, s = j\}$; $\lim_{t \rightarrow \infty} r(t) = i$, $\lim_{t \rightarrow \infty} s(t) = j$, where j represents total customers presented in the queue and i represents the state of the server. Let the probability vector be $\Pi = (\Pi_0, \Pi_1, \Pi_2, \dots)$, where $\Pi_0 = \pi_{00}$ and $\Pi_j = (\pi_{j1}, \pi_{j2})$, $j = 1, 2, 3, 4, \dots$. By utilizing $PQ = O$, we get the static probability row matrix.

$$\Pi_0 B_{00} + \Pi_1 B_{10} = 0 \quad (1)$$

$$\Pi_0 B_{01} + \Pi_1 A_1 + \Pi_2 C_1 = 0 \quad (2)$$

$$\Pi_1 B_1 + \Pi_2 A_1 + \Pi_3 C_1 = 0 \quad (3)$$

$$\Pi_2 B_1 + \Pi_3 A_1 + \Pi_4 C_1 = 0 \quad (4)$$

$$\Pi_{i-1} B_1 + \Pi_i A_1 + \Pi_{i+1} C_1 = 0 \quad (5)$$

and

$$\Pi_i = \Pi_1 R^{i-1} \text{ for } i \geq 2 \quad (6)$$

Let us denote rate matrix be R . Using matrix equation, it is irreducible nonnegative solution.

$$\Pi_0 [R^2 C_1 + R A_1 + B_1] = 0 \quad (7)$$

Using Eqs. (1) and (2), we get

$$(\Pi_0 \ \Pi_1) \begin{pmatrix} B_{00} & B_{10} \\ B_{01} & A_1 + R C_1 \end{pmatrix} = 0 \quad (8)$$

The balancing mathematical statement is

$$\Pi_0 e + \Pi_1 (1 - R)^{-1} e = 1 \quad (9)$$

Here, “ e ” is unit matrix.
The static condition takes format $\rho < 1$, where $\rho = \frac{\lambda}{\mu_1+\mu_2}$.
Using Eq. (7) we found R matrix by iteration method and from that we found probability vectors using Eqs. (6) and (8).

3 Numerical Study

Consider a student’s hostel have only two water heaters that heat the water with different heating speed rates, with limited holding area. Students prefer to get the heat water from the faster water heater (server). The students arrive in the Markovian process queue pattern. If the heat water cannot get immediately, they will wait in a line. Suppose the waiting duration of students over runs their predicted stay, they may renege and quit the line without obtaining the heat water with probability p . The hostel allows a few of student’s retention, with a probability $q (= 1 - p)$. The students are appeared with the rate of arrival λ and serviced by the heating speed rates μ_1, μ_2 with reneging rate χ .

Now we show the effect of the parameters on system features. By the above-mentioned procedures, we interpret with some variations of λ and the remaining parameters are constant.

Interpretation 1

We take $\lambda = 0.05, \mu_1 = 0.5, \mu_2 = 0.4, =0.2, R = \begin{pmatrix} 0.0714 & 0 \\ 0 & 0.0833. \end{pmatrix}$

The row vectors $\Pi_0 = \mathbf{0.8564}$ and $\Pi_1 = (0.0618, 0.0707)$ are gotten by using the matrix R in Eq. (8) and the normalization equation $\Pi_0 e + \Pi_1(1 - R)^{-1}e = \mathbf{1}$. The remaining vectors Π_j ‘s are obtained from $\Pi_i = \Pi_1 R^{i-1}, i \geq 2$. In Table 1, the second and third columns hold the pair values of $\Pi_j, j = 0, 1, 2, \dots$. Hence, the sum of probabilities is affirmed to be $0.9999 \approx \mathbf{1}$.

Interpretation 2

If $\lambda = 0.1, \mu_1 = 0.5, \mu_2 = 0.4, =0.2, R = \begin{pmatrix} 0.1429 & 0 \\ 0 & 0.1667 \end{pmatrix}$

Table 1 Probability vector

Π_j			Total
Π_0	0.8564		0.8564
Π_1	0.0618	0.0707	0.1325
Π_2	0.0044	0.0059	0.0103
Π_3	0.0003	0.0004	0.0007
		Total	0.9999

Table 2 Probability vector

Π_j			Total
Π_0	0.7316		0.7316
Π_1	0.1046	0.1220	0.2266
Π_2	0.0149	0.0203	0.0352
Π_3	0.0021	0.0034	0.0055
Π_4	0.0003	0.0006	0.0009
Π_5	0.0000	0.0001	0.0001
		Total	0.9999

Table 3 Probability vector

Π_j			Total
Π_0	0.5133		0.5133
Π_1	0.1529	0.1818	0.3347
Π_2	0.0437	0.0606	0.1043
Π_3	0.0125	0.0202	0.0327
Π_4	0.0036	0.0067	0.0103
Π_5	0.0010	0.0022	0.0032
Π_6	0.0003	0.0007	0.0010
Π_7	0.0001	0.0002	0.0003
Π_8	0.0000	0.0001	0.0001
		Total	0.9999

The row vectors $\Pi_0 = \mathbf{0.7316}$ and $\Pi_1 = (0.1046, 0.1220)$ are gotten by using the matrix R in Eq. (8) and the normalization equation $\Pi_0 e + \Pi_1 (I - R)^{-1} e = \mathbf{1}$. The remaining vectors Π_j 's are obtained from $\Pi_i = \Pi_1 R^{i-1}$, $i \geq 2$. In Table 2, the second and third columns hold the pair values of Π_j , $j = 0, 1, 2, \dots$. Hence the sum of probabilities is affirmed to be $0.9999 \approx \mathbf{1}$.

Interpretation 3

If $\lambda = 0.2$, $\mu_1 = 0.5$, $\mu_2 = 0.4$, $\omega = 0.2$, $R = \begin{pmatrix} 0.2857 & 0 \\ 0 & 0.3333 \end{pmatrix}$

The row vectors $\Pi_0 = \mathbf{0.5133}$ and $\Pi_1 = (0.1529, 0.1818)$ are gotten by using the matrix R in Eq. (8) and the normalization equation $\Pi_0 e + \Pi_1 (I - R)^{-1} e = \mathbf{1}$. The remaining vectors Π_j 's are obtained from $\Pi_i = \Pi_1 R^{i-1}$, $i \geq 2$. In the Table 3, the second and third columns hold the pair values of Π_j , $j = 0, 1, 2, \dots$. Hence, the sum of probabilities is affirmed to be $0.9999 \approx \mathbf{1}$.

Interpretation 4

If $\lambda = 0.3$, $\mu_1 = 0.5$, $\mu_2 = 0.4$, $\omega = 0.2$, $R = \begin{pmatrix} 0.2857 & 0 \\ 0 & 0.3333 \end{pmatrix}$

The row vectors $\Pi_0 = \mathbf{0.3658}$ and $\Pi_1 = (0.1670, 0.1709)$ are gotten by using the matrix R in Eq. (8) and the normalization equation $\Pi_0 e + \Pi_1 (I - R)^{-1} e = \mathbf{1}$. The remaining vectors Π_j 's are obtained from $\Pi_i = \Pi_1 R^{i-1}$, $i \geq 2$. In Table 4, the second and third columns hold the pair values of Π_j , $j = 0, 1, 2, \dots$. Hence, the sum of probabilities is affirmed to be $0.9995 \approx \mathbf{1}$.

Table 4 Probability vector

Π_j			Total
Π_0	0.3658		0.3658
Π_1	0.1670	0.1709	0.3379
Π_2	0.0716	0.0855	0.1571
Π_3	0.0307	0.0427	0.0734
Π_4	0.0131	0.0214	0.0345
Π_5	0.0056	0.0107	0.0163
Π_6	0.0024	0.0053	0.0077
Π_7	0.0010	0.0027	0.0037
Π_8	0.0004	0.0013	0.0017
Π_9	0.0002	0.0007	0.0009
Π_{10}	0.0001	0.0003	0.0004
Π_{11}	0.0000	0.0001	0.0001
		Total	0.9995

Attainment Appraises

With respect to the Steady-state probabilities

Probability that the system is idle:

$$P(I) = \Pi_0 \quad (10)$$

Probability that the system is busy:

$$P(B) = 1 - \Pi_0 \quad (11)$$

Probability that the students got the service (rate of retention) from heater 1:

$$P(S_1) = \sum_{j=0}^{\infty} j\pi_{j1} \quad (12)$$

Probability that the students got the service (rate of retention) from heater 2:

$$P(S_2) = \sum_{j=1}^{\infty} j\pi_{j2} \quad (13)$$

Expectation of students in the system:

$$E(N) = P(B) + P(S_1) + P(S_2) \quad (14)$$

Aggregate accomplishment secured is:

$$\delta = \mu_1 \sum_{j=0}^{\infty} j\pi_{j1} + \mu_2 \sum_{j=1}^{\infty} j\pi_{j2} \quad (15)$$

Λ	With the effect of λ			
	0.05	0.1	0.2	0.3
P(I)	0.8564	0.7316	0.5133	0.3658
P(B)	0.1436	0.2684	0.4867	0.6342
P(S ₁)	0.0715	0.1419	0.2997	0.5101
P(S ₂)	0.0837	0.1757	0.4078	0.6806
E(N)	0.2988	0.5860	1.1942	1.8249
δ	0.0692	0.1412	0.3130	0.5273

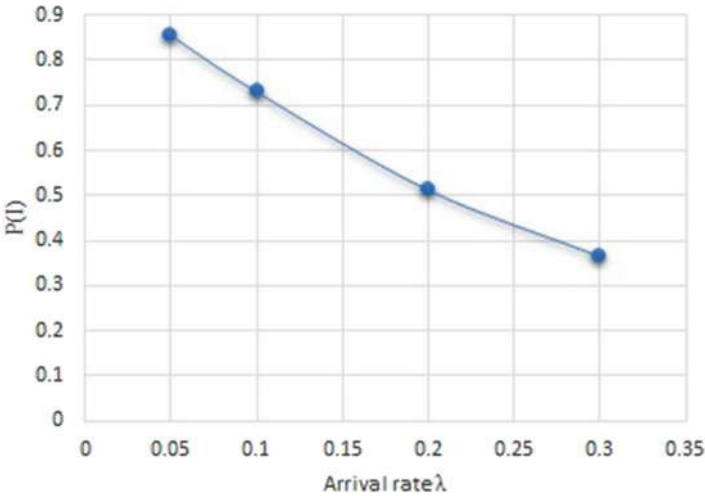


Fig. 3 Arrival versus P(I)

Out of the above figures, we acquired few attainment appraises for the sake of λ , which are the probabilities of that the system is idle, system is busy, students got the service from heater 1, students got the service from heater 2, expectation of students throughout the system, and aggregate accomplishment secured all at once. Figure 3 shows that when the arrival rate increases, the probability of the system being idle decreases. Figures 4, 5, 6, 7, and 8 shows when the arrival rate increases, the probability of the system is busy, students got the service from heater 1, students got the service from heater 2, expectation of students throughout the system, and aggregate accomplishment secured increases.

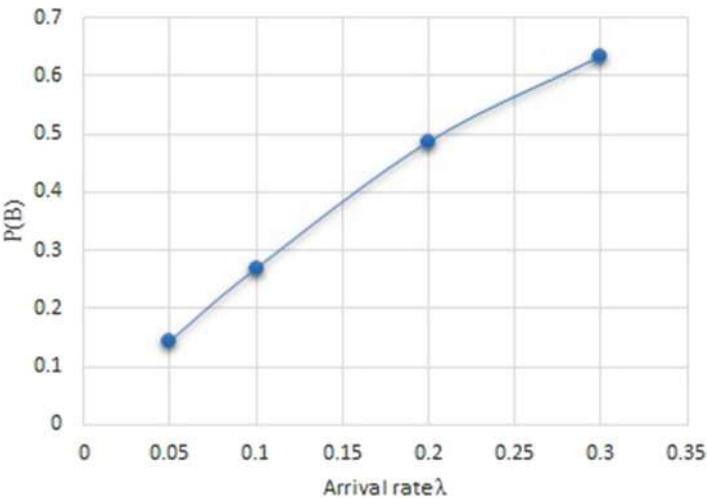


Fig. 4 Arrival versus $P(B)$

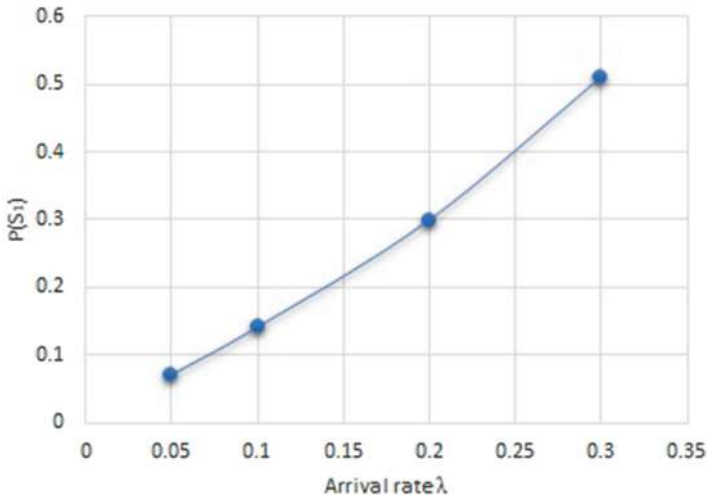


Fig. 5 Arrival versus $P(S_1)$

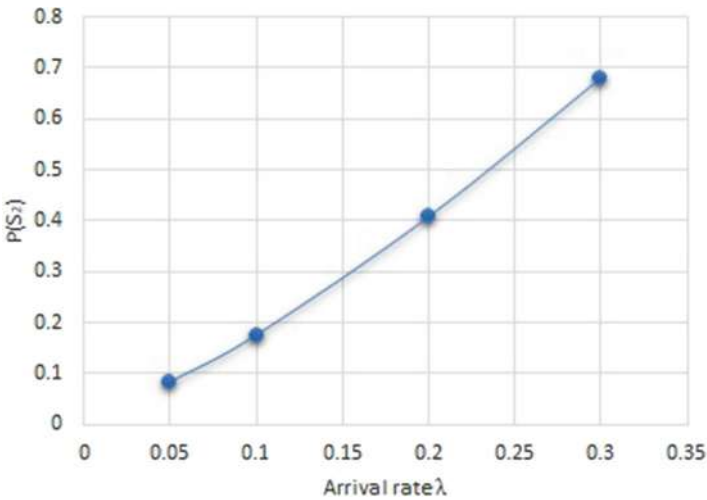


Fig. 6 Arrival versus $P(S_2)$

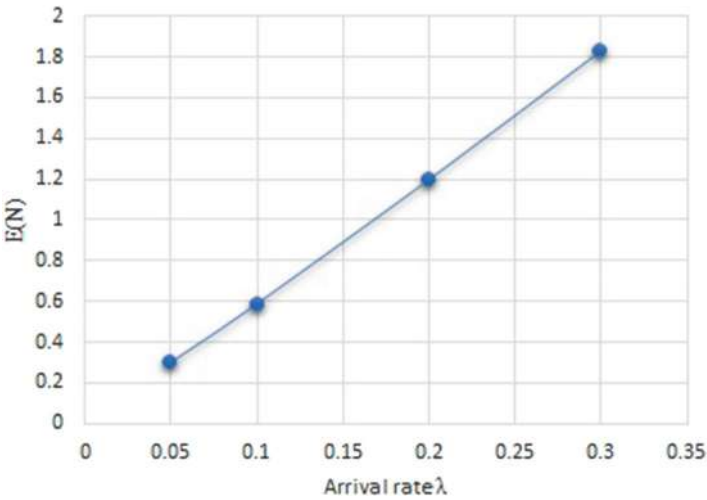


Fig. 7 Arrival versus $E(N)$

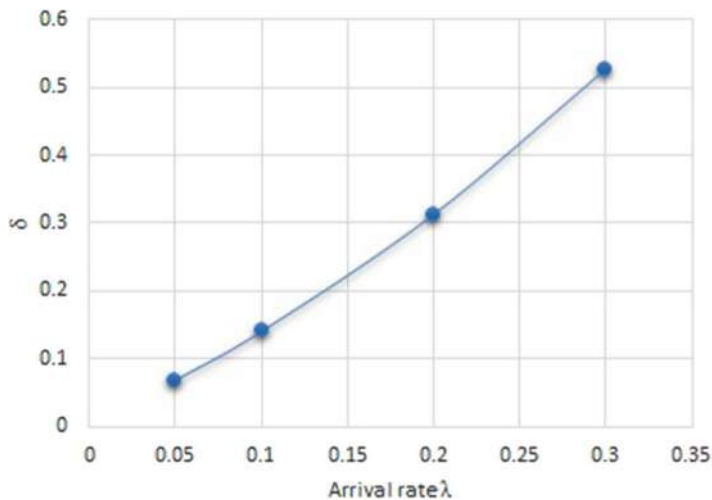


Fig. 8 Arrival versus δ

4 Output

This chapter granted an oval gambit for finding the attainment expedients of two Markovian heterogeneous water heaters in queueing models by matrix geometric method. By this method, we got very precise appraises. So, this technique is ample superior than the other. We can wield this procedure in numerous queues structures, which is also very appropriate and essential for upcoming inventions.


References

1. Ancker Jr., C. J., and Gafarian, A. V., "Some Queueing Problems with balking and Reneging", *Operations Research*, 11 (1963) 88–100.
2. Ancker, Jr., C. J., and Gafarian, A. V., "Some Queueing Problems with Balking and Reneging II.", *Operations Research*, 11 (1963) 928–937.
3. Bose S, "An Introduction to Queueing systems", Kluvar Academic / Plenum Publishers, New Yark – 2008.
4. Cooper R, "Introduction to Queueing Theory", 3rd Edition, CEE Press, Washington, 1990.
5. Haight F.A., "Queueing with Reneging", *Metrika*, 2 (1959) 186–197.
6. Janos S, "Basic Queueing Theory", Globe Edit Publishers, Omni scriptum GMBH, Germany 2016.
7. Kumar R, and Sharma S, "Transient and Steady State Behavior of a Two-Heterogeneous servers Queueing System with Balking and Retention of Reneging Customers", *Performance prediction and analytics of fuzzy, reliability and queueing models*, 2018, August, 251–264.
8. Kumar R and Sharma S, "Transient solution of a Two-Heterogeneous servers Queueing System with Retention of Reneging Customers", *Bulletin of Malaysian Mathematical Sciences Society*, 2019, January vol 42, no 1223–240.

9. Ramesh R and Srinivasan R, "Performance Measures of Two Heterogeneous Servers Queueing Models Under Trisectional Fuzzy Trapezoidal Approach", *Malaya Journal of Matematik*, Vol. S, No. 1, 392–396, 2020.
10. Seenivasan M et al. "Retrial queueing model with two heterogeneous server using matrix geometry method", *Materials Science and Engineering*, 1070 (2021).
11. Wang et al. "An Experimental Study on Heat Conduction and Thermal Contact Resistance for the AlN Flake", *Advances in Materials Science and engineering*, Volume 2013, Article ID 352173, 7 pages <http://dx.doi.org/10.1155/2013/352173> (2013).
12. Xiong, K., Yan, Z., Xie, Y. *et al.* "Regulating heat conduction of complex networks by distributed nodes masses". *Sci Rep* 11, 5501 (2021).

Identification of Medicinal Plants Using Inception V3 Model



Manoj Kumar Mahto , Vinjamoori Manaswini, D. Akshara, and Indhirala Jayasree

Abstract Medicinal plants have been utilized for centuries, and their importance in the healthcare industry is growing. Yet, manually classifying and identifying medicinal plants may be time-consuming and blunder. To label this issue, this project proposes the development of a vision-based system for accurate identification and classification of various medicinal plant species using the Inception V3 model. The system uses deep convolutional neural networks to learn and extract relevant features and patterns that distinguish different species from the large dataset of images of various medicinal plants. To enhance the system's accuracy and robustness, several image preprocessing and feature extraction techniques are employed. The effectiveness of the proposed system will be assessed using a variety of metrics, such as recall, accuracy, precision, and F1-score. Our proposed system achieves good accuracy of 93.3%. The system's potential applications include identifying medicinal plants in natural environments, tracking plant growth, and monitoring plant diseases. The proposed vision-based system has the potential to provide an accurate and efficient solution for identifying and classifying medicinal plants, contributing to the fields of botany, agriculture, and medicine.

Keywords Vision · classification · Inception V3 · Deep convolutional neural networks · Medicinal plants

1 Introduction

India is a developing country with a diverse ecosystem that plays a critical role in sustaining the natural resources and increasing ecosystem productivity. In underdeveloped nations, more than 60% of the people uses the traditional botanicals to alleviate a range of illnesses [1, 2]. Nevertheless, due to their toxic and dose-

M. K. Mahto (✉) · V. Manaswini · D. Akshara · I. Jayasree
Vignan Institute of Technology and Science, Hyderabad, Telangana, India

dependent nature, medicinal plants posed considerable concerns throughout the twentieth century. As a result, there is rising interest in using medicinal plants as a complementary method of treatment for a variety of diseases. Yet, classifying and identifying medicinal plants may be difficult, particularly for those who are not botany specialists [3].

To address this issue, computer vision technologies have been developed to automate the process of identifying and classifying medicinal plants. In this research project, we present a vision-based system for identifying and categorizing therapeutic plants using the Inception V3 model [4, 5]. The proposed system uses a dataset of images of various medicinal plants, which were collected from different regions around the world. The Inception V3 model was implemented to preprocess the photos and extract the features. A classification model was then trained using the extracted features to accurately identify and classify medicinal plants.

Eighty percent of the world's population receives their main medical treatment from traditional herbal medicine according to the World Health Organization (WHO). Twenty-five percent of medications in the United States and 80% of drugs in China and India come from plants [6]. Yet, only a small number of professionals, rural residents, and local people are knowledgeable about indigenous plants, and it is difficult to get the supporting documentation. To encourage a healthy herbal lifestyle and stop the extinction of prospective plants, it is essential to spread knowledge about herbs to both the general population and researchers [7].

The proposed system has several advantages over traditional methods of medicinal plant identification, including the capacity to quickly handle massive volumes of data accurately. The system can also be used by individuals who are not experts in botany, allowing them to identify and classify medicinal plants with ease. In conclusion, this research project's proposed system has the potential to enhance the field of medicinal plant research and development and contribute to the sustainable use of natural resources.

2 Literature Survey

Venation-based plant leaves classification using Google Net and VGG 8 [8] proposed a Google Net CNN model for classifying plants using leaf venation. They trained and tested the model using Dleaf, Flavia, and Leaf1 datasets with both CNN and SVM classifiers. They compared the performance of fine-tuned GoogLeNet, finetuned VGG-16, and Dleaf CNN models [9]. *Findings:* The fine-tuned GoogLeNet CNN model outperformed the other models, achieving a five-fold cross-validation accuracy of 99.2% on the Leaf1 dataset using the SVM classifier. Deep learning for plant species classification using leaf vein morphometric [10] proposed a new automated plant species identification system named D-Leaf. The system is based on deep learning, which is known for its robustness in feature extraction from images. The researchers used three different CNN models, including pretrained Alex Net, fine-tuned Alex Net, and D-Leaf to extract features from a

set of preprocessed leaf images [11]. The extracted features were then classified using five different machine learning techniques, including support vector machine (SVM), artificial neural network (ANN), k-nearest neighbor (k-NN), Naïve Bayes (NB), and CNN. The benchmark method used in this study was the conventional morphometric method, which computed morphological measurements based on Sobel segmented veins [12].

Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost proposed by Smith [13] plant species recognition model is based on morphological features extracted from leaves' images using the support vector machine (SVM) with adaptive boosting technique. The study uses various morphological features, including centroid, major axis length, minor axis length, solidity, perimeter, and orientation, extracted from digital images of different categories of leaves. Transfer learning is also used in the feature extraction process, as suggested by previous studies. The study employs various classifiers, such as the kNN, decision trees, and multilayer perceptron (with and without AdaBoost), on the opensource dataset, FLAVIA [14]. The classification of leaves was accomplished by constructing a CNN architecture with two FCLs in a five-layer network, using the MalayaKew Leaf dataset [15]. This approach demonstrated a high accuracy of 97.70%, surpassing the conventional method [16], and provided a more effective representation of leaf features. As the CNN architecture enables the automatic extraction of local and global features of an input image, it has proven to exhibit remarkable performance in numerous machine learning and image processing tasks. In the context of identifying plant species through vein features, a three-layer CNN model was proposed and tested on the INTA custom dataset, containing images of three distinct leaf species. The results showed an acceptable accuracy of 92% [17]. Due to the absence of standardized images in the dataset, the accuracy of classification has been adversely affected. However, the AlexNet model coupled with an SVM classifier achieved an accuracy of 96.76% when tested on a dataset of 2400 images, representing 40 species of Indian herbs. Each class in the dataset consisted of 60 images, including both front and back views of each leaf face, and all sample images were manually edited using the GIMP editor. Notably, details regarding model evaluation time (both for training and testing), real-time testing, and automation of input image editing were not provided [18].

3 Methodology

In the existing system, we are using machine learning techniques. Machine learning techniques on images take a lot of time to train the model. The accuracy obtained is also not satisfactory.

- The efficiency of machine learning algorithms is not suitable for larger datasets.
- The project proposes identifying and classifying medicinal plants using the Inception V3 model.



Fig. 1 Sample leaves of medicinal plant

3.1 Dataset

Proper identification of medicinal plants is essential for agronomists, ayurvedic medicinal practitioners, and for ayurvedic medicines industry. Even though many plant leaf databases are available publicly, no specific standardized database is available for Indian ayurvedic plant species. We applied a publicly available annotated database of Indian plant leaf images named as MepcoTropicLeaf. Presently, this study used 50,000 leaves and distinguished specific types of leaf species, as seen in selected samples in Fig. 1. In this study, the dataset has been updated with 50 leaf species of Indian ayurvedic plants, 25 spinach species, and another 250 leaf species from all over the world. These images are taken using mobile cameras and most of the leaves are found in the foothills of Western Ghats in Tamil Nadu, India [4].

3.2 *Used Tools*

3.2.1 **Tensor Flow**

Google created the open-source machine learning library known as Tensor Flow. It offers a platform for creating and training deep neural network models, among other machine learning models.

3.2.2 **Keras**

Python-based Keras is an API for high-level neural networks. It features support for both convolutional and recurrent neural networks and enables quick experimentation with deep neural networks. Tensor Flow, Theano, or CNTK may all be used as the base for Keras (Microsoft Cognitive Toolkit).

3.2.3 **Inception V3 Model**

Google researchers unveiled Inception V3, a deep convolutional neural network structure built for image identification applications in 2015. It is a development of the original Inception models and was made to be more accurate and computationally efficient.

The Inception V3 architecture includes several features that help it achieve high accuracy on image classification tasks, including the following:

Inception modules

To extract characteristics from input photos, the model employs many convolutional layers with various filter sizes and pooling procedures.

Factorization into smaller convolutions

Inception V3 uses factorization into smaller convolutions to reduce the number of parameters required for convolutional layers, which in turn reduces the computational cost.

Auxiliary classifiers

The architecture includes auxiliary classifiers at intermediate layers to encourage the model to learn more robust features and reduce overfitting.

Batch normalization

The model uses batch normalization to reduce internal covariate shift and stabilize the training process.

Pooling of global averages

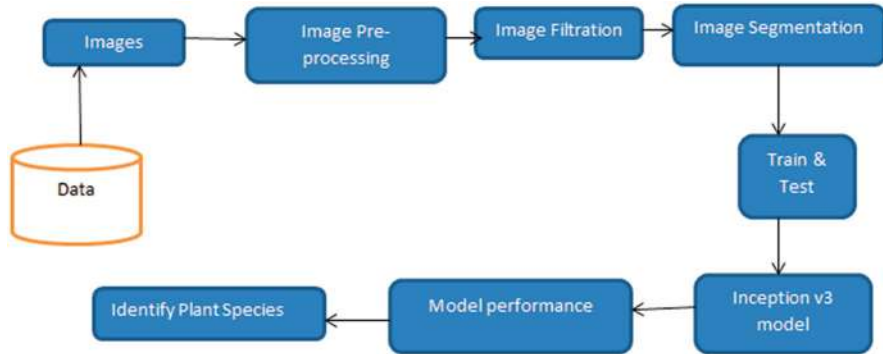


Fig. 2 System architecture

The final layer of the network performs global average pooling, which helps to reduce overfitting and makes the model more robust to changes in input image size.

For a variety of picture classification challenges, the Inception V3 model has, overall, attained state-of-the-art performance in a system architecture as shown in Fig. 2. It has been extensively utilized in several computer vision applications, including picture segmentation, object identification, and captioning.

3.3 Evaluation Metrics

Many statistical measures are employed to assess the effectiveness of the suggested architecture in addition to our unique metric, known as the corona score.

3.3.1 Accuracy

A measurement of the method's ability to identify the right expected cases is accuracy.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

True Positive

The quantity of accurately anticipated positive cases is the same as the True Positive.

False Positive

The number of incorrectly forecasted positive instances is equivalent to how many false positives are there.

True Negative

The quantity of correctly anticipated negative situations is the same as the true negative.

False Negative

The number of incorrectly anticipated negative instances is equal to how many false negatives were there.

3.3.2 Recall

The method's sensitivity is recall.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

3.3.3 Precision

The ratio of unnecessary positive instances to all confirmed samples is known as precision.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

3.3.4 Specificity

Specificity is defined as the proportion of properly expected negatives to negative observations.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

3.3.5 F1-Score

The effectiveness of detection is gauged by the F1-Score.

$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \tag{5}$$

$$= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4 Result Analysis

4.1 Loss Graph for Inception Model

The loss graph during training is shown in Fig. 3. The loss was defined as the cross entropy between predicted value and the actual one. With the iteration of training, the loss dramatically decreases and finally tends to zero. Horizontal axis: iteration times. Vertical axis: loss value.

From the plot of the loss, you can see that the model has comparable performance on both train and validation datasets (labeled test).

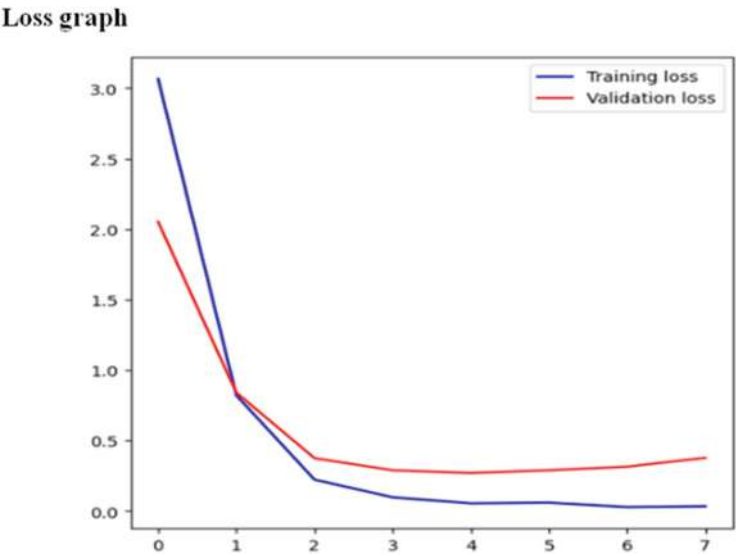


Fig. 3 Loss graph

Accuracy graph

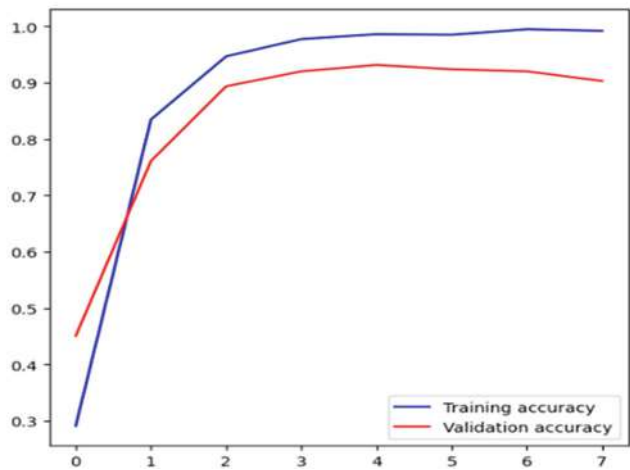


Fig. 4 Accuracy graph

Table 1 Accuracy table

Model	Accuracy
Efficient model	75%
Inception V3 model	93.3%

4.2 Accuracy Graph for Inception Model

From the plot of the accuracy shown in Fig. 4, you can see that the model could probably be trained a little more as the trend for accuracy on both datasets is still rising for the last few epochs.

4.3 Accuracy Table

From Table 1 we can understand that our proposed model gives good accuracy compared to existing efficient model. Fig. 5 shows a specific medicinal plant leaf.

4.4 Comparison of Models

Fig. 6 is the comparison graph between existing model and proposed model where the proposed model gives 93.3% accuracy.

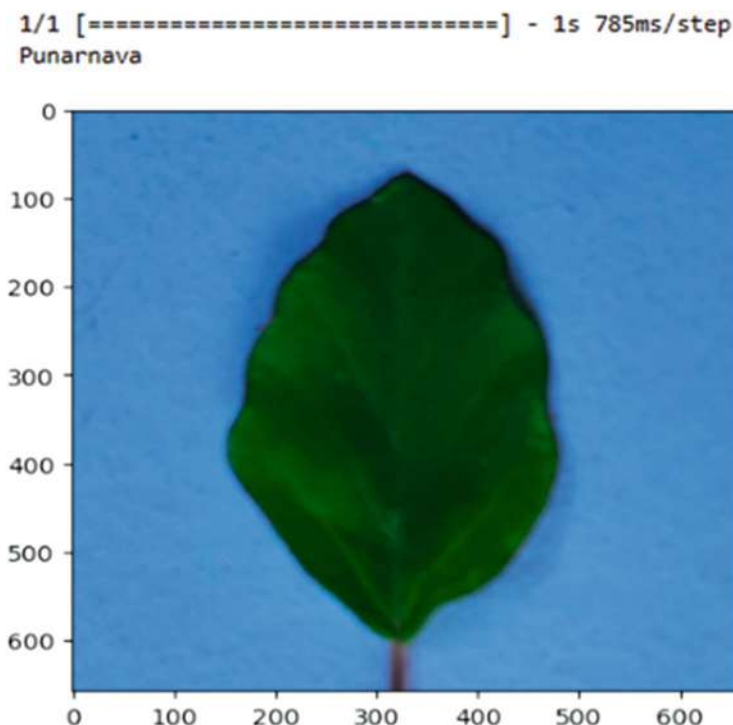


Fig. 5 Predicting medicinal plant name

5 Conclusion

The proposed computer vision-based system using the Inception V3 model achieved high accuracy of 93.83% in identifying medicinal plants from images, improving the efficiency and accuracy of plant identification process. The study emphasizes the use of computer vision and deep learning approaches for plant identification across a range of industries, including botany, agriculture, and pharmaceuticals.

Incorporation of Other Deep Learning Architectures While the Inception V3 model is a powerful architecture for image recognition, there are other deep learning models that can be used in conjunction with it to improve accuracy and performance. For example, incorporating attention mechanisms or recurrent neural networks can help to better capture the temporal and spatial features of plant growth.

Integration with Other Data Sources In addition to images, other data sources such as environmental data, soil data, and weather data can be incorporated into the system to improve accuracy and help predict plant growth patterns.

Development of a Mobile Application The system can be integrated into a mobile application that can be used by farmers, researchers, and other stakeholders to

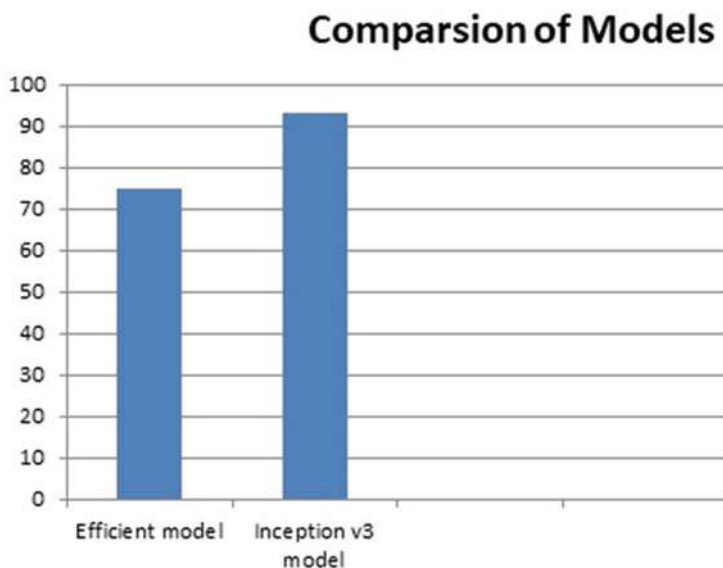


Fig. 6 Comparison of two models

monitor plant growth and health. The app can include features such as real-time plant recognition, recommendations for plant care, and an interface for uploading and analyzing images.

References

1. Bir, P., Kumar, R., & Singh, G. (2020, October). Transfer learning-based tomato leaf disease detection for mobile applications. In 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 34–39). IEEE.
2. Dyrmann, M., Karstoft, H., & Midtby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems engineering*, 151, 72–80.
3. Adedoja, A. O., Owolawi, P. A., Mapayi, T., & Tu, C. (2022). Intelligent Mobile Plant Disease Diagnostic System Using NASNet-Mobile Deep Learning. *IAENG International Journal of Computer Science*, 49(1), 216–231.
4. MepcoTropicLeaf, (2023). <https://www.kaggle.com/datasets/ahilaprem/mepco-tropic-leaf>
5. Salih, T. A. (2020). Deep learning convolution neural network to detect and classify tomato plant leaf diseases. *Open Access Library Journal*, 7(05), 1.
6. Cheng, Q., Zhao, H., Wang, C., & Du, H. (2018, December). An android application for plant identification. In 2018 IEEE 4th information technology and mechatronics engineering conference (ITOEC) (pp. 60–64). IEEE.
7. Jasitha, P., Dileep, M. R., & Divya, M. (2019, May). Venation based plant leaves classification using GoogLeNet and VGG. In 2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT) (pp. 715–719). IEEE.

8. Kaya, A., Keceli, A. S., Catal, C., Yalic, H. Y., Temucin, H., & Tekinerdogan, B. (2019). Analysis of transfer learning for deep neural network based plant classification models. *Computers and electronics in agriculture*, 158, 20–29.
9. Bodhwani, V., Acharjya, D. P., & Bodhwani, U. (2019). Deep residual networks for plant identification. *Procedia Computer Science*, 1.
10. Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition*, 71, 1–13.
11. Mukti, I. Z., & Biswas, D. (2019, December). Transfer learning based plant diseases detection using ResNet50. In 2019 4th International conference on electrical information and communication technology (EICT) (pp. 1–6). IEEE.
12. Wäldchen, J., & Mäder, P. (2018). Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering*, 25, 507–543.
13. Smith, L. N. (2017, March). Cyclical learning rates for training neural networks. In 2017 IEEE winter conference on applications of computer vision (WACV) (pp. 464–472). IEEE.
14. Liu, C., Cao, Y., Luo, Y., Chen, G., Vokkarane, V., & Ma, Y. (2016). Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment. In *Inclusive Smart Cities and Digital Health: 14th International Conference on Smart Homes and Health Telematics, ICOST 2016, Wuhan, China, May 25–27, 2016. Proceedings 14* (pp. 37–48). Springer International Publishing.
15. Petrovska, B. B. (2012). Historical review of medicinal plants' usage. *Pharmacognosy reviews*, 6(11), 1.
16. Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015, September). Deep-plant: Plant identification with convolutional neural networks. In 2015 IEEE international conference on image processing (ICIP) (pp. 452–456). IEEE.
17. Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and electronics in agriculture*, 127, 418–424..
18. Dileep, M. R., & Pournami, P. N. (2019, October). AyurLeaf: a deep learning approach for classification of medicinal plants. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)* (pp. 321–325). IEEE.

Smart Gardening Using Internet of Things



M. G. Mahesh, S. Afsana, K. Divya, P. Menaka, G. Prasad,
and M. A. Suhanulla Khan

Abstract Agriculture has an important role in India's economy. Consequently, the agricultural sector is an essential component of our nation. Agriculture relies on monsoon. In agriculture, the plants get their required amount of water according to the composition of the soil. In the field of agriculture, it is essential to first determine the level of moisture present in the soil and then learn as much as possible about the fertility of the ground. Irrigation refers to any technique that allows water to trickle slowly down to the roots of a plant; this is accomplished by using a solenoid valve. The intelligent irrigation system helps to use less water and time, is ready to be used, uses energy effectively, and is cost-effective. In this research, we make use of many sensors, some of which are light sensors, infrared sensors, and moisture sensors for the soil. Email and SMS (short messaging service) notifications are provided to the user if the plants demand more water.

Keywords Embedded system · NodeMCU · Wi-Fi · Sensor · Fertility

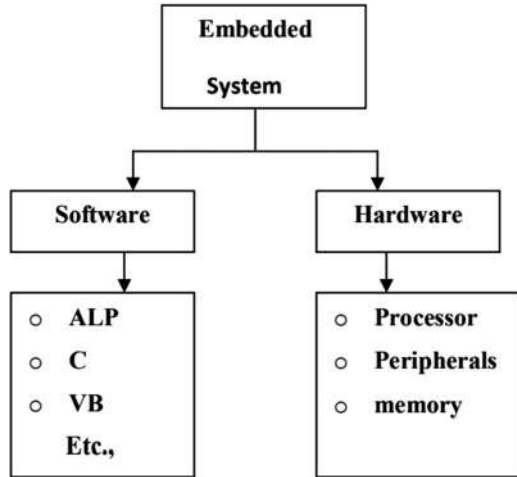
1 Introduction

1.1 Introduction to Embedded Systems (Fig. 1)

An embedded system is a system that is going to accomplish a preset stated job, and it is even characterized as a mix of both software and hardware. An embedded system is a system that is going to execute the work [1]. Devices that are used to control, monitor, or aid in the functioning of equipment, machinery, or plant are referred to as embedded systems. This is a general-purpose definition of embedded systems. The term “embedded” refers to the fact that they are an essential

M. G. Mahesh (✉) · S. Afsana · K. Divya · P. Menaka · G. Prasad · M. A. S. Khan
Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering
and Technology, Chittoor, India

Fig. 1 Block diagram of embedded system



component of the overall system. On the other hand, a general-purpose computer may be employed to manage the functioning of a vast and complicated industrial facility, in which case its presence would be readily apparent [2].

Computers and other types of microprocessors are included in every embedded system. On the other hand, in comparison to a home computer, some of these computers are really simple systems.

The embedded systems that are the least complex are those that are only capable of carrying out a single function or collection of operations in order to fulfil a single intended goal. In more complex systems, the operation of the embedded system is determined by an application program. This application program is what makes it possible for the embedded system to be used for a particular purpose within a particular application [3]. Because it is able to run different programs, a single embedded system can serve a wide variety of functions, provided that the appropriate program is loaded. If a microprocessor is created in such a manner that application software for a specific purpose may be added to the basic software in a second process, after which it is not feasible to make any more modifications to the program, then the microprocessor is said to be constructed in such a way that it is intended for a particular purpose. Firmware is a term that is occasionally used to refer to the application software that is included on such CPUs [4].

The simplest devices have a single microprocessor, also known as a "chip," which may or may not be bundled with other chips as part of a hybrid system or application-specific integrated circuit (ASIC). Its output is sent to a switch or activator, which (for example) may initiate or halt the operation of a machine or, by opening a valve, may regulate the flow of gasoline to an engine. Its input comes from a detector or sensor, and its output is sent to a switch or activator. Due to the fact that the embedded system is a mix of software and hardware [5].

The software is concerned with the languages such as ALP, C, and VB, among others, whereas the hardware is concerned with processors, peripherals, and memory [6].

Memory: It may be used to store information or addresses.

Peripherals: These are the gadgets from the outside that are linked.

Processor: It is a piece of IC that is used in the execution of a certain job.

2 Literature Survey

Nowadays, automation has taken over the globe. It is a method that involves making use of computers or mobile phones in order to monitor and regulate the fundamental aspects of one's day-to-day existence. The routine of using automation for relatively straightforward tasks is going to be beneficial to the quality of life that we lead. We are able to create sensors that are strong in automation by using the notion of IOT to make them connect with each other. The fact that this prototype is both cost-effective and risk-free is the most significant thing about it. People, when they try to make plantings and set up their own garden, are careful in the maintenance at only the beginning stages of their efforts. The lack of care that has been performed on the plants causes their destruction with the passing of each day. People will be able to automatically monitor the parameters with the aid of this prototype, which will also guarantee that the garden is maintained. It is an important component that provides a beneficial environment for plant growth [7]. The Internet of Things (IoT) offers answers to a variety of issues and makes it possible for items to be sensed or controlled remotely through network architecture.

Typically, it is composed of a microcontroller at the heart of the system to which other components are attached. The NodeMCU serves as a hub in the smart garden, and a variety of sensors, including moisture and humidity sensors, temperature sensors, and ultrasonic sensors, are attached to it. The ultrasonic sensor was attached to a water tank and provided feedback on the amount of water currently contained inside the tank. Other sensors are attached to their respective placements, and these sensors relay the data that they collect to NodeMCU, which has a Wi-Fi technology as part of its internal architecture. The real-time readings of the sensor are updated once every second in the database known as Firebase, which is accessible over the internet [8]. The program known as android studio is used in the development of Android applications. The communication between the application and firebase will be established inside of the program itself. Therefore, the user is able to monitor the parameters regardless of their location. The amount of water that a garden needs depends on the kind of soil. Therefore, the values of the sensors are hardcoded into the software in order to facilitate automation of the process. When the user determines that the garden needs watering, a toggle switch inside the program will take care of the operation automatically. This contributes to the garden's overall care and upkeep.

3 Conclusion

NodeMCU is used to create a smart irrigation system for indoor and outdoor plants. We also linked it with the relay power module in addition to the infrared sensor, the moisture sensor, and the light sensor. Every sensor is given an appropriate threshold value, which is determined by experimental testing. The relay is under the direction of NodeMCU, which determines whether it is ON or OFF at any given moment. The relay acts as a measuring device for the solenoid valve. The method that has been proposed is both realistic and affordable, and it has the potential to be used in gardens and other forms of agriculture on a smaller scale. Control of the irrigation system may be achieved via the use of an android interface app on a mobile device. The user will get the information in the form of an email as well as a text message. Because of the use of this technology, the amount of human involvement is decreased. It requires less energy and less time to complete.

References

1. Tzu-Ching Chia "Design and implementation of the microcontroller control system for vertical garden application" Department of special design Kun Shan University, KSU Tainan, , Chun-Li Lu* Department of computer and communication, Kun Shan University, KSU Tainan, Taiwan, 2011,
2. Tang Li-fang "Application of auto control technology in water-saving Garden irrigation" Department of Computer Engineering Cangzhou Normal College Cangzhou, Hebei Province, 2012, China, E-mail
3. AravindAnil*, Aravind R Thampi*, Prathap John M* and Shanthi K J† "Project HARITHA-An Automated Irrigation System for Home Gardens" *Dept. of Electronics and Communication Sree Chitra Thirunal College of Engineering, Trivandrum, Kerala, 2012.
4. A. R. Al-Ali, Murad Qasaimeh, Mamoun Al-Mardinia, Suresh Radder and I. A. Zualkernan "ZigBee-Based Irrigation System for Home Gardens" Department of Computer Science and Engineering, American University of Sharjah, UAE P. O. Box 26666, 2015, Sharjah, 2012, UAE. Email:
5. Siwakorn Jindarat "SmartFarm Monitoring Using Raspberry Pi and Aurdino" Faculty of Information Technology King Mongkut's University of Technology North Bangkok, Thailand Pongpisitt Wuttidittachotti, North Bangkok, Bangkok, Thailand, 2015,
6. Shweta B. Saraf "IOT Based Smart Irrigation Monitoring And Controlling System" NBN, Sinhgad School of Engineering, Dhanashree H Gawali, 2017,
7. Vaishali S, Suraj S, Vignesh G, Dhivya S and Udhayakumar S. " Mobile Integrated Smart Irrigation Management and Monitoring System Using IOT", 2017.
8. Filipe Caetano "Intelligent management of urban garden irrigation" Faculty of Engineering UBI-University of Beira Interior Covilhã, Portugal RuiPitarmaa, Pedro Reisb, ESTG-School of Technology and Management IPG-Polytechnic Institute of GuardaGuarda, Portugal, 2017,

Predictive Analytics of Blood Donor Risk Assessment Using Machine Learning Methods



V. Kakulapati , Avula Sravan Reddy, S. Mani Teja, and Acha Vamshi

Abstract Donating blood is a selfless gesture with the possibility of saving lives. Several techniques exist for giving blood. Each kind contributes to meeting certain medical needs. Donating blood is completely safe. By providing clean, single-use equipment to each donor, we can ensure that no one gets an infection from donating blood. Most healthy people can give half a liter of blood (a pint) without risking their health. The body quickly replaces the fluids lost after donating blood. In 2 weeks, the body will replenish the lost red blood cells. For the safety of donors and patients, it is essential to identify and manage risks in the donation process and blood products. In this study, we create models for clumped datasets by enhancing performance significantly. The need for transfusions as a result of accidents, procedures, diseases, and other circumstances is the driving force behind this work effort. Realistic assessments of the number of blood donors allow healthcare professionals to establish strategies for motivating people to donate blood to meet needs. With the goal of predicting high-risk donors, this research attempts to assess the risk analysis of blood donors by using DL (deep learning) techniques based on a recurrent neural network (RNN) and a co-relational network to automatically identify blood donors. The model incorporates RNNs to track patients' progress over time, improving the accuracy of future projections. In order to provide more precise risk assessments and appropriate donor referrals, a probabilistic neural network (PNN) is fed data from a co-relational neural network (CRNN) that has identified the most crucial aspects of blood donation. The experimental findings demonstrate improved accuracy in assessing total risk in both donors and recipients.

Keywords Deep learning · Donors · Donation · Recipient · Disease · Machine learning · Accuracy · Correlation · Medical · Health

V. Kakulapati (✉) · A. S. Reddy · S. M. Teja · A. Vamshi
Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

1 Introduction

The availability of blood is crucial to the long-term health and financial stability of national healthcare systems. It has a high initial cost and a low rate of return [1]. Governments may help increase public blood donation awareness by guaranteeing the public's access to safe and reliable blood supplies. In Italy, giving blood is a completely voluntary and nameless gesture due to the fact that blood cannot be "addressed" for safety concerns. It is a collective effort to make the world a better place for everyone [2].

Regular donations from healthy volunteers are the main means of satisfying needs. Predictive analytics in the context of blood donor risk assessment is the foundation of this research. Hence, the requirement for blood and its transfer is an important factor to consider. Also, numerous illnesses are shown to be spreading in our daily lives. An instance of one of these is thalassemia. One who is afflicted by this illness also requires weekly blood transfusions. In this chapter, an effort has been made to classify blood-borne illnesses according to how they influence a person's decision to donate blood. The topic of predictive analytics is one of data science and machine learning that is expanding quickly. Large datasets are analyzed, and future results are predicted using machine learning and statistical algorithms. Predictive analytics can be used in the medical sector to evaluate the risk of various illnesses and disorders. Predictive analytics may be applied in this situation to evaluate the risk involved with blood donation [3].

It is the responsibility of blood banks to provide enough safe blood to the public. Donors may be classified into three broad categories: those who give of themselves voluntarily; those who are compensated for their services; and those who act only as replacements for other family members. Blood obtained from paid donors has a higher risk of transmitting infections via transfusion. The demand for cash makes people more inclined to lie about their health [4].

Blood donation, a vital component of healthcare, can save many lives. The caliber of the given blood, however, is what determines how safe blood transfusions are. Ensuring the blood is clear of infectious illnesses and other health issues that might endanger the receiver is what determines how safe blood transfusions are. Ensure the blood is clear of infectious illnesses and other health issues that might endanger the receiver. To protect the security of the blood supply, blood banks and transfusion facilities must evaluate the risk involved with blood donation. It is possible to create prediction models that evaluate the risk involved in blood donation using machine learning techniques. Large datasets may be analyzed by these models to find variables that raise the chance of contracting infectious illnesses and other medical disorders. Blood banks and transfusion facilities may detect high-risk donors using machine learning algorithms and take the necessary precautions to protect the blood supply.

In this study, the proposed model is used to investigate the application of machine learning techniques for the predictive analytics of blood donor risk assessment in this setting. We will examine the most recent research on using machine learning for

the risk assessment of blood donors and talk about the advantages and disadvantages of this strategy. We will also discuss a case study that shows how machine-learning techniques may be used to assess the risk of blood donors [5]. The significance of blood donor risk assessment: Blood transfusions is vital in treating patients with various medical conditions. However, if the donated blood is not safe, it can lead to infections and other complications, posing a significant risk to the recipient's health. Therefore, blood donor risk assessment is critical to ensuring the safety of the blood supply [6]. Traditional methods of blood donor risk assessment: Blood banks and transfusion centers use several traditional methods to assess the risk associated with blood donation, such as medical history questionnaires, physical exams, and laboratory tests. However, these methods have limitations, and there is a need for more efficient and accurate methods [7].

The potential of predictive analytics: Predictive analytics can help overcome the limitations of traditional methods and enable blood banks and transfusion centers to identify high-risk donors accurately. Machine learning methods can analyze vast amounts of data and identify patterns that may be missed by humans [8]. Methods for evaluating blood donors' safety DT (decision trees), RF (random forests), SVM (support vector machines), and NN (neural networks) are only some of the ML algorithms that may be used to evaluate donors' safety. There are benefits and drawbacks to each algorithm, and whatever algorithm a blood bank or transfusion center ultimately chooses will depend on its own requirements [9]. *The importance of data quality:* The effectiveness of machine learning models for blood donor risk assessment depends on the quality of the data used. The data should be accurate, complete, and representative of the donor population. Therefore, cleaning of data and preprocessing are critical phases in the ML pipeline [10]. *Ethical and legal considerations:* The use of machine learning methods for blood donor risk assessment raises several ethical and legal issues, such as data privacy, transparency, and bias. These issues need to be addressed to ensure that the use of predictive analytics is ethical and beneficial to all stakeholders. *Future directions* [11]: The use of predictive analytics for blood donor risk assessment is still in its early stages, and there is much room for improvement. Future research can focus on developing more accurate and efficient models, incorporating new data sources, and addressing ethical and legal issues.

The fields of ML and AI (artificial intelligence) have advanced greatly in recent years. Image analysis, computerized diagnosis, synthesizing images, classification of images, segmenting approaches, feature extraction from X-ray pictures, picture-led care, image recovery, and analysis are just some of the healthcare applications that have greatly benefited from the use of ML and AI. Medical professionals may improve their risk assessment and early detection with the use of ML object identification algorithms from photographic pictures, which are a robust and reliable way to present content. In this study, we utilize convolutional neural networks, recurrent neural networks, and other DL methods for risk analysis [12].

The study is organization as follows: The suggested conceptual model, study objectives, and research evaluation are all presented in Sect. 2. In Sect. 3, we provide our methodology, which includes our study design, data gathering, and analytic

procedures. The proposed methodology empirical findings are given in Sect. 4. In Sect. 5, concluding remarks and some prospects and management implications are discussed.

2 Related Work

The chance of a person donating blood is calculated by considering their gender, race or ethnicity, religion, socioeconomic position, age, occupation, distance, directional difficulties, personal motivations, medical anxiety, and the time since their last donation. Donor scores [23] are calculated using a logistic regression model, although this method is not particularly trustworthy owing to overfitting and the absence of a sufficient training dataset.

To determine the factors that motivate people to donate blood [24] and to determine whether young people at an Indian state institution are donors, the study used a variety of visualization and categorization methods. The KNN (k-nearest neighbor classification) technique is used to learn the correlation between blood donor characteristics and make accurate forecasts. It is now possible to address the critical issues that discourage prospective blood donors from giving blood and satisfy the critical need for blood to save lives.

The scholarly literature categorized the results into a few groups. Second, to better understand the reasons behind volunteer donors' donations, blood banks routinely conduct surveys of these donors. For instance, the primary motivators of recurrent blood donation among experienced donors were intention [13], perceived control, expected guilt, moral standard, age, and past donation frequency. Conversely, among new donors, the only variables predicting recurrent blood donations were age and intention. Others have developed research to better understand the motivations behind blood donation. Sojka and Sojka found that friend influence was the most often reported motivation (47.2%), followed by media demands (23.5%), in their 2008 research of more than 500 contributors. Ultimately, they discovered that blood donors' altruism (40.3%), social responsibility (19.7%), and friend influence (17.9%) will drive blood donation in the future. As was previously stated, just 5% of those who are eligible to give do so [14]. Experts in social and behavioral sciences regularly analyze the causes of greater public participation [15]. The theory of planned behavior (TPB), which is still being researched in this sector, is being advanced by multiple investigations, as we have found.

Previous and related work on predictive analytics of blood donor risk assessment using machine learning methods can be divided into three categories: studies that have developed machine learning models for blood donor risk assessment; studies that have evaluated the effectiveness of these models; and studies that have explored the ethical and legal implications of using machine learning in blood donor risk assessment. A decision tree algorithm was utilized to develop a predictive model for blood donor suitability. The model was based on data collected from 13,768 donors and was able to accurately identify high-risk donors [16].

A support vector machine algorithm was used to develop a predictive model for the risk of hepatitis B virus infection among blood donors. The model was based on data from 10,000 donors and was able to identify donors with a high risk of infection [17]. In a study published in the *Journal of Transfusion Medicine*, researchers evaluated the effectiveness of a predictive model for blood donor suitability developed using a neural network algorithm. The model was based on data from 1499 donors and was able to accurately identify high-risk donors [18]. Another study published in the *Journal of Blood Transfusion* evaluated the effectiveness of a predictive model for the risk of HIV infection among blood donors. The model was developed using a random forest algorithm and was based on data from 11,674 donors. The study found that the model was able to accurately identify high-risk donors [19]. The ethical implications of using machine learning are explored in blood donor risk assessment. The study highlighted the need to ensure that the use of machine learning is transparent, accountable, and respectful of donor privacy [20]. The legal implications of using machine learning in blood donor risk assessment: The study highlighted the need to ensure that the use of machine learning is compliant with data protection and privacy laws.

Classifying abilities and the motivations of blood donors are examined using sophisticated models and strategies [25], such as the MLP (multilayer perceptron) and the PNN (probabilistic neural network). Individuals hospitalized for the hospital's transfusion needs are predicted using one of four ML algorithms: LR (logistic regression), RF (random forest), NN (neural networks), or GBT (gradient boosting trees).

Overall, previous and related work on predictive analytics of blood donor risk assessment using machine learning methods has shown promising results in developing accurate and efficient models. However, ethical and legal implications need to be considered to ensure that the use of machine learning is beneficial and does not harm donors or patients.

3 Methodology

The methodology of predictive analytics for blood donor risk assessment using machine learning methods involves several steps (shown in Fig. 1). The following are some of the key steps that can be included:

3.1 Data Collection

The first step is to collect data from blood donors. The data can include demographic information, medical history, laboratory test results, and any other relevant information that can be used to assess the risk associated with blood donation.

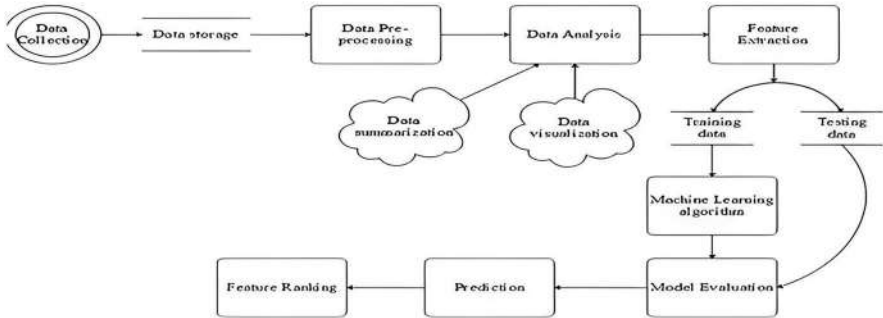


Fig. 1 Framework of the proposed methodology

3.2 Data Cleaning and Preprocessing

The collected data needs to be cleaned and preprocessed to ensure that it is accurate, complete, and consistent. This step involves removing duplicates, filling in missing values, and converting categorical variables to numerical ones.

3.3 Feature Selection

The next step is to select the most relevant features that can be used to predict the risk associated with blood donation. This step involves analyzing the data and identifying the features that have the most significant impact on the outcome.

3.4 RNNs

RNNs can be used to solve problems with estimating time series. They have an input layer, one or more hidden layers, and an output layer. By communicating between the ostensibly concealed levels, the connection in time sequences is established, and the data from the bottom layer is partitioned into the ones below it [21]. Using *RNNs*, data may be maintained despite extensive transmission because of a hidden layer sent via time. This flexibility in working with sequences of varied lengths makes them ideal for use in prediction applications.

3.5 Probabilistic Neural Networks (PNNs)

PNNs are used for the bulk of the analysis since they are superior at training. To explain microbial contamination in the absence of favorable growth circumstances, it combines quantitative principles with machine learning. Because of mutation, the nature of labor has changed [22]. In PNN, nonlinear mappings and clustering are used to identify statistical patterns. To train in one pass, the neural feed-forward network uses a supervised PNN algorithm with three layers. PNN can train on data that is otherwise difficult to collect.

3.6 Model Selection

Once the features are selected, the next step is to choose a suitable ML algorithm that can be used to develop a predictive model. Several algorithms can be used, such as decision trees, random forests, support vector machines, and neural networks.

3.7 Model Training and Testing

The selected machine learning algorithm needs to be trained on the collected data. This step involves splitting the data into training and testing sets and using the training set to train the model. The model's performance is then evaluated on the testing set to determine its accuracy.

3.8 Model Evaluation and Refinement

After testing the model, its performance is evaluated using various performance metrics, such as precision, recall, and F1 score. The model may need to be refined by adjusting its parameters or using a different algorithm to improve its performance.

The methodology of predictive analytics for blood donor risk assessment using machine learning methods involves collecting and preprocessing data, selecting relevant features, choosing a suitable machine learning algorithm, training and testing the model, evaluating and refining the model, addressing ethical and legal considerations, and deploying the model in blood banks and transfusion centers.

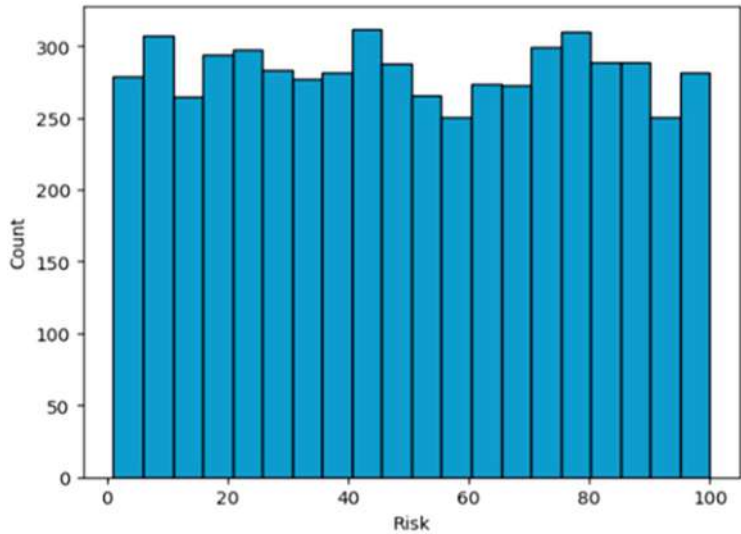


Fig. 2 Premedical history of the risk percent based on the patient count

4 Implementation Results

The data used in this study was prepared using a variety of datasets found online. This dataset contains the blood transfusion details, including all the patient health details and risks of the transfusion (shown in Table 1 and Fig. 2).

- Patient_id is the number given to the patient, which is unique.
- Donor is the blood group of the donor in the transfusion.
- Recipient is the blood group of the recipient in the transfusion.
- CFF_D is the trueness of having a cold, flu, or fever in the donor.
- HIV_D is the trueness of having HIV in the donor.
- AIDS_D is the truth of having AIDS in a donor.

The risk percentage in blood transfusions:

- Age_D is the age of the donor.
- Age_R is the age of the recipient.
- CFF_R is the trueness of having a cold, flu, or fever in the recipient.
- HIV_R is the trueness of having HIV in the recipient.
- AIDS_R is the trueness of having AIDS in the recipient.

All the above attributes of dataset and their predicted values and performance measures are shown in Table 2 and Figs. 3, 4, and 5.

Table 1 The sample of the dataset

Patient_id	Donor	Recipient	CFF_D	HIV_D	AIDS_D	Risk	Age_D	Age_R	CFF_R	HIV_R	AIDS_R
1	A+	B+	1	0	0	24	44	71	1	0	1
2	A+	B+	1	0	1	56	45	59	0	0	0
3	O+	B+	0	0	1	2	31	58	1	1	0
4	A+	A+	0	1	1	80	38	74	1	1	1
5	A+	B+	1	1	1	11	27	23	0	1	0

Table 2 The summary of all recipient blood groups

Patient_id	CFF_R	HIV_R	AIDS_R	Recipient_A+	Recipient_A−	Recipient_B+
653	0	0	0	0	0	1
5541	0	0	0	0	0	1
151	0	0	1	1	0	0
5317	1	0	1	0	0	1
4590	0	0	1	0	0	0

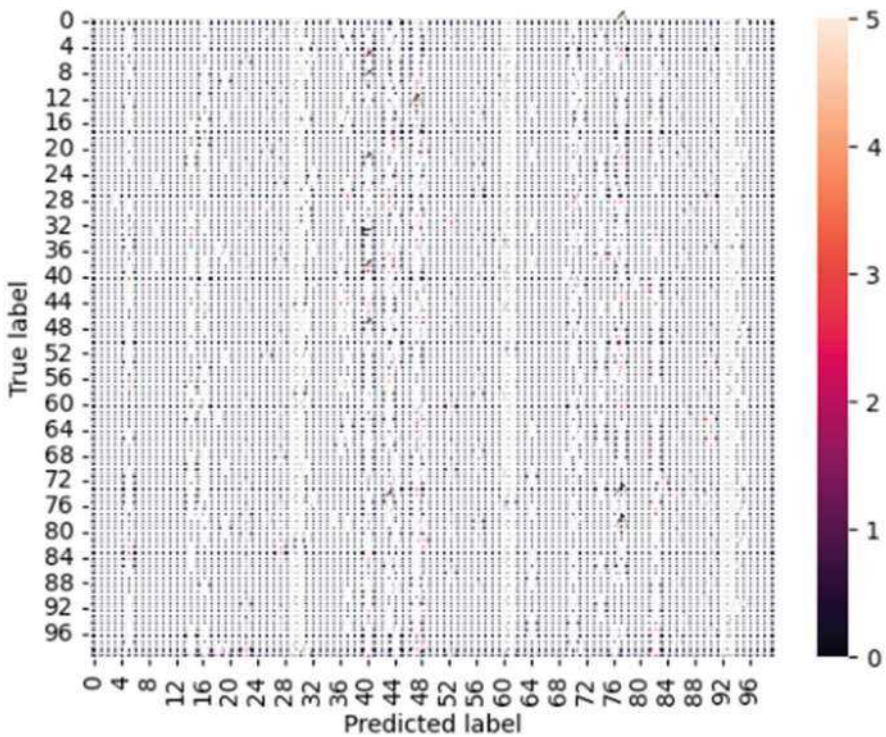


Fig. 3 Confusion matrix on predicted label vs. true label

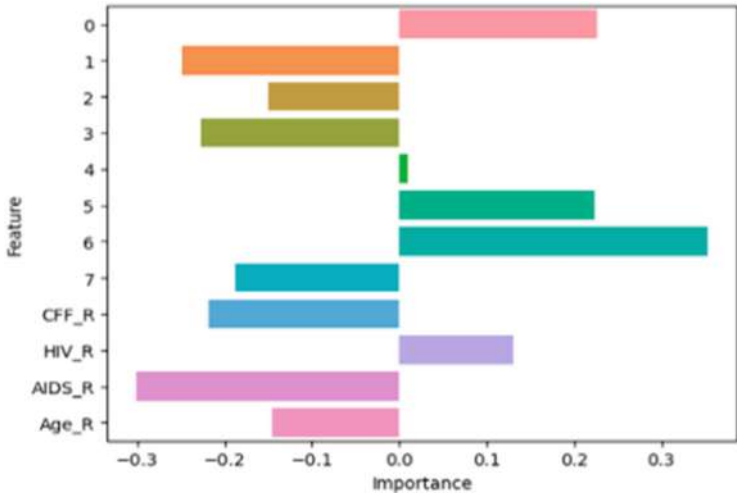


Fig. 4 Plot on features such as CFF, HIV, and importance

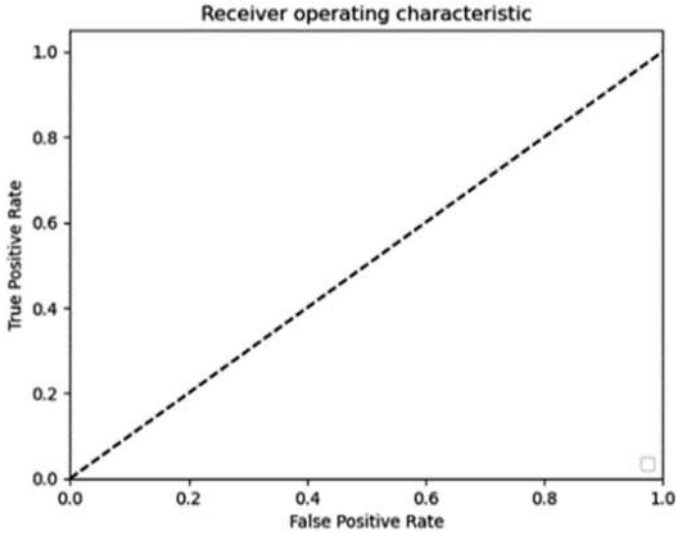


Fig. 5 ROC (receiver operating characteristic) on FPR vs. TPR

5 Conclusion

In conclusion, predictive analytics of blood donor risk assessment using machine learning methods has the potential to improve the safety and efficacy of blood transfusions. Machine learning models can be used to predict the risk associated with blood donation by analyzing donor data, such as demographic information, medical history, and laboratory test results. These models can identify high-risk donors and prevent the transmission of infectious diseases through blood transfusions.

Previous and related work has shown promising results in developing accurate and efficient machine learning models for blood donor risk assessment. However, ethical and legal implications need to be considered to ensure that the use of machine learning is transparent, accountable, and respectful of donor privacy. Data protection and privacy laws need to be followed to prevent any harm to donors or patients. Data gathering, preprocessing, feature selection, model selection, training, testing, evaluation, refinement, and deployment are just a few of the processes in the methodology of predictive analytics for the risk assessment of blood donors using ML techniques. Through employing this method, accurate and efficient machine learning models can be developed and deployed in blood banks and transfusion centers. Overall, the use of predictive analytics for blood donor risk assessment using machine learning methods has the potential to improve the safety and efficacy of blood transfusions. However, it is crucial to consider the ethical and legal implications and follow a rigorous methodology to ensure the benefits of using machine learning outweigh any potential harm.

6 Future Enhancement

There are several ways in which the predictive analytics of blood donor risk assessment using machine learning methods can be enhanced in the future. Here are some potential areas for improvement: (1) *Incorporation of new data sources*: The predictive models can be enhanced by incorporating new data sources, such as genetic data or electronic health records. This can present a more wide-ranging view of the donor's fitness status and improve the accuracy of the risk assessment. (2) *Integration with other healthcare systems*: The predictive models can be integrated with other healthcare systems, such as electronic medical records or clinical decision support methods, to afford a more holistic view of the patient's health status and improve the accuracy of the risk assessment. (3) *Personalized risk assessment*: Machine learning models can be trained to provide personalized risk assessments based on individual donor characteristics. This can help identify donors who may be at higher risk for certain infectious diseases and provide tailored recommendations for risk mitigation. (4) *Real-time risk assessment*: Predictive models can be updated in real-time as new data becomes available. This can enable blood banks and transfusion centers to quickly identify high-risk donors and take appropriate measures to prevent the transmission of infectious diseases. (5) *Ex and interpretability*: Machine learning models can be made more transparent and interpretable by using techniques such as feature importance ranking, partial dependence plots, or local interpretable model-agnostic explanations (LIME). This can help build trust in the models and facilitate their adoption in clinical practice. In the future, a precise forecast and personalized risk assessments, risk mitigation in real-time, and better patient outcomes are all possible due to ML-based predictive analytics in the blood donor risk assessment process.

References

1. WHO. Global status report on blood safety and availability 2016. World Health Organization 2017; Retrieved from: <https://apps.who.int/iris/bitstream/handle/10665/254987/9789241565431-eng.pdf;jsessionid=0D9313B2A0BD8363BE9A3903106FCB7B?sequence=1>
2. Bas S, et al., Management of Blood Donation System: literature review and research perspectives. In Springer Proceedings in Mathematics and Statistics 2016; 169:121–132.
3. Chauhan R, et al., A study to assess the knowledge, attitude, and practices about blood donation among medical students of a medical college in North India. J Family Med Prim Care. 2018 Jul-Aug;7(4):693–697. https://doi.org/10.4103/jfmpe.jfmpe_54_17.
4. Nwogoh B, Aigberadion U, Nwannadi AI. Knowledge, attitude, and practice of voluntary blood donation among healthcare workers at the University of Benin Teaching Hospital, Benin City, Nigeria. J Blood Transfus. 2013;2013:1–6.
5. Flegel WA, Besenfelder W, Wagner FF. Predicting a donor's likelihood of donating within a preselected time interval. Transfusion Medicine. 2000 Sep 1; 10(3):181–92.
6. Stephen Marsland,—Machine Learning—An Algorithmic Perspective, Second Edition, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series, 2014.

7. Tom M Mitchell,—Machine Learning, First Edition, McGraw Hill Education, 2013.
8. Raj A, Gupta A & Selvaraj P. 2018. “Predicting Donor’s Likelihood of donating blood given various factors”. *International Journal of Pure and Applied Mathematics*.118.
9. Jason Bell,—Machine learning—Hands on for Developers and Technical Professionals||, First Edition, Wiley, 2014.
10. Ethem Alpaydin,—Introduction to Machine Learning 3e (Adaptive Computation and Machine Learning Series), Third Edition, MIT Press, 2014.
11. <http://www.americasblood.org/donate-blood/blood-donation-101.aspx>.
12. Kakulapati, V., Prince, A. (2021). Feature Extraction of Coronavirus X-Ray Images by RNN, Correlational Networks, and PNN. In: Oliva, D., Hassan, S.A., Mohamed, A. (eds) *Artificial Intelligence for COVID-19. Studies in Systems, Decision, and Control*, vol 358. Springer, Cham. https://doi.org/10.1007/978-3-030-69744-0_15.
13. Godin G, Conner M, Sheeran P, Bélanger-Gravel A, Germain M. Determinants of repeated blood donation among new and experienced blood donors. *Transfusion*. 2007 Sep;47(9):1607–15. <https://doi.org/10.1111/j.1537-2995.2007.01331.x>
14. Katsaliaki, Korina. (2008). Cost-effective practices in the blood service sector. *Health policy (Amsterdam, Netherlands)*. 86. 276–87. <https://doi.org/10.1016/j.healthpol.2007.11.004>.
15. Ferguson E, France CR, Abraham C, Ditto B, Sheeran P. Improving blood donor recruitment and retention: integrating theoretical advances from social and behavioral science research agendas. *Transfusion*. 2007 Nov;47(11):1999–2010. <https://doi.org/10.1111/j.1537-2995.2007.01423.x>.
16. Deepti Bahel et al., “Predicting Blood Donations Using Machine Learning Techniques”.
17. Harabor V, et al., Machine Learning Approaches for the Prediction of Hepatitis B and C Seropositivity. *Int J Environ Res Public Health*. 2023 Jan 29;20(3):2380. <https://doi.org/10.3390/ijerph20032380>.
18. Ljubomir Buturovic, et al., Evaluation of a Machine Learning-Based Prognostic Model for Unrelated Hematopoietic Cell Transplantation Donor Selection, *Biology of Blood and Marrow Transplantation*, Volume 24, Issue 6, 2018, Pages 1299–1306, ISSN 1083-8791. <https://doi.org/10.1016/j.bbmt.2018.01.038>.
19. Darwiche, et al., (2010). Prediction of blood transfusion donation. 51–56. <https://doi.org/10.1109/RCIS.2010.5507363>.
20. Blood Donor Counselling: Implementation Guidelines. Geneva: World Health Organization; 2014. 4, Ethical and legal considerations in blood donor counseling. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK310579>.
21. Galvez, R.L., et al.: Object detection using convolutional neural networks. In: *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*. Institute of Electrical and Electronics Engineers Inc.; 2019. pp. 2023–2027.
22. Fekrazad, F.: The best approach in intrusion detection for computer network PNN/GRNN/RBF. *Int. J. Comput. Sci. Issues* 11, 182 (2014).
23. Raj A, Gupta A & Selvaraj P. 2018. “Predicting Donor’s Likelihood of donating blood given various factors”. *International Journal of Pure and Applied Mathematics*.118.
24. Pabreja, K. & Bhasin, A. (2021). A Predictive Analytics Framework for Blood Donor Classification. *International Journal of Big Data and Analytics in Healthcare (IJBDAH)*, 6(2), 1–14. <https://doi.org/10.4018/IJBDAH.20210701.oa1>.
25. B M Shashikala., M P Pushpalatha., B Vijaya. (2019). Machine learning approaches for potential blood donor prediction, *Emerging research in electronics, computer science, and technology, Lecture Notes in Electrical Engineering*. 545. https://doi.org/10.1007/978-981-13-5802-9_44

Risk Analysis of COVID-19 Patients Mortality Rate in Emergency Ward



V. Kakulapati , Gadala Praveen, G. Dheeraj, and E. Nagaraju

Abstract A worldwide epidemic of COVID-19 occurred in the year 2020, and it was unprecedented. Contributions in computer science are mostly focused on the creation of techniques for the identification, detection, and prediction of COVID-19 instances. The method most often employed in this field is machine learning (ML). Automatic methods are required for the detection of this illness, given the pace of COVID-19 dissemination. The primary function of an ML model is to forecast death rates based on various symptoms and patient conditions. Health professionals gather data from a variety of patients. The dataset includes several variables that describe the various symptoms and degrees of patient immunity. The dataset includes several variables that describe different signs and degrees of resistance associated with patients. Training the model using several strategies to estimate people's death rates yielded the anticipated result. The data collected includes various factors such as the severity of symptoms, hypertension, and diabetes, and it also includes the count of white blood cells and platelet count as such. A model may be constructed as there is some association between these factors. The accuracy of the gradient matrices created for each technique allows us to choose the best one

Keywords COVID-19 · ML · Pandemic · Mortality · Gradient matrix · Risk · Factors · Algorithms · Health · Model

1 Introduction

The World Health Organization (WHO) first reported COVID-19 as a pandemic in the year 2020. According to several studies, it has infected at least 116 million individuals and killed about 2.5 million people (almost twice the population of Hawaii) [1]. Over 300 million vaccination doses have been given, and over

V. Kakulapati (✉) · G. Praveen · G. Dheeraj · E. Nagaraju
Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

65 million individuals have received all the recommended doses of the vaccinations that have been licensed.

In terms of the direct effects of the public health crisis on health and mortality, there have already been significant repercussions for impoverished and vulnerable populations around the globe. In the last 10 years, innovation for development has gained popularity as a method of attaining development and humanitarian objectives in more inventive and creative ways that may have a greater effect on more people [2]. Only the most obvious types of innovation have occurred in recent years, and they are often simpler to describe and value financially. To combat the pandemic's indirect or ancillary effects, several improvements have been made.

COVID-19 has caused a significant increase in emergency room visits, particularly among patients with severe symptoms. Mortality rates for COVID-19 patients in the emergency ward vary depending on several factors, including age, underlying health conditions, and the severity of the illness [3]. Accurately predicting mortality rates can help healthcare professionals identify high-risk patients and provide appropriate care. To effectively forecast death rates based on these risk variables, a model to analyze the risk factors related to the mortality rates of COVID-19 patients in the emergency department was established in this study. The results of this study can be useful in improving patient outcomes and reducing the burden on healthcare systems [4].

A wide range of technological innovations that try to replicate human thinking and smart behavior fall under the umbrella term “artificial intelligence” (AI). The discipline of machine learning, or ML, is a branch of artificial intelligence that focuses on methods that let computers infer complicated associations or trends from concrete information without having to be defined in advance [5, 6]. The primary capability of the recommended ML technique is the capacity to forecast death rates based on multiple signs and conditions experienced by patients. Health professionals get information from a variety of victims. The dataset includes several variables that describe various ailments and immune rates [7].

It is necessary to impart awareness among the medical personnel and the patients regarding the appropriate usage of large language models based on artificial intelligence, as it is important for them to clearly understand and recognize the advantages and capabilities of the model and make sure that this technology is used ethically and morally to make conclusions and put it to use in real-world applications shortly for betterment [8].

2 Literature Survey

Multiple investigations have explored the potential of ML models to forecast mortality rates among COVID-19 patients. For instance, Liang et al. carried out a study that leveraged a model to predict COVID-19 patient transience in emergency wards, achieving an accuracy level of 87.9% and an AUC-ROC score of 0.927, which suggests that the model could be useful in clinical contexts [9].

An ML-based methodology to forecast COVID-19 patient mortality in intensive care units is employed by utilizing numerous clinical and laboratory parameters, such as age, sex, comorbidities, and laboratory values. The model recorded an AUC-ROC score of 0.883, indicating its capability to predict mortality [10].

A similar investigation to forecast COVID-19 patient mortality in emergency wards using an ML-based model was developed. The model relied on various clinical variables, such as age, sex, comorbidities, and laboratory values, and it achieved an AUC-ROC score of 0.902, implying its potential usefulness in clinical settings [11].

COVID-19 patient mortality in emergency departments was predicted using ML-based models that included numerous clinical and laboratory parameters, including age, sex, comorbidities, and laboratory values. The model achieved an accuracy rate of 85.8% and an AUC-ROC score of 0.912, demonstrating its potential to forecast mortality among COVID-19 patients [12].

ML-based models are used to predict COVID-19 patient outcomes, including mortality, based on clinical and laboratory data. The model achieved an AUC-ROC score of 0.854, indicating its potential utility in clinical settings [13].

A diagnostic model [17] for COVID-19 detection in chest X-rays utilizes convolutional neural networks, radial basis function networks, deep belief networks, and generative belief models. The study's findings were extremely promising for the development of COVID-19 diagnostic tools. VGG16 was used to train ANN, RNN, DNN, and GMB models throughout many epochs. RNN achieved a higher accuracy of 96.6% and an F-measure of 94%, DNN achieved an accuracy of 94.8% and an F-measure of 92%, ANN achieved an accuracy of 93.9% and an F-measure of 91%, and GBM achieved an accuracy of 93.8% and an F-measure of 94%, demonstrating the efficacy of the applied ML approach.

The COVID-19 persistent disease patient database relies on a classification system and the likelihood method of ML techniques [18]. In this case, coronary artery disease, diabetes, renal disease, hypertension, and malignancy are considered as examples of chronic conditions. The mortality rate of these affected individuals is higher than the mortality rate of normally healthy people because the virus attacks the immune system. The findings will aid in the increased monitoring and care given to those with COVID-19 who have had serious medical conditions.

3 Methodology

This work intends to determine the mortality rate of individuals who have contracted the novel coronavirus COVID-19. To achieve this, our model (shown in Fig. 1) gathers data from patients such as age, comorbidities, symptoms, and vital signs such as white blood cell count and hemoglobin levels. Using this data, our model creates a heat map to optimize the process and then trains the collected data using ten different algorithms to estimate the mortality rate of $x\%$ of people. The data collected includes various factors such as the severity of symptoms, ventilation, and

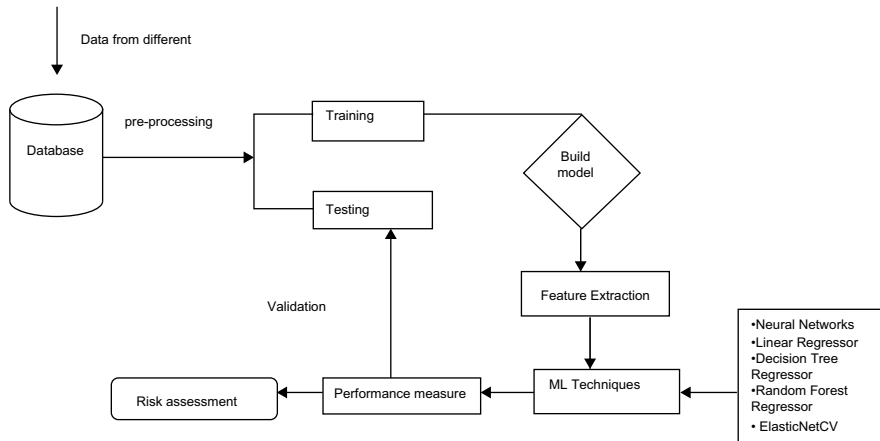


Fig. 1 Framework of proposed work

underlying health conditions such as hypertension, diabetes, cardiovascular disease, COPD, cancer, liver disease, and cerebrovascular disease. The collected data is then analyzed and processed using sophisticated algorithms to generate accurate predictions for the mortality rate.

This study created several data metrics and now needs to apply them to different algorithms, which include:

- NN (neural networks)
- LR (linear regression)
- DTR (decision tree regressor)
- RFR (random forest regressor)
- ElasticNetCV

The proposed model consists of three phases. Each phase will perform the necessary tasks on the data to get the right results.

Phase 1

Collecting the data from Kaggle. This dataset consists of a variety of attributes that explain different symptoms and different immunity levels in patients. It contains the data of patients and converts the data into a heat map and Excel format for further use (shown in Fig. 2).

Phase 2

- After collecting the data, our model will pass it to different algorithms for the best results.
- The algorithms that are used.

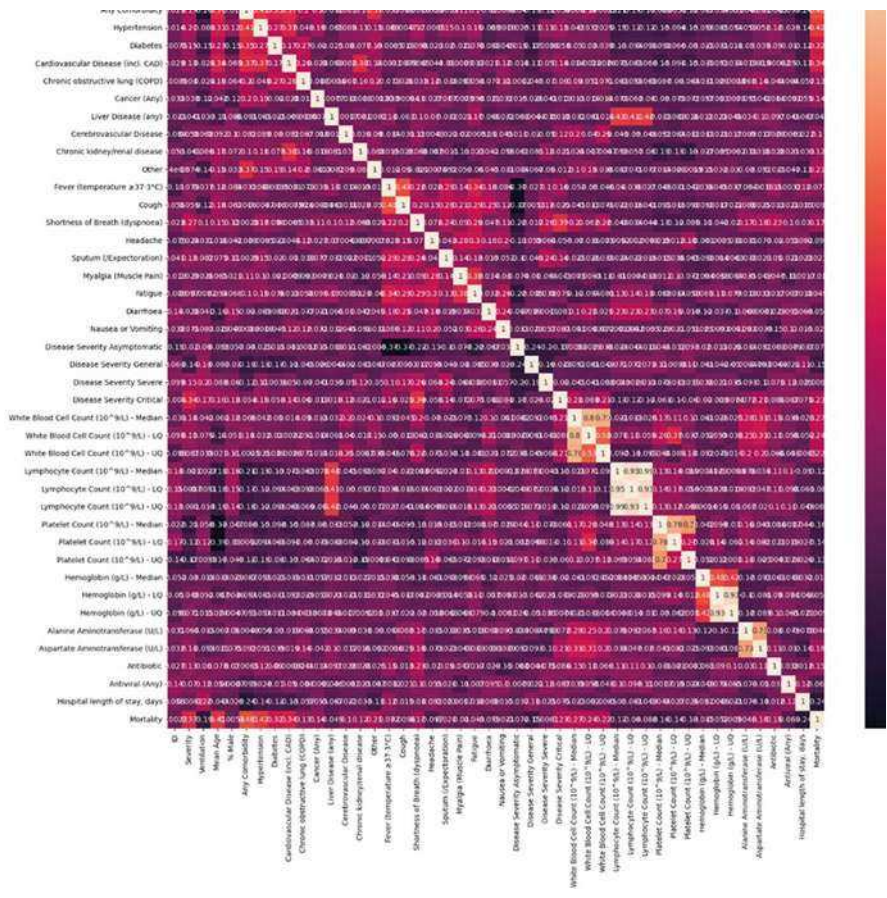


Fig. 2 Heat map of the data based on different comorbidities of patients

3.1 RFR (Random Forest Regressor)

This is an ML technique that uses a combination of several kinds of decision trees (DTs) to provide reliable forecasts. By employing different features and data subsets to build each tree, this method mitigates overfitting and makes the model more resilient. The final prediction is made by aggregating the predictions of all the trees. It is commonly used for classification and regression tasks in various industries because of its accuracy, resilience to noise, and ability to handle large datasets [14].

RF algorithm takes two parameters:

- 1. Number of decision trees
- 2. Random state

In the proposed work, the number of decision trees is 1000 and random states is 42.

Then use a fit function with the RF algorithm to calculate the result on the dataset (shown in Table 4).

3.2 *NN (Neural Networks)*

It consists of interrelated nodes that can be trained from data to formulate forecasts. They can improve over time through training by adjusting their connections. Neural networks are used for various applications, such as image and speech recognition [15].

The neural network has three dense layers

1. First dense—8 neurons
2. Second dense—2670 neurons
3. Final dense—1 neuron (with linear activation function)

3.3 *LR (Linear Regression)*

In this statistical method, a linear equation is used to represent the connection between the dependent variable and independent variables. Predictions are made using scikit-learn and the variable interest. Predictions are made for the test data based on the model's fit to the training data. The linear regression instance is sent to the console along with the root-mean-squared error (RMSE) between the actual and anticipated values (shown in Table 5).

3.4 *DTR (Decision Tree Regressor)*

It is an ML algorithm used for modeling regression problems. It is a nonparametric method that can handle nonlinear relationships between variables and is useful for predicting continuous values. The algorithm aims to minimize the variance in each terminal node of the tree [16]. An instance of the DTR class is created and fitted to the training data using fit(). The predict() method is called on the instance to generate predicted values, and the RMSE is calculated between the actual and predicted values. The Decision Tree Regressor instance is printed to the console. Set the DTR random state to 0 (shown in Table 6).

3.5 *ElasticNetCV*

ElasticNetCV is a method in the scikit-learn library for performing regression analysis using elastic net regularization. Elastic net is an LR model that combines the L1 and L2 regularization penalties in an attempt to get the best of both. ElasticNetCV uses cross-validation to tune the hyperparameters of the model, finding the best combination of L1 and L2 regularization coefficients. This method is useful for dealing with high-dimensional datasets with a large number of features (shown in Table 7) and can handle cases where the number of features exceeds the number of samples [18].

The values we got for using ElasticNetCV in our model are as follows:

[0.12130651	0.15862142	0.05089887	0.10305398	0.12145833	0.04300186
0.06157361	0.25960197	0.06516016	0.1072916	0.06432617	−0.03675575
−0.01513184	0.3788423	−0.18684309	0.15238246	−0.06890483	0.319122
0.01846749	−0.27476106	0.30895667	0.04375987	−0.04392547	0.25972755
−0.18899641	−0.01845868	0.00421478	−0.04214094	0.05606519	0.00445828
0.1993514	0.07414587	0.09708011	0.03385755	0.0836632	0.07012356
0.10185778	0.09856545	0.3206385	0.10398064	0.17350724	0.09061816
0.09061816	0.09061816	0.18608221	0.28154888	0.11060682	0.0549081
0.16745439	0.20142134	0.13756241	0.16796171	0.21979378]	

Phase 3

During phase 3, evaluate the performance of multiple algorithms and select the best one. Then, using this algorithm, estimate the mortality rate of COVID-19 patients with the highest accuracy.

4 Implementation Analysis

The dataset is taken from Kaggle, which is publicly available. After collecting the dataset, preprocess the dataset to remove any redundancy elements. Divide the dataset into 70% training and 30% testing examples. After applying different algorithms to the data, the scores have been recorded and stated in Tables 1, 2, and 3.

Table 1 The dataset contains patient details

All	211
Severe/critical only	144
Mild only	91
Both	5
Mild	2
Severe	2
Severe/critical only	2
Asymptomatic only	1

Table 2 Number of patients on ventilation and non-ventilation

Both	197
Non-ventilation only	66
Ventilation only	19
Yes	12
NA	4
No	3

Table 3 The length of stay in hospital

[11 7.5 12 15 14 16 23.2 nan 17 13 18.5 4 14.5 10 8.4 8.6 18 19.5 13.8
16.4 17.22 17.41 19.43 15.7 15.14 12.67 5 17.81 8 20 16.68 16.5 10.5 12.2
11.6 19 18.3 26.5 21.25 11.2 2 7 12.71 10.56 9 12.5 10.45 11.16 9.27 6.5
12.7 53 51 55 20.8 21 23 4.6 30 13.5 27.5 39 16.63 17.27 14.94 28 15.08
22 26 23.7 24.9 21.9 25.3 31 20.1 'na' 12.9 11.4 15.4 22.5 5.36 0 7.31]

RFR (Table 4)

Table 4 Classification report of RFR

Precision	Recall	F1-score	Support
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
0.67	0.92	0.33	17,979
Accuracy			0.70
Macro avg		0.21	0.09
weighted avg		0.19	0.70

NN

The modeling process employs the Adam optimizer, an indicator, and the mean-squared error (MSE) as the loss function. After training on 32 and 70 epochs of training data, the RMSE between the actual and anticipated outputs is computed. To quantify how far off estimates are from reality, the root-mean-squared error (RMSE) is often used.

[0.12865150],	[0.12865150],	[0.12865150],	[0.10364410],	[0.05995990],	[0.04639436],
[0.05088538],	[0.09222879],	[0.04892305],	[0.05935900],	[0.03815826],	[0.07247222],
[0.07902139],	[0.12865150].				

LR (Table 5)

Table 5 Classification report of LR

Precision	Recall	F1-score	Support
0.22	0.92	0.35	17,979
0.15	0	0.01	12,825
0.21	0.16	0.18	16,068
0.2	0.09	0.13	14,822
0.18	0.05	0.08	12,116
0.25	0.01	0.02	10,362
0.23	0	0	7263
Accuracy			0.66
Macro avg		0.09	0.07
Weighted avg		0.19	0.66

DTR (Table 6)

Table 6 Classification report of DTR

Precision	Recall	F1-score	Support
0.27	1	0.43	17,979
0	0	0	12,825
0.29	0.39	0.33	16,068
0.2	0.06	0.09	14,822
0.22	0.12	0.16	12,116
0.2	0.03	0.06	10,362
Accuracy			0.65
Macro avg.		0.07	0.09
Weighted avg		0.17	0.65

ElasticNetCV (Table 7)

- Random forest: 70.65423345340001
- Linear regression: 66.16835155500729
- Decision tree: 65.59837651922213
- ElasticNet: 66.62837652126366
- Mortality rate = 11%

The results of our study on the risk analysis of COVID-19 patients’ mortality rate in the emergency ward indicate that the Random Forest model is the most accurate model and method for predicting mortality rates, with an R-squared value of 0.706542. This is followed by the elastic net model, linear regression, and decision tree with 0.666, 0.6616, and 0.655, respectively (shown in Figs. 3 and 4 and values are presented in Table 8).

The final mortality rate predicted by the models was 11%, which can be interpreted as the average mortality rate for COVID-19 patients in the emergency ward based on the data used to train the models. These findings can be useful in identifying high-risk patients and providing appropriate care by doctors, physicians, and hospital management.

Table 7 Classification report of ElasticNetCV

Precision	Recall	F1-score	Support
0.22	0.92	0.35	17,979
0.15	0	0.01	12,825
0.21	0.16	0.18	16,068
0.2	0.09	0.13	14,822
0.18	0.05	0.08	12,116
0.25	0.01	0.02	10,362
0.23	0	0	7263
Accuracy			0.66
Macro avg		0.09	0.05
Weighted avg		0.19	0.13

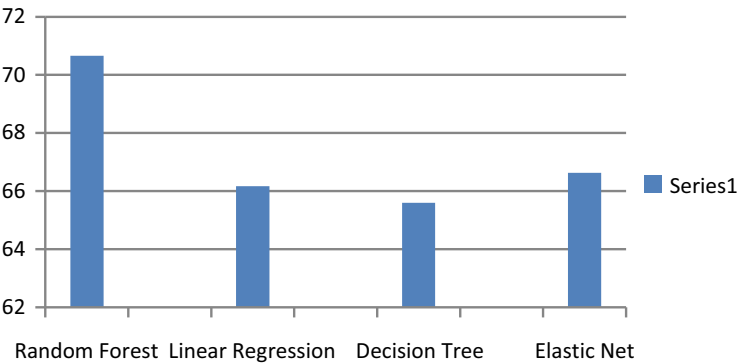


Fig. 3 Different classification regressor model’s accurateness

Fig. 4 Different classification performance measures

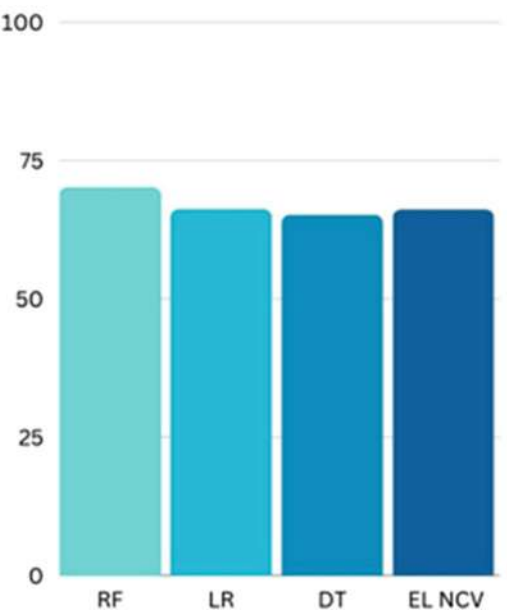


Table 8 The classification accuracies of different algorithms

S. no.	Precision	Recall	F1-score	Support
1	0.67	0.92	0.33	17,979
2	0.58	0	0.01	12,825
3	0.55	0.09	0.03	16,068
4	0.43	0.77	0.08	14,822
5	0.34	0.43	0.02	12,116

5 Discussion

Due to the novel characteristics of coronavirus disease and patients’ other ailments and comorbidities, clinicians and doctors find it challenging to tackle the situation. Healthcare workers in the emergency department and on the front lines had to deal with an unexpected surge of patients while having neither the personnel nor the facilities to properly care for them. The increases in the number of cases with many limitations are responsible for increased mortality in patients with other comorbidities.

Death rates in emergency rooms changed significantly throughout the epidemic, and the findings indicated that the pandemic had a major effect on immediate care. It will aid researchers and policymakers from all over the globe in their pursuit of a solution to the problem of emergency room deaths caused by pandemics. Policymakers, doctors, and hospital managers will all be able to use the study’s findings as a starting point for mitigating and preparing for any future effects.

6 Conclusion

The proposed method examines the potential risks connected with the ER fatality rate among COVID-19 patients. The results of this study on risk analysis of COVID-19 patients' mortality rate in the emergency ward indicate that the Random Forest model is the most accurate model and method for predicting mortality rates, with an R-squared value of 0.706542. This is followed by the elastic net model, linear regression, and decision tree with 0.666, 0.6616, and 0.655, respectively. The final mortality rate predicted by the models was 10.999, which can be interpreted as the average mortality rate for COVID-19 patients in the emergency ward based on the data used to train the models. These findings can be useful in identifying high-risk patients and providing appropriate care by doctors, physicians, and hospital management. Further studies will also help the policymakers take the necessary steps regarding this and be prepared in the future for any hazardous or unexpected situation like the pandemic, which took many lives.

7 Future Enhancement

In the future, more advanced ML approaches, such as deep learning, may be used to better forecast the complicated links between patient features and mortality risk. In order to create a more accurate and robust mortality risk prediction system, future research might investigate the use of ensemble techniques to aggregate the predictions of various models. By adjusting their hyperparameters or trying out new models, the updates will aid healthcare practitioners and policymakers by increasing the precision, clarity, and practicality of the mortality risk of COVID-19 patients treated in the emergency room

References

1. <https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global-excess-mortality>.
2. Renu N. Technological advancement in the era of COVID-19. SAGE Open Med. 2021 Mar 9;9:20503121211000912. DOI: <https://doi.org/10.1177/20503121211000912>. PMID: 33786181; PMCID: PMC7958161.
3. Rieg S, von Cube M, Kalbhenn J, Utzolino S, Pernice K, Bechet L, Baur J, Lang CN, Wagner D, Wolkewitz M, Kern WV, Biever P; COVID UKF Study Group. COVID-19 in-hospital mortality and mode of death in a dynamic and non-restricted tertiary care model in Germany. PLoS One. 2020 Nov 12;15(11):e0242127. <https://doi.org/10.1371/journal.pone.0242127>. PMID: 33180830; PMCID: PMC7660518.
4. Goletti O, Nessi C, Testa A, Albano G, Torri V, Beretta GD, Castoldi M, Bombardieri E. Factors Affecting Mortality in 1022 COVID-19 Patients Referred to an Emergency Department

- in Bergamo during the Peak of the Pandemic. *SN Compr Clin Med*. 2020;2(9):1313–1318. <https://doi.org/10.1007/s42399-020-00444-4>. Epub 2020 Aug 17. PMID: 32838194; PMCID: PMC7430544.
5. Neeru S. Redhu, Zoozeal Thakur, Shikha Yashveer, Poonam Mor, Artificial intelligence: a way forward for agricultural sciences, Editor(s): Pradeep Sharma, Dinesh Yadav, Rajarshi Kumar Gaur, Bioinformatics in Agriculture, Academic Press, 2022, Pages 641–668, ISBN 9780323897785, <https://doi.org/10.1016/B978-0-323-89778-5.00007-6>.
 6. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>
 7. <https://www.hsph.harvard.edu/nutritionsource/nutrition-and-immunity/>
 8. Malik YS, Sircar S, Bhat S, Ansari MI, Pande T, Kumar P, Mathapati B, Balasubramanian G, Kaushik R, Natesan S, Ezzikouri S, El Zowalaty ME, Dhama K. How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future. *Rev Med Virol*. 2021 Sep;31(5):1–11. <https://doi.org/10.1002/rmv.2205>. Epub 2020 Dec 19. PMID: 33476063; PMCID: PMC7883226.
 9. Yan, L., Zhang, HT., Goncalves, J., *et al.* An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell*2, 283–288 (2020). <https://doi.org/10.1038/s42256-020-0180-7>.
 10. Jamshidi E, Asgary A, Tavakoli N, Zali A, Setareh S, Esmaily H, Jamaldini SH, Daaee A, Babajani A, Sendani Kashi MA, Jamshidi M, Jamal Rahi S, Mansouri N. Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. *Front Digit Health*. 2022 Jan 13;3:681608. <https://doi.org/10.3389/fdgth.2021.681608>. PMID: 35098205; PMCID: PMC8792458.
 11. Mahdavi M, Choubdar H, Zabeh E, Rieder M, Safavi-Naeini S, Jobbagy Z, *et al.* (2021) A machine learning-based exploration of COVID-19 mortality risk. *PLoS One* 16(7): e0252384. <https://doi.org/10.1371/journal.pone.0252384>.
 12. Jamshidi E, Asgary A, Tavakoli N, Zali A, Setareh S, Esmaily H, Jamaldini SH, Daaee A, Babajani A, Sendani Kashi MA, Jamshidi M, Jamal Rahi S, Mansouri N. Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. *Front Digit Health*. 2022 Jan 13;3:681608. <https://doi.org/10.3389/fdgth.2021.681608>. PMID: 35098205; PMCID: PMC8792458 <https://www.hindawi.com/journals/sp/2021/5587188>.
 13. <https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/>.
 14. https://www.sas.com/en_us/insights/analytics/neural-networks.html.
 15. <https://www.knowledgehut.com/blog/data-science/linear-regression-for-machine-learning>
 16. <https://www.mastersindatascience.org/learning/machine-learning-algorithms/decision-tree/>
 17. V. Kakulapati, Chapter 16 – Risk analysis of coronavirus patients who have underlying chronic cancer, Editor(s): Yu-Dong Zhang, Arun Kumar Sangaiah, In *Cognitive Data Science in Sustainable Computing, Cognitive Systems and Signal Processing in Image Processing*, Academic Press, 2022, Pages 359–374, ISBN 9780128244104, <https://doi.org/10.1016/B978-0-12-824410-4.00001-5>.
 18. V. Kakulapati *et al.*, Risk analysis of coronaviruses caused death by the probability of patients suffering from chronic diseases – a machine learning perspective. *International Journal Critical Reviews*. 2020; 7(14): 2626–2633. Doi:<https://doi.org/10.31838/jcr.07.14.499>.

Machine Learning Classification Analysis on Leaf Disease Data



Dileep Kumar Kadali , M. Srinivasa Rao , Podagatlapalli Vinay ,
G. A. K. S. Rajeev Kumar , D. V. Naga Raju , and G. Ratnakanth

Abstract Agriculture plays a major role everywhere because humans cannot be left without food in their lives. Many people depend on agricultural cultivation, either directly or indirectly. Leaves actually play an important role in the fast-growing of crops and give a good product but they are facing a major issue of low productivity because of plant leaf disease. Leaf disease occurs due to bacteria, viruses, fungi, pests, and so on. Detecting leaf disease plays an important role in increasing crop productivity. Manual detection is a time-consuming process. This chapter proposes automatic detections using machine learning and image processing techniques for detecting leaf disease. Machine learning techniques are better when compared with image processing techniques. Machine learning classification techniques are very helpful and quickly identify diseases. Machine learning using various classification techniques and evaluation metrics such as accuracy, precision, recall, and F1-Score will show disease progress on the leaf.

Keywords Agriculture · Leaf diseases · Machine learning techniques · Accuracy · Precision · Recall · F1-Score

1 Introduction

India is one of the largest producers of crops in agriculture. Many people depend on agricultural cultivation, either directly or indirectly. Rice may be affected by bacteria, viruses, and fungal diseases [1, 2]. Bacterial leaf streak, leaf scald, and bacterial blight diseases are caused by bacteria. Bakanae, brown spot, narrow-leaf spot, and rice blast are some of the diseases caused by fungus [5, 6]. Grassy stunts and tungro are the diseases caused by the virus. The effect of bacterial leaf streak

D. K. Kadali (✉) · M. S. Rao · P. Vinay · G. A. K. S. R. Kumar · D. V. N. Raju · G. Ratnakanth
Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India
e-mail: msrinivasaraoit@svecw.edu.in; pvinayit@svecw.edu.in; gaksit@svecw.edu.in;
dvnraju@svecw.edu.in; ratnakanth@svecw.edu.in

is the leaf changes from green to light brown. Symptoms of leaf scald are leaf edges or tips turning dark brown. Bacterial blight disease symptoms are white or yellow strips on leaf blades. These may lead to the death of the plant. The diseases in tomatoes are caused mainly by fungus and bacteria [6]. Fungal diseases are anthracnose, black mold, early blight, fusarium wilt, gray mild, leaf mild, septoria leaf spot, and target spot. Bacterial diseases are bacterial canker, bacterial speck, and bacterial spot. These diseases affect either leaves or fruits. Mango leaves are affected by diseases of fungi and bacteria [8, 9]. Another type of disease is the algal leaf spot in which orange color circles will appear on the leaf. Fungal diseases are anthracnose, Phoma blight, pink disease, powdery mildew, and sooty mild. Guava leaves are affected by diseases of fungi [7, 10]. Fungal diseases include anthracnose and pseudo-cercosporin leaf spot. Another disease that affects the leaf is the algal leaf spot. Leaves play a very important role in the fast-growing of crops and give a good product. Plant growth depends mainly on leaves and roots. Leaf disease spoils the total plant. Plant leaf disease may cause great economic loss to farmers [12, 14]. Detecting the diseased leaf is very difficult for farmers and this will be done by experts. They will identify the disease visually and suggest the pesticide for it. The issue is a slow process and the cost is very high. Early identification of leaf diseases can minimize the loss of crops [13, 15]. Machine learning techniques have efficient methods to detect leaf diseases using leaf data [19]. Machine learning is an efficient system with various algorithms models with evaluation metrics such as accuracy, precision, recall, and F1-Score [18, 19]. These will improve their knowledge or performance with experience and show disease progress on the leaf [20, 26]. In lazy prediction, instead of training a model from scratch, we use a pretrained model or a model that has already been trained on a similar dataset [21, 22]. This approach is useful when we have limited time or computational resources [23]. The LazyClassifier will automatically try out multiple classification algorithms and provide with a summary of the results, including accuracy, F1-Score, and training time [24, 25]. It saves you the effort of manually trying out different models and tuning hyperparameters [27, 28].

2 Literature Survey

Many researchers have used computational techniques to detect leaf disease detection. Sardogan et al. detect leaf diseases found in tomato plants [16]. Tomato plants are affected by diseases such as bacterial spots, late blight, septoria leaf spots, and yellow-curved leaf diseases [3]. Goswami et al. explained different types of diseases that are caused by fungi, bacteria, and viruses. To increase the processing time and accuracy change, the segmentation and feature extraction techniques are used. For better accuracy, k-means segmentation and edge detector are used to detect the closer portion of the disease. In feature extraction, the support vector machine (SVM) gained better accuracy compared with other techniques [4]. Hidayatulloh et al. used SqueezeNet Architecture to detect the disease in tomato leaves [3].

Increasing the data can improve the system for better results. The advantage of using the SqueezeNet model is a good choice for implementing mobile devices [5]. Jiang et al. detect diseases in apple leaves [4]. These are spots of Alternaria leaf, gray, brown, mosaic, and rust. INAR-SSD model provides the best result than other existing models. The drawback is that the identification of Alternaria and gray spot is difficult due to similar characteristics. This leads to low accuracy. The advantage of this system avoids the fusion among classes [11]. Kurale et al. detect leaf diseases such as early blight, late blight, and black rot [11]. KNN classifier gives better results and also reduces time when compared to other methods. The disadvantage of using a k-means clustering algorithm is that it uses only for a single database at a time [16]. Shetty et al. detect the leaf diseases of four classes [17]. Results obtained from different methods range between 70% and 90%. The algorithm is developed in the Python tool [17]. Fuentes et al. proposed a leaf disease classification method based on a combination of color and texture features extracted from leaf images. They utilized a support vector machine (SVM) classifier to distinguish between healthy and diseased leaves [29]. Kaur et al. study a deep learning-based approach using convolutional neural networks (CNNs), which was employed for the identification and classification of various crop diseases. The authors trained a CNN architecture on a large dataset of leaf images, achieving high accuracy in disease classification [30]. Brahimi et al. introduced a deep learning model based on a combination of CNN and recurrent neural networks (RNNs) for leaf disease classification. The proposed model effectively captured spatial and temporal dependencies in leaf images and achieved improved classification performance [31]. Ghosal et al. focused on the identification and classification of mango leaf diseases using deep learning techniques. The authors employed a pretrained CNN model, followed by a fine-tuning process on their dataset [33]. The proposed approach achieved accurate disease classification and outperformed traditional machine learning algorithms [32].

3 Proposed Work

The proposed work has evaluated metrics for all types of classification techniques in the supervised learning models.

3.1 Evaluation Metrics

Evaluation metrics are very useful and secure for machine learning tasks. Using metrics from the classification, that are used proceeds the further process here gets the metrics are accuracy, F1-Score, and precision-recall, which are useful for multiple tasks. Using these metrics for concert evaluation for every classification model, it had better improve the model's general predictive power based on earlier

and spool it on sale for making on hidden data. Deprived of concern, an appropriate evaluation process of the machine learning classification model with evaluation metrics, and here individual conditional on the accuracy can prime value to a problematic procedure when the certain model is organized on hidden facts and may conclude in underprivileged final predictions.

3.2 Classification

Classification is approximately predicting the target specified based on contribution data. In binary variable classification, individual dual likely yield classes are close together. The process of the multi-class classification models is done by classification system architecture for leaf disease, further, than two likely classes Jerry can stay extant shown in Fig. 1.

Confusion Matrix A confusion matrix is well-defined as a table that is frequently used to define the concept of a classification model following an established of test data for which the true values are known, which is shown in Table 1. It is enormously convenient for measuring the precision, recall, AUC-ROC curves, and accuracy.

Accuracy: Accuracy simply refers to how frequently the model classifier appropriately expects the prediction, and accuracy is the proportion of the count of correct expects of the values of the prediction and the total count of overall prediction values.

Fig. 1 Classification system architecture for leaf disease

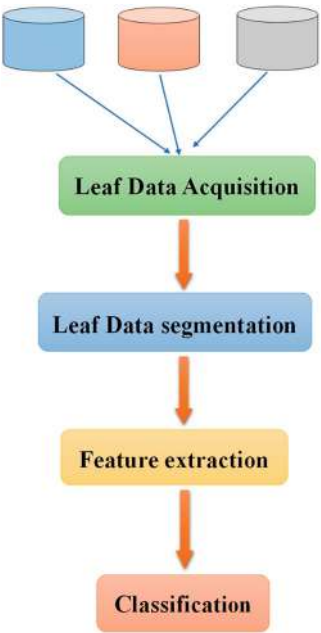


Table 1 Confusion matrix

Predicted values	Actual values		
		Positive (1)	Negative (0)
	True (1)	TPTrue positive	TNTrue negative
	False (0)	FPFalse positive	FNFalse negative

True Positive (TP): The count value of predicted positive from the cross table and its value is true.

True Negative (TN): The count value of predicted negative from the cross table and its value is true.

False Positive (FP) (Type 1 Error): The count value of predicted positive from the cross table and its value is false.

False Negative (FN) (Type 2 Error): The count value of predicted negative from the cross table and its value is false.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision: The precision value from the target is well-defined by way of the count of true positive values divided by the overall count of predicted positive values.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

Recall (Sensitivity): The recall value from the target is well-defined by way of the count of true positives divided by the total count of actual positives values.

$$\text{Accuracy} = \frac{TP}{TP + FN} \quad (3)$$

F1-Score: The F1-Score will give a united knowledge of recall and precision metrics. It is thoroughgoing that the eo recall.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3.3 Leaf Disease Classification Techniques

In leaf disease classification, machine learning and image processing techniques are used. The following are different types of machine learning techniques:

- *Naïve Bayes classifier*: It is a classification system based on the Bayes theorem. The naïve Bayes model is informal to shape and mostly useful for large datasets. It is also called a generative learning model.
- *Support vector machine (SVM)*: SVM is frequently applied for pattern recognition and classification. It is also used to classify the images.

- *k-nearest neighbor* (KNN): k-nearest neighbor is a simple algorithm that stores data and classifies based on similarity measures.
- *Decision trees*: It is a tree representation to solve the problem. Class labels are represented at leaf nodes and attributes at internal nodes.
- *Random forest*: It is an algorithm that randomly generates and combines multiple decision trees into one forest.

4 Experiment Result

Considering leaf diseases dataset that have various kinds of leaves and that’s disease information based on various features. Here, we observed 30 types of leaf information affected by the disease. The dataset can be split into train data and

Table 2 Evaluation metrics result for various classification models

Classification model	Accuracy	Balanced accuracy	ROC AUC	F1-Score	Time taken
LabelSpreading	0.55	0.55	0.55	0.55	3.20
LabelPropagation	0.54	0.54	0.54	0.54	2.23
BernoulliNB	0.54	0.54	0.54	0.45	0.13
AdaBoostClassifier	0.54	0.54	0.54	0.54	0.22
RandomForestClassifier	0.54	0.54	0.54	0.54	1.63
ExtraTreesClassifier	0.53	0.53	0.53	0.53	0.63
LGBMClassifier	0.52	0.52	0.52	0.52	0.35
XGBClassifier	0.52	0.52	0.52	0.52	0.93
PassiveAggressiveClassifier	0.52	0.52	0.52	0.45	0.02
NuSVC	0.52	0.52	0.52	0.51	4.21
BaggingClassifier	0.52	0.52	0.52	0.51	0.23
SVC	0.51	0.51	0.51	0.46	3.56
DecisionTreeClassifier	0.51	0.51	0.51	0.51	0.05
NearestCentroid	0.51	0.51	0.51	0.46	0.01
Perceptron	0.51	0.51	0.51	0.51	0.01
ExtraTreeClassifier	0.51	0.51	0.51	0.51	0.02
GaussianNB	0.50	0.51	0.51	0.40	0.01
QuadraticDiscriminantAnalysis	0.50	0.50	0.50	0.48	0.02
KNeighborsClassifier	0.50	0.50	0.50	0.50	0.03
CalibratedClassifierCV	0.50	0.50	0.50	0.33	0.78
DummyClassifier	0.50	0.50	0.50	0.33	0.01
LogisticRegression	0.49	0.50	0.50	0.35	0.03
LinearDiscriminantAnalysis	0.49	0.50	0.50	0.35	0.08
RidgeClassifier	0.49	0.50	0.50	0.35	0.04
RidgeClassifierCV	0.49	0.50	0.50	0.35	0.03
LinearSVC	0.49	0.50	0.50	0.35	0.26
SGDClassifier	0.44	0.44	0.44	0.41	0.02

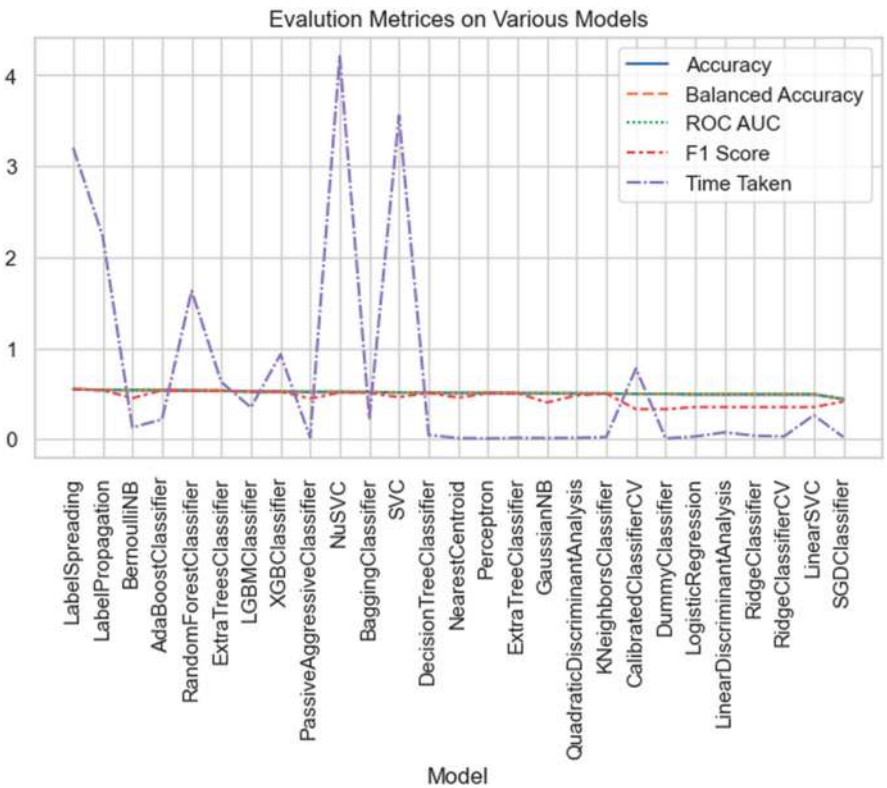


Fig. 2 Evaluation matrices comparison for various classification models

test data from more than 10,000 records about various leaves. Then, we apply the machine learning classification techniques. The classification techniques are shown in Table 2 for evaluation metrics information. Based on that evaluating the matrices are balanced accuracy, accuracy ROC, and F1-Score with maximum evaluation time. Here, we used lazy predict supervised learning, and it is a very useful and simple technique to evaluate metrics and it gives additional information time complexity of each classification model with existing matrices.

In the comparison of each model matrices along with time complexity, the NuSVC model can be taken with the highest time for the worst complexity model. Many models can give results with less and the best time complexity as shown in Fig. 2. Based on this information, we can identify a good model for disease identification purposes.

The total count of the model accuracy statics will be fit into values between 0 to 1 based on pictorial information for each mode in descending order as shown in Fig. 3.

Accuracy can be improved using autonomous learning based on supervised machine learning classification models. The lazy predict learning approaches

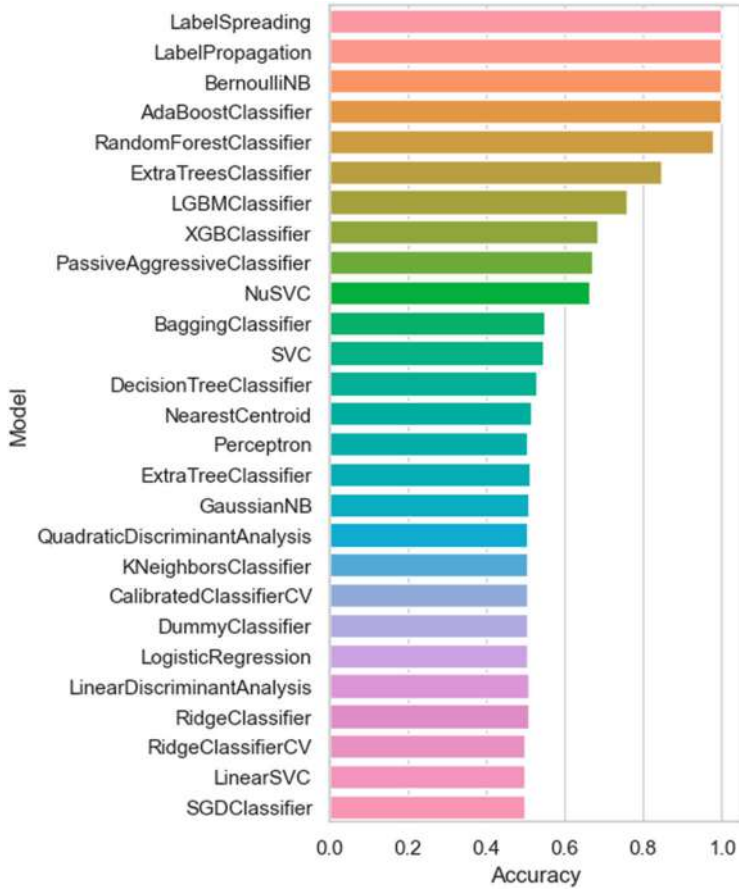


Fig. 3 Accuracy comparison for various classification models

evaluate all classification models by increasing the dataset size so that the size of the training set and testing set will be increased and acquired with high clarity as shown in Fig. 4. It shows capabilities for plant leaf disease detection and classification can be improves in the real environment.

5 Conclusion

This chapter summarizes the evaluation metrics for all classification models in machine learning. To do that process, various leaf disease parameter values data such as bacteria, viruses, fungi, and pests are considered. Machine learning classification techniques are very helpful in quickly identifying diseases. Finally, it

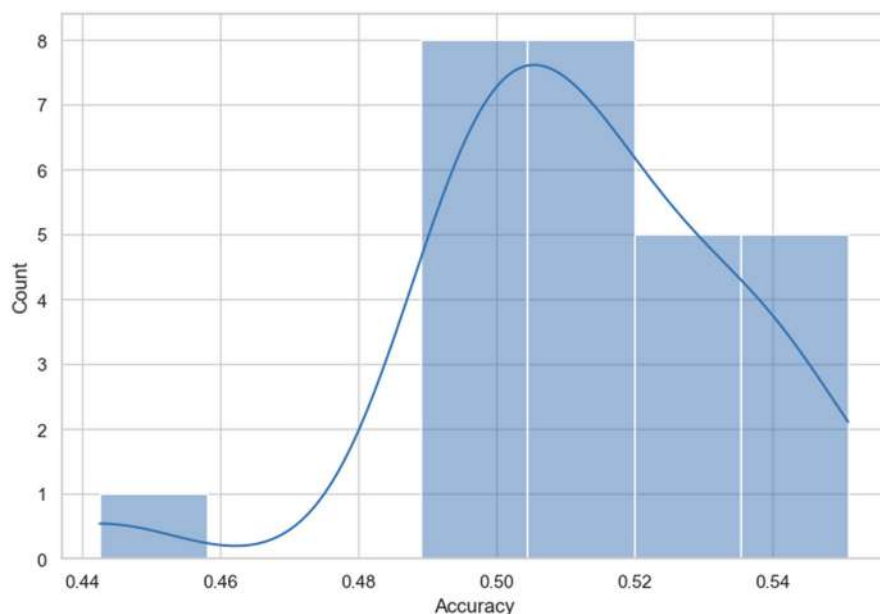


Fig. 4 Accuracy density representation for various classification models

summarizes the evaluation metrics values such as precision, accuracy, recall, and F1-Score. These show the progression of disease on the leaf.

References

1. Devaraj, A., Rathan, K., Jaahnavi, S., & Indira, K. (2019, April). Identification of Plant Disease using Image Processing Technique. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0749-0753). IEEE.
2. Goswami, M., Maheshwari, S., & Poonia, A. (2018). Performance Analysis of Classifiers and Future Directions for Image Analysis Based Leaf Disease Detection. In Recent Findings in Intelligent Computing Techniques (pp. 519-525). Springer, Singapore.
3. Hidayatuloh, A., Nursalman, M., & Nugraha, E. (2018, October). Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model. In 2018 International Conference on Information Technology Systems and Innovation (ICITSI) (pp. 199-204). IEEE.
4. Jiang, P., Chen, Y., Liu, B., He, D., & Liang, C. (2019). Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, 7, 59069-59080.
5. RNV Jagan Mohan, Dileep Kumar Kadali and et al. "Crime data optimization using Neutrosophic logic based game theory", *Concurrency and Computation: Practice and Experience*, 2022.
6. Kadali, Dileep Kumar et al. 'Machine Learning Approach for Corona Virus Disease Extrapolation: A Case Study'. 1 Jan. 2022: 219—227.

7. Kaur, S., Pandey, S., & Goel, S. (2019). Plants disease identification and classification through leaf images: A survey. *Archives of Computational Methods in Engineering*, 26(2), 507-530.
8. Dileep Kumar Kadali, R.N.V. Jagan Mohan presented a paper titled "Shortest Route Analysis for High level Slotting Using Peer-to-Peer", *The Role of IoT and Blockchain: Techniques and Applications* CRC press, 2022.
9. Kumari, C. U., Prasad, S. J., & Mounika, G. (2019, March). Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1095-1098). IEEE.
10. Kadali, D.K., Venkata Naga Raju, D., Venkata Rama Raju, P., "Cluster Query Optimization Technique Using Blockchain". In: Kumar, A., Ghinea, G., Merugu, S. (eds) *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing. ICCIC 2022. Cognitive Science and Technology*. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_65
11. Kurale, N. G., & Vaidya, M. V. (2018, July). Classification of Leaf Disease Using Texture Feature and Neural Network Classifier. In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 1-6). IEEE.
12. Dileep Kumar Kadali, M Chandra Naik, RNV Jagan Mohan "Estimation of Data Parameters Using Cluster Optimization", *EasyChair and Springer Nature*, 2022.
13. Dileep Kumar Kadali, R.N.V. Jagan Mohan, M. Chandra Naik, Unsupervised based crimes cluster data using decision tree classification. *J. Solid State Technol.* 63(5), 5387-5394, 2020.
14. Raghavendra, B. K. (2019, March). Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 313-316). IEEE.
15. Dileep Kumar Kadali, R.N.V. Jagan Mohan, M. Srinivasa Rao, Cluster optimization for similarity process using de-duplication. *IJSRD-Int. J. Sci. Res. Dev.* 4(06), 2321-0613. ISSN (online) (2016)
16. Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 382-385). IEEE.
17. Suma, V., Shetty, R. A., Tated, R. F., Rohan, S., & Pujar, T. S. (2019, June). CNN based Leaf Disease Identification and Remedy Recommendation System. In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 395-399). IEEE.
18. D. K. Kadali, R. Mohan, and M. C. Naik, "Enhancing crime cluster reliability using Neutrosophic logic and a Three-Stage model," *Journal of Engineering Science and Technology Review*, vol. 16, no. 4, pp. 35-40, Jan. 2023, <https://doi.org/10.25103/jestr.164.05>.
19. Vaishnnave, M. P., Devi, K. S., Srinivasan, P., & ArutPerumJothi, G. (2019, March). Detection and Classification of Groundnut Leaf Diseases using KNN classifier. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-5). IEEE.
20. Dileep Kumar Kadali, R.N.V. Jagan Mohan, Optimizing the duplication of cluster data for similarity process. *ANU J. Phys. Sci.* 2. ISSN: 0976-0954 (2014).
21. Dileep Kumar Kadali, R.N.V. Jagan Mohan, Y. Vamsidhar, "Similarity based query optimization on map reduce using Euler angle oriented approach". *Int. J. Sci. Eng. Res.* 3(8). ISSN 2229-5518 (2012).
22. Ani Brown Mary, N., Robert Singh, A., Athisayamani, S. "Classification of Banana Leaf Diseases Using Enhanced Gabor Feature Descriptor". In: Ranganathan, G., Chen, J., Rocha, Á. (eds) *Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems*, vol 145. Springer, Singapore, 2021. https://doi.org/10.1007/978-981-15-7345-3_19
23. A. Jenifa, R. Ramalakshmi and V. Ramachandran, "Classification of Cotton Leaf Disease Using Multi-Support Vector Machine," 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Tamilnadu, India, 2019, pp. 1-4, <https://doi.org/10.1109/INCOS45849.2019.8951356>.

24. Dinesh, A., Maragatharajan, M., Balakannan, S.P. "Survey on Machine Learning Algorithm for Leaf Disease Detection Using Image Processing Techniques". In: Suma, V., Baig, Z., Kolandapalayam Shanmugam, S., Lorenz, P. (eds) *Inventive Systems and Control. Lecture Notes in Networks and Systems*, vol 436. Springer, Singapore, 2022. https://doi.org/10.1007/978-981-19-1012-8_47
25. Deepalakshmi P., et al. "Plant Leaf Disease Detection Using CNN Algorithm." *IJISMD* vol.12, no.1 2021: pp.1–21. <https://doi.org/10.4018/IJISMD.2021010101>
26. D. P. Rajan, K. S. Kannan, P. Divya and S. Velliangiri, "Comparative Analysis of Liver diseases by using Machine Learning Techniques," 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2022, pp. 01–05, <https://doi.org/10.1109/ESCI53509.2022.9758256>.
27. Babu, S., Maravarman, M., Pitchai, R. "Detection of Rice Plant Disease Using Deep Learning Techniques" *Journal of Mobile Multimedia*, Volume 18, Issue 3, 2022, Pages 757–770, <https://doi.org/10.13052/jmm1550-4646.18314>
28. C. S. K. Raju, K. Pranitha, S. Sundari M, P. Samyuktha and J. Madhumathi, "Prediction of COVID 19- Chest Image Classification and Detection using RELM Classifier in Machine Learning," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 1184–1188, <https://doi.org/10.1109/ICACCS54159.2022.9785131>.
29. Satyanarayana, B.V.V., Kumar, G.P., Varma, A.K.C., Dileep, M., Srinivas, Y., Budumuru, P.R. "Alzheimer's Disease Detection Using Ensemble of Classifiers. In: Gupta, N., Pareek, P., Reis, M. (eds) *Cognitive Computing and Cyber Physical Systems. IC4S 2022. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 472. Springer, Cham. https://doi.org/10.1007/978-3-031-28975-0_5.
30. Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2018). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 18(11), 3765.
31. Kaur, H., Khanna, S., Khanna, A., & Vatsa, M. (2019). Plant disease classification using deep convolutional neural network. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.
32. Brahimi, M., Djamaa, S., & Adjabi, S. (2020). Leaf disease identification using deep learning. *Multimedia Tools and Applications*, 79(45), 33271-33290.
33. Ghosal, S., Gupta, D., Gupta, M., & Agarwal, P. (2021). Mango leaf disease detection using deep learning techniques. In *Innovations in Intelligent Computing and Applications* (pp. 325-334). Springer.

Conversion of Type-2 Intuitionistic Fuzzy Sets into Interval-Valued Intuitionistic Fuzzy Sets and Its Implementation in Decision-Making



V. Sireesha, N. Annapurna, and V. Anusha

Abstract The chapter aims to study the conversion of Type-2 intuitionistic fuzzy set (T2IFS) to interval-valued intuitionistic fuzzy set (IVIFS). A conversion method of T2IFS to IVIFS is presented without any loss of indeterminacy degree exists in intuitionistic fuzzy sets (IFS). The necessary properties related to the conversion are examined thoroughly with proofs, and the conversion is illustrated through numerical examples. Furthermore, the proposed conversion method is implemented to a random T2IF problem from the literature, and the results are analyzed.

Keywords Conversion · IVIFSs · T2IFSs · Ranking · Decision-making

1 Introduction

Zadeh [1] introduced standard fuzzy sets (FS), modeling the uncertainty using membership grade of each element. Later, Atanassov [2] proposed a generalization of FS in which the non-belongingness amount of each element also considered and named it as an intuitionistic fuzzy (IF) set. The most suitable way of information collection in fuzzy decision theory is to congregate the data considering the degree of nonmembership along with membership. Therefore, more generalizations of intuitionistic fuzzy set (IFS), interval-valued intuitionistic fuzzy sets (IVIFSs) [3], and Type-2 intuitionistic fuzzy sets (T2IFSs) [4, 5] were offered.

IVIFS is a combining concept of IFS and IVFS. For their advantage of coping with more imprecise information, they have been adequately applied in decision-making. The literature survey suggests that decision-making problems with IVIF information significantly impacted the uncertain system and received much attention from researchers. Furthermore, the IVIF decision information provides better

V. Sireesha (✉) · N. Annapurna · V. Anusha

Department of Mathematics, GITAM Institute of Science, GITAM (Deemed to be University),
Visakhapatnam, Andhra Pradesh, India

e-mail: sveerama@gitam.edu

results in evaluating alternatives for both dependent and independent criteria. Thus, numerous decision-making methods were developed to solve IVIF multicriteria decision-making (MCDM) problems.

T2IFS is a generalized form of Type-2 fuzzy set and IFS having primary membership, secondary membership grades along with primary nonmembership and secondary nonmembership grades. While modeling decision-making problems, it enables more degree of freedom; offering decision-maker a vast scope to represent the performance values by considering T2IFSs [6, 7]. The ability to capture the degrees of membership of relevant membership values and vice-versa for non-membership, handling the uncertainty is more accurately is stated as fundamental superiority of T2IFS.

Research has been conducted considering the switching and interrelationships between several types of fuzzy information (Vague sets and Atanassov IFSs [11], Pythagorean Fuzzy Sets (PFSs) and IFSs [10], PFSs and IFSs [12], IFS and IVIFS [14], Type-2 fuzzy sets and IFSs [8]). By connecting two types of fuzzy information, one can discuss the conclusions for the given problem using both types of fuzzy information.

Even though the implementation of the IVIF in decision-making is intuitive and simple, the Type-2 intuitionistic fuzzy sets possess great expressive power and are thus conceptually quite appealing [15]. However, the use of T2IFS is limited as the computational complexity of dealing with them is greater than those for dealing with IVIFS [15]. When a fuzzy number fails to produce accurate results or a decision-maker perceives ambiguity in the overall outcome, one can employ the conversion method to study the same problem in a different fuzzy number by converting it into another fuzzy number. Thus, the aim of the chapter is to develop a conversion process of a fuzzy number T2IF to IVIF. Moreover, using the proposed conversion, one can study a T2IF decision problem in IVIF environment, which has proven to be a flexible mathematical tool to depict and handle uncertainty in decision-making problems.

The chapter is summarized as follows. Section 2 covers the preliminaries of the fuzzy sets used in the study. In Sect. 3, the conversion is proposed, necessary properties are verified, and the conversion is demonstrated with numerical examples. In Sect. 4, a problem from literature is discussed and results are analyzed. Section 5 is the conclusion.

2 Preliminaries

The section briefs about basic definitions and operations of IVIFSs and T2IFSs.

Definition 1: Type-2 Intuitionistic Fuzzy Set (T2IFS) [5]

A T2IFS T on the universal set X is defined as

$$T = \{ \langle x, (\mu_T(x), g_x(\mu_T), \vartheta_T(x), h_x(\vartheta_T)) \rangle / x \in X, \mu_T \in [0, 1], \vartheta_T \in [0, 1] \}$$

where μ_T and ϑ_T denote the primary membership function (PMF) and primary nonmembership function (PNMF) such that $0 \leq \mu_T(x) + \vartheta_T(x) \leq 1$ and $g_x(\mu_T)$ membership grade of the PMF is named as secondary membership function (SMF) and $h_x(\vartheta_T)$ is the nonmembership grade of the PNMF named as and secondary nonmembership function (SNMF) such that $0 \leq g_x(\mu_T) + h_x(\vartheta_T) \leq 1$.

Definition 2: Variance Margin Function of T2IFS [5]

The variance margin function (VMF) of T2IFS is defined as the difference between PMF and PNMF; SMF and SNMF. It is denoted by ξ and η , i.e., for T2IFS T , variance margin functions are

$$\xi_T(x) = |\mu_T(x) - g_x(\mu_T)| \text{ and } \eta_T(x) = |\vartheta_T(x) - h_x(\vartheta_T)| \text{ for all } x \in X.$$

Definition 3: Set operations of T2IFSs

For any two T2IFS $S = \{ \langle x, (\mu_S(x), g_x(\mu_S), \vartheta_S(x), h_x(\vartheta_S)) \rangle / x \in X, \mu_S \in [0, 1], \vartheta_S \in [0, 1] \}$ and $T = \{ \langle x, (\mu_T(x), g_x(\mu_T), \vartheta_T(x), h_x(\vartheta_T)) \rangle / x \in X, \mu_T \in [0, 1], \vartheta_T \in [0, 1] \}$,

- (i) $S \cup T = \{ \max(\mu_S(x), \mu_T(x)), \max(g_x(\mu_S), g_x(\mu_T)), \min(\vartheta_S(x), \vartheta_T(x)), \min(h_x(\vartheta_S), h_x(\vartheta_T)) \}$
- (ii) $S \cap T = \{ \min(\mu_S(x), \mu_T(x)), \min(g_x(\mu_S), g_x(\mu_T)), \max(\vartheta_S(x), \vartheta_T(x)), \max(h_x(\vartheta_S), h_x(\vartheta_T)) \}$

Definition 4: Interval-Valued Intuitionistic Fuzzy Sets (IVIFSs) [3]

Let $E = \{x_1, x_2, \dots, x_n\}$ be a subset of universal set X , then the set

$$\tilde{I} = \left\{ \langle x_i, [\mu_I^L(x_i), \mu_I^U(x_i)], [\vartheta_I^L(x_i), \vartheta_I^U(x_i)] \rangle / x_i \in E \right\}$$

$$[\mu_I^L(x_i), \mu_I^U(x_i)] \subseteq [0, 1] \text{ and } [\vartheta_I^L(x_i), \vartheta_I^U(x_i)] \subseteq [0, 1], \text{ satisfying}$$

$$0 \leq \mu_I^U(x_i) + \vartheta_I^U(x_i) \leq 1 \text{ is an IVIFS.}$$

For each $x_i \in E$, the interval hesitancy degree of any IVIFS \tilde{I} is defined as

$$\pi_I(x_i) = [1 - \mu_I^U(x_i) - \vartheta_I^U(x_i), 1 - \mu_I^L(x_i) - \vartheta_I^L(x_i)]$$

Definition 5: Set Operation on IVIFSs [9]

For any two IVIFSs, $\tilde{I}_1 = \left\{ \langle x_i, [\mu_{I_1}^L(x_i), \mu_{I_1}^U(x_i)], [\vartheta_{I_1}^L(x_i), \vartheta_{I_1}^U(x_i)] \rangle / x_i \in E \right\}$ and $\tilde{I}_2 = \left\{ \langle x_i, [\mu_{I_2}^L(x_i), \mu_{I_2}^U(x_i)], [\vartheta_{I_2}^L(x_i), \vartheta_{I_2}^U(x_i)] \rangle / x_i \in E \right\}$

$$\begin{aligned}
 \text{(i)} \quad \tilde{I}_1 \cap \tilde{I}_2 &= \left\{ \left\langle x_i, \left[\begin{array}{l} \min \left(\mu_{\tilde{I}_1}^L(x_i), \mu_{\tilde{I}_2}^L(x_i) \right), \\ \min \left(\mu_{\tilde{I}_1}^U(x_i), \mu_{\tilde{I}_2}^U(x_i) \right) \end{array} \right] \right. \right. \\
 &\quad \left. \left. \left[\begin{array}{l} \max \left(\vartheta_{\tilde{I}_1}^L(x_i), \vartheta_{\tilde{I}_2}^L(x_i) \right), \\ \max \left(\vartheta_{\tilde{I}_1}^U(x_i), \vartheta_{\tilde{I}_2}^U(x_i) \right) \end{array} \right] \right. \right. \right\} \\
 \text{(ii)} \quad \tilde{I}_1 \cup \tilde{I}_2 &= \left\{ \left\langle x_i, \left[\begin{array}{l} \max \left(\mu_{\tilde{I}_1}^L(x_i), \mu_{\tilde{I}_2}^L(x_i) \right), \\ \max \left(\mu_{\tilde{I}_1}^U(x_i), \mu_{\tilde{I}_2}^U(x_i) \right) \end{array} \right] \right. \right. \\
 &\quad \left. \left. \left[\begin{array}{l} \min \left(\vartheta_{\tilde{I}_1}^L(x_i), \vartheta_{\tilde{I}_2}^L(x_i) \right), \\ \min \left(\vartheta_{\tilde{I}_1}^U(x_i), \vartheta_{\tilde{I}_2}^U(x_i) \right) \end{array} \right] \right. \right. \right\}
 \end{aligned}$$

Definition 6: Jaccard Distance on T2IFS

For any two T2IFS S, T , the Jaccard distance between S and T denoted by $d_J(S, T)$ is defined as follows:

$$d_J(S, T) = 1$$

$$\frac{\left| \left\{ \min(\mu_P(x), \mu_Q(x)), \min(g_x(\mu_P), g_x(\mu_Q)), \min(\xi_P(x), \xi_Q(x)), \right. \right.}{\left. \left. \max(\vartheta_P(x), \vartheta_Q(x)), \max(h_x(\vartheta_P), h_x(\vartheta_Q)), \max(\eta_P(x), \eta_Q(x)) \right\} \right|}{\left| \left\{ \max(\mu_P(x), \mu_Q(x)), \max(g_x(\mu_P), g_x(\mu_Q)), \max(\xi_P(x), \xi_Q(x)), \right. \right.}{\left. \left. \min(\vartheta_P(x), \vartheta_Q(x)), \min(h_x(\vartheta_P), h_x(\vartheta_Q)), \min(\eta_P(x), \eta_Q(x)) \right\} \right|}$$

The Jaccard distance of any T2IFS S from the ideal T2IFS $I = \{ \langle x, (1, 1, 0, 0) \rangle : x \in X, \mu_T \in [0, 1], \vartheta_T \in [0, 1] \}$ is given by

$$d_J(S, I) = 1$$

$$\frac{\left| \left\{ \min(\mu_S(x), \mu_I(x)), \min(g_x(\mu_S), g_x(\mu_I)), \min(\xi_S(x), \xi_I(x)), \right. \right.}{\left. \left. \max(\vartheta_S(x), \vartheta_I(x)), \max(h_x(\vartheta_S), h_x(\vartheta_I)), \max(\eta_S(x), \eta_I(x)) \right\} \right|}{\left| \left\{ \max(\mu_S(x), \mu_I(x)), \max(g_x(\mu_S), g_x(\mu_I)), \max(\xi_S(x), \xi_I(x)), \right. \right.}{\left. \left. \min(\vartheta_S(x), \vartheta_I(x)), \min(h_x(\vartheta_S), h_x(\vartheta_I)), \min(\eta_S(x), \eta_I(x)) \right\} \right|}$$

Definition 7: Jaccard Distance on IVIFS [13]

For any IVIFSs \tilde{I}_1, \tilde{I}_2 on E , the Jaccard distance on \tilde{I}_1, \tilde{I}_2 is defined as:

$$d_J(\tilde{I}_1, \tilde{I}_2) = 1$$

$$\frac{\left\| \begin{bmatrix} \min \left(\mu_{I_1}^{\sim L}(x_i), \mu_{I_2}^{\sim L}(x_i) \right), \\ \min \left(\mu_{I_1}^{\sim U}(x_i), \mu_{I_2}^{\sim U}(x_i) \right) \end{bmatrix}, \begin{bmatrix} \max \left(\vartheta_{I_1}^{\sim L}(x_i), \vartheta_{I_2}^{\sim L}(x_i) \right), \\ \max \left(\vartheta_{I_1}^{\sim U}(x_i), \vartheta_{I_2}^{\sim U}(x_i) \right) \end{bmatrix} \right\|}{\left\| \begin{bmatrix} \max \left(\mu_{I_1}^{\sim L}(x_i), \mu_{I_2}^{\sim L}(x_i) \right), \\ \max \left(\mu_{I_1}^{\sim U}(x_i), \mu_{I_2}^{\sim U}(x_i) \right) \end{bmatrix}, \begin{bmatrix} \min \left(\vartheta_{I_1}^{\sim L}(x_i), \vartheta_{I_2}^{\sim L}(x_i) \right), \\ \min \left(\vartheta_{I_1}^{\sim U}(x_i), \vartheta_{I_2}^{\sim U}(x_i) \right) \end{bmatrix} \right\|}$$

Here, mod defines the Euclidean distance from the set to the origin.

The Jaccard distance of any IVIFS \tilde{I}_1 from the ideal IVIFS $I = \{<x, [1, 1], [0, 0] > / x \in X, \mu_T \in [0, 1], \vartheta_T \in [0, 1]\}$ is given by

$$d_J \left(\tilde{I}_1, \tilde{I} \right) = 1 - \frac{\left\| \begin{bmatrix} \min \left(\mu_{I_1}^{\sim L}(x), \mu_I^{\sim L}(x) \right), \\ \min \left(\mu_{I_1}^{\sim U}(x), \mu_I^{\sim U}(x) \right) \end{bmatrix}, \begin{bmatrix} \max \left(\vartheta_{I_1}^{\sim L}(x), \vartheta_I^{\sim L}(x) \right), \\ \max \left(\vartheta_{I_1}^{\sim U}(x), \vartheta_I^{\sim U}(x) \right) \end{bmatrix} \right\|}{\left\| \begin{bmatrix} \max \left(\mu_{I_1}^{\sim L}(x), \mu_I^{\sim L}(x) \right), \\ \max \left(\mu_{I_1}^{\sim U}(x), \mu_I^{\sim U}(x) \right) \end{bmatrix}, \begin{bmatrix} \min \left(\vartheta_{I_1}^{\sim L}(x), \vartheta_I^{\sim L}(x) \right), \\ \min \left(\vartheta_{I_1}^{\sim U}(x), \vartheta_I^{\sim U}(x) \right) \end{bmatrix} \right\|}$$

Definition 8: Ranking of IVIFSs [13]

For any two IVIFSs \tilde{I}_1 and \tilde{I}_2 , the ranking is given as follows:

- (i) If $d_J \left(\tilde{I}_1, \tilde{I} \right) < d_J \left(\tilde{I}_2, \tilde{I} \right)$ then $\tilde{I}_1 > \tilde{I}_2$
- (ii) If $d_J \left(\tilde{I}_1, \tilde{I} \right) > d_J \left(\tilde{I}_2, \tilde{I} \right)$ then $\tilde{I}_1 < \tilde{I}_2$
- (iii) If $d_J \left(\tilde{I}_1, \tilde{I} \right) = d_J \left(\tilde{I}_2, \tilde{I} \right)$ then $\tilde{I}_1 = \tilde{I}_2$

The ranking of T2IFS is similar to the above, where the Jaccard distance is computed using Definition 6.

3 Proposed Conversion Methodology

T2IFSs are represented by type-2 membership and nonmembership functions, whereas IVIFSs are represented by its membership and nonmembership function that lies within an interval. In this section, the conversion method is explained and illustrated through examples. Supporting statements are proved.

3.1 Steps Involved in the Method of Conversion

In this section, the steps involving the proposed conversion of T2IFSs into IVIFSs are explained.

Let $\tilde{T} = \{ (x, (\mu_{\tilde{T}}, g_x(\mu_{\tilde{T}}), \vartheta_{\tilde{T}}, h_x(\vartheta_{\tilde{T}}))) \} / x \in X, \mu_{\tilde{T}} \in [0, 1], \vartheta_{\tilde{T}} \in [0, 1] \}$ be a T2IFS, then it is converted into an IVIFS as follows:

Step 1: Identify the primary membership function $\mu_{\tilde{T}}$, nonmembership function $\vartheta_{\tilde{T}}$, and the corresponding hesitancy $\pi_{\tilde{T}} = 1 - \mu_{\tilde{T}} - \vartheta_{\tilde{T}}$ [1].

Similarly identify the secondary membership function $g_x(\mu_{\tilde{T}})$, nonmembership function $h_x(\vartheta_{\tilde{T}})$, and the corresponding hesitancy $\varpi_{\tilde{T}} = 1 - g_x(\mu_{\tilde{T}}) - h_x(\vartheta_{\tilde{T}})$ [1].

Step 2: Construct an IVIFS from $(\mu_{\tilde{T}}, \vartheta_{\tilde{T}}, \pi_{\tilde{T}})$ as follows:

$$\text{IVIFS} = \left\{ \left(x, \left[\mu_{\tilde{T}} - \alpha_1 \pi_{\tilde{T}} \mu_{\tilde{T}}, \mu_{\tilde{T}} + \alpha_1 \pi_{\tilde{T}} \mu_{\tilde{T}} \right], \left[\vartheta_{\tilde{T}} - \beta_1 \pi_{\tilde{T}} \vartheta_{\tilde{T}}, \vartheta_{\tilde{T}} + \beta_1 \pi_{\tilde{T}} \vartheta_{\tilde{T}} \right] \right) / x \in X \right\} \quad (1)$$

where $\alpha_1 = \mu_{\tilde{T}} + 0.5\pi_{\tilde{T}}$ and $\beta_1 = \vartheta_{\tilde{T}} + 0.5\pi_{\tilde{T}}$.

Step 3: Construct the other IVIFS from $\left(g_x\left(\mu_T^\sim\right), h_x\left(\vartheta_T^\sim\right), \varpi_T^\sim\right)$ as follows:

$$\text{IVIFS} = \left(x, \left[g_x\left(\mu_T^\sim\right) - \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right), g_x\left(\mu_T^\sim\right) + \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right) \right], \right. \\ \left. \left[h_x\left(\vartheta_T^\sim\right) - \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right), h_x\left(\vartheta_T^\sim\right) + \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right) \right] \right) / x \in X \quad (2)$$

where $\alpha_2 = g_x\left(\mu_T^\sim\right) + 0.5\varpi_T^\sim$ and $\beta_2 = h_x\left(\vartheta_T^\sim\right) + 0.5\varpi_T^\sim$.

Step 4: The required IVIFS of \tilde{T} is the average of both IVIFS obtained in Step 2 and Step 3.

$$\text{IVIFS } \tilde{T} = \left(x, \left[\mu_T^{\sim L}(x), \mu_T^{\sim U}(x) \right], \left[\vartheta_T^{\sim L}(x), \vartheta_T^{\sim U}(x) \right] \right) / x \in X \\ = \left\{ \left(x, \left[\frac{\left(\mu_T^\sim - \alpha_1 \pi_T^\sim \mu_T^\sim + g_x\left(\mu_T^\sim\right) - \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right) \right)}{2}, \frac{\left(\mu_T^\sim + \alpha_1 \pi_T^\sim \mu_T^\sim + g_x\left(\mu_T^\sim\right) + \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right) \right)}{2} \right], \right. \right. \\ \left. \left. \left[\frac{\left(\vartheta_T^\sim - \beta_1 \pi_T^\sim \vartheta_T^\sim + h_x\left(\vartheta_T^\sim\right) - \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right) \right)}{2}, \frac{\left(\vartheta_T^\sim + \beta_1 \pi_T^\sim \vartheta_T^\sim + h_x\left(\vartheta_T^\sim\right) + \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right) \right)}{2} \right] \right) \right\} / x \in X \quad (3)$$

where

$$\mu_T^{\sim L}(x) = \frac{\left(\mu_T^\sim - \alpha_1 \pi_T^\sim \mu_T^\sim + g_x\left(\mu_T^\sim\right) - \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right) \right)}{2}, \\ \mu_T^{\sim U}(x) = \frac{\left(\mu_T^\sim + \alpha_1 \pi_T^\sim \mu_T^\sim + g_x\left(\mu_T^\sim\right) + \alpha_2 \varpi_T^\sim g_x\left(\mu_T^\sim\right) \right)}{2} \\ \vartheta_T^{\sim L}(x) = \frac{\left(\vartheta_T^\sim - \beta_1 \pi_T^\sim \vartheta_T^\sim + h_x\left(\vartheta_T^\sim\right) - \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right) \right)}{2} \text{ and} \\ \vartheta_T^{\sim U}(x) = \frac{\left(\vartheta_T^\sim + \beta_1 \pi_T^\sim \vartheta_T^\sim + h_x\left(\vartheta_T^\sim\right) + \beta_2 \varpi_T^\sim h_x\left(\vartheta_T^\sim\right) \right)}{2} \quad (4)$$

Remark 3.1 The values $\alpha_1 = \mu_T^\sim + 0.5\pi_T^\sim$ and $\beta_1 = \vartheta_T^\sim + 0.5\pi_T^\sim$ are such that, $\alpha_1 + \beta_1 = 1$.

Proof

$$\alpha_1 + \beta_1 = \mu_T^\sim + 0.5\pi_T^\sim + \vartheta_T^\sim + 0.5\pi_T^\sim = \mu_T^\sim + \vartheta_T^\sim + \pi_T^\sim = 1.$$

Remark 3.2 The values $\alpha_2 = g_x\left(\mu_T^\sim\right) + 0.5\varpi_T^\sim$ and $\beta_2 = h_x\left(\vartheta_T^\sim\right) + 0.5\varpi_T^\sim$ are such that $\alpha_2 + \beta_2 = 1$.

Proof Follows as Remark 3.1 proof.

Theorem 3.1

Equation (3) is an IVIFSs for all T2IFSs.

1. $\left[\mu_T^{\sim L}(x), \mu_T^{\sim U}(x)\right] \subseteq [0, 1]$ and $\left[\vartheta_T^{\sim L}(x), \vartheta_T^{\sim U}(x)\right] \subseteq [0, 1]$
2. $\mu_T^{\sim U}(x) + \vartheta_T^{\sim U}(x) \leq 1$

Proof

$$1. \mu_T^{\sim L}(x) = \frac{(\mu_T^{\sim} - \alpha_1 \pi_T^{\sim} \mu_T^{\sim} + g_x(\mu_T^{\sim}) - \alpha_2 \varpi_T^{\sim} g_x(\mu_T^{\sim}))}{2},$$

The lower limit of $\mu_T^{\sim L}(x)$ can be gained by minimizing the membership grades

$$\lim_{\mu_T^{\sim}, g_x(\mu_T^{\sim}) \rightarrow 0} \frac{(\mu_T^{\sim} - \alpha_1 \pi_T^{\sim} \mu_T^{\sim} + g_x(\mu_T^{\sim}) - \alpha_2 \varpi_T^{\sim} g_x(\mu_T^{\sim}))}{2} = 0$$

The upper limit of $\mu_T^{\sim L}(x)$ can be gained by maximizing the membership grades

$$\lim_{\substack{\mu_T^{\sim}, g_x(\mu_T^{\sim}) \rightarrow 1 \\ \pi_T^{\sim}, \varpi_T^{\sim} \rightarrow 0}} \frac{(\mu_T^{\sim} - \alpha_1 \pi_T^{\sim} \mu_T^{\sim} + g_x(\mu_T^{\sim}) - \alpha_2 \varpi_T^{\sim} g_x(\mu_T^{\sim}))}{2} = 1$$

Therefore, we have $0 \leq \mu_T^{\sim L}(x) \leq 1$. Similarly, $0 \leq \mu_T^{\sim U}(x) \leq 1$.

Then, $\left[\mu_T^{\sim L}(x), \mu_T^{\sim U}(x)\right] \subseteq [0, 1]$.

Thus, $\left[\mu_T^{\sim L}(x), \mu_T^{\sim U}(x)\right] \subseteq [0, 1]$ and $\left[\vartheta_T^{\sim L}(x), \vartheta_T^{\sim U}(x)\right] \subseteq [0, 1]$ are satisfied.

$$2. \mu_T^{\sim U}(x) + \vartheta_T^{\sim U}(x)$$

$$\begin{aligned} &= \frac{(\mu_T^{\sim} - \alpha_1 \pi_T^{\sim} \mu_T^{\sim} + g_x(\mu_T^{\sim}) - \alpha_2 \varpi_T^{\sim} g_x(\mu_T^{\sim}))}{2} + \frac{(\vartheta_T^{\sim} + \beta_1 \pi_T^{\sim} \vartheta_T^{\sim} + h_x(\vartheta_T^{\sim}) + \beta_2 \varpi_T^{\sim} h_x(\vartheta_T^{\sim}))}{2} \\ &= \frac{\mu_T^{\sim} + g_x(\mu_T^{\sim}) + \vartheta_T^{\sim} + h_x(\vartheta_T^{\sim}) + (\alpha_1 \mu_T^{\sim} + \beta_1 \vartheta_T^{\sim}) \pi_T^{\sim} + (\alpha_2 g_x(\mu_T^{\sim}) + \beta_2 h_x(\vartheta_T^{\sim})) \varpi_T^{\sim}}{2} \\ &\leq \frac{\mu_T^{\sim} + g_x(\mu_T^{\sim}) + \vartheta_T^{\sim} + h_x(\vartheta_T^{\sim}) + \pi_T^{\sim} + \varpi_T^{\sim}}{2} \quad (\text{By Remarks 3.1 and 3.2}) \\ &= \frac{(\mu_T^{\sim} + \vartheta_T^{\sim} + \pi_T^{\sim}) + (g_x(\mu_T^{\sim}) + h_x(\vartheta_T^{\sim}) + \varpi_T^{\sim})}{2} = \frac{1+1}{2} = 1 \end{aligned}$$

Therefore, $\mu_T^{\sim U}(x) + \vartheta_T^{\sim U}(x) \leq 1$

Thus, Eq. (3) is an IVIFS.

The flow of the proposed conversion process is shown in Fig. 1.

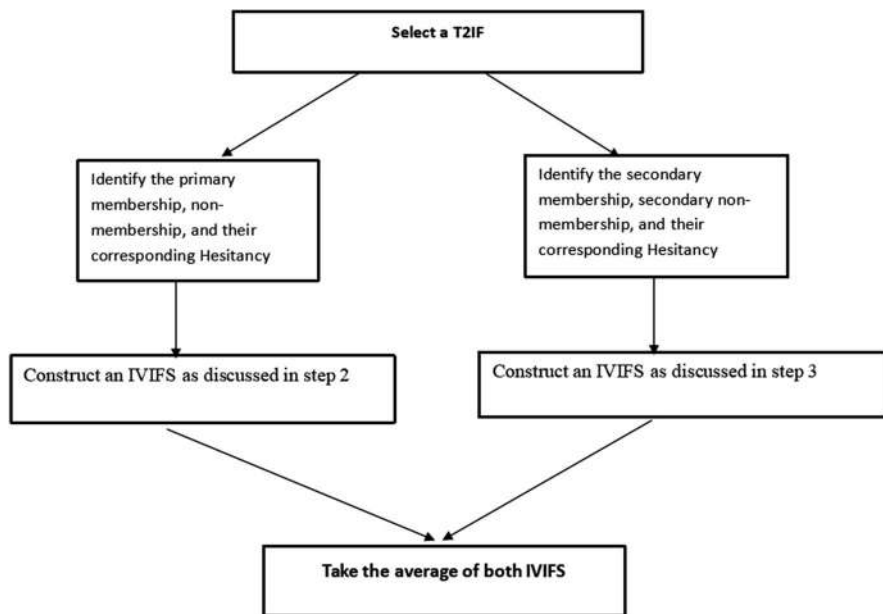


Fig. 1 Flow of the proposed conversion method (T2IFS to IVIFS)

3.2 Numerical Examples

In this section, we have taken some random examples of T2IFS and the proposed conversion is explained.

Example 1 Let X be a universal set, consider a T2IFS of $x \in X$ as $\tilde{S} = \{x, (0.7, 0.6, 0.2, 0.2) / x \in X\}$.

Step 1: The primary membership $\mu_{\tilde{S}}$ and nonmembership grades of \tilde{S} are $(0.7, 0.2)$, the corresponding hesitancy is $\pi_{\tilde{S}} = 0.1$.

Similarly, the secondary membership and nonmembership grades of \tilde{S} are $(0.6, 0.2)$, the corresponding hesitancy is $\varpi_{\tilde{S}} = 0.2$.

Step 2: Using Eq. (1), calculate α_1 and β_1 .

$$\alpha_1 = \mu_{\tilde{T}} + 0.5\pi_{\tilde{T}} = 0.7 + (0.5 * 0.1) = 0.75$$

$$\beta_1 = \varpi_{\tilde{T}} + 0.5\pi_{\tilde{T}} = 0.2 + (0.5 * 0.1) = 0.25$$

Therefore, $\alpha_1 = 0.75$ and $\beta_1 = 0.25$

Calculate the IVIFS corresponding to (0.7, 0.2, 0.1) using Eq. (1)

$$\begin{aligned} & \left(\left[\mu_T^\sim - \alpha_1 \pi_T^\sim \mu_T^\sim, \mu_T^\sim + \alpha_1 \pi_T^\sim \mu_T^\sim \right], \left[\vartheta_T^\sim - \beta_1 \pi_T^\sim \vartheta_T^\sim, \vartheta_T^\sim + \beta_1 \pi_T^\sim \vartheta_T^\sim \right] \right) \\ &= ([0.7 - (0.75 * 0.1 * 0.7), 0.7 + (0.75 * 0.1 * 0.7)], \\ & \quad [0.2 - (0.25 * 0.1 * 0.2), 0.2 + (0.25 * 0.1 * 0.2)]) \\ &= ([0.648, 0.753], [0.195, 0.205]) \end{aligned}$$

Step 3: Using Eq. (2), calculate α_2 and β_2 .

We get, $\alpha_2 = 0.7$ and $\beta_2 = 0.3$

Then calculate the IVIFS corresponding to (0.6, 0.2, 0.2) using Eq. (2)

The obtained IVIFS is ([0.52, 0.68], [0.18, 0.212])

Step 4: The required IVIFS obtained using Eq. (3) is

$$IVIFS = (x, ([0.58, 0.72], [0.19, 0.20]))$$

Example 2 This example is about the conversion of ideal (best) T2IFS to ideal IVIFS.

Consider the ideal (best) T2IFS $\tilde{I} = \{(x, (1, 1, 0, 0)) / x \in X\}$.

Here, we get $\alpha_1 = 1$ and $\beta_1 = 0$ and the corresponding IVIFS is ([1, 1], [0, 0]).

Similarly, we get $\alpha_2 = 1$ and $\beta_2 = 0$ and the corresponding IVIFS is ([1, 1], [0, 0]).

Then, the required IVIFS is obtained as ([1, 1], [0, 0]).

Hence, by the proposed conversion, the ideal T2IFS has been converted into ideal IVIFS.

Example 3 This example is about the conversion of null (worst) T2IFS to ideal IVIFS.

Consider the null (worst) T2IFS $\tilde{N} = \{(x, (0, 0, 1, 1)) / x \in X\}$.

Here, we get $\alpha_1 = 0$ and $\beta_1 = 1$ and the corresponding IVIFS is ([0, 0], [1, 1]).

Similarly, we get $\alpha_2 = 0$ and $\beta_2 = 1$ and the corresponding IVIFS is ([0, 0], [1, 1]).

Then, the required IVIFS is obtained as ([0, 0], [1, 1]).

Hence, by proposed conversion, the null T2IFS has been converted into null IVIFS.

Thus, the conversion is validated by examining the cases of the best and worst sets of the sets.

4 Implementation of Proposed Conversion in MCDM

In this section, we implement the proposed conversion in a group decision-making problem. The study is primarily concentrated on the results obtained when studied simultaneously in both T2IF and IVIF form.

4.1 Solving MCDM Problem Using Proposed Conversion

Assume that the decision-maker provides the performance ratings of each alternative in terms of a T2IF decision matrix. Let \tilde{p}_{ij}^l represents the performance of the alternative A_s , $s = 1, 2, \dots, n$ for criteria Y_t , $t = 1, 2, \dots, m$ with respect to decision-makers dm_l , $l = 1, 2, \dots, r$. The weight vector of decision-makers is represented by $(w_{dm_1}, w_{dm_2}, \dots, w_{dm_r})$ where $w_l \geq 0$, $l = 1, 2, \dots, r$, and $\sum_{l=1}^r w_l = 1$, and the weight vector of criterion is represented by $(w_{y_1}, w_{y_2}, \dots, w_{y_m})$ where $w_{y_t} \geq 0$, $t = 1, 2, \dots, m$ and $\sum_{t=1}^m w_{y_t} = 1$.

The decision of a multicriteria problem is taken through the following steps.

Step 1: Convert the T2IF decision matrix into IVIF decision matrix as stated in Sect. 3.1. and Fig. 1.

Step 2: Compute the distance matrix by finding the distance of each alternative from ideal IVIF $d_J \left(\tilde{p}_{ij}^l, \tilde{I} \right)$ using Definition 7.

Step 3: Compute the aggregated weighted decision matrix for each alternative.

$$\begin{aligned} d_1^l &= w_{y_1} \tilde{p}_{11}^l + w_{y_2} \tilde{p}_{12}^l + w_{y_3} \tilde{p}_{13}^l + \dots + w_{y_m} \tilde{p}_{1m}^l \\ d_2^l &= w_{y_1} \tilde{p}_{21}^l + w_{y_2} \tilde{p}_{22}^l + w_{y_3} \tilde{p}_{23}^l + \dots + w_{y_m} \tilde{p}_{2m}^l \\ &\vdots \\ d_s^l &= w_{y_1} \tilde{p}_{s1}^l + w_{y_2} \tilde{p}_{s2}^l + w_{y_3} \tilde{p}_{s3}^l + \dots + w_{y_m} \tilde{p}_{sm}^l \end{aligned}$$

where d_s^l represents the element in the decision matrix corresponding to decision-maker l to the alternative s .

Step 4: Calculate the overall assessment of alternatives $OA(A_s)$.

$$\begin{aligned} OA(A_1) &= w_{dm_1} (d_1^1) + w_{dm_2} (d_1^2) + \dots + w_{dm_r} (d_1^r) \\ OA(A_2) &= w_{dm_1} (d_2^1) + w_{dm_2} (d_2^2) + \dots + w_{dm_r} (d_2^r) \\ &\vdots \\ OA(A_s) &= w_{dm_1} (d_s^1) + w_{dm_2} (d_s^2) + \dots + w_{dm_r} (d_s^r) \end{aligned}$$

Step 5: Choose the best alternative by ranking the overall assessment of each alternative obtained in Step 4. The smallest distance from ideal set indicates the best alternative.

Step 1 is omitted when the problem is solved in T2IF environment, and the distance calculation and ranking are followed as per T2IFS definitions given in Sect. 2.

4.2 Example

In this section, an example from Singh [5] is considered for the study. A T2IF decision-making problem in which an investor wants to choose the best company in the market to invest in has been taken from Ref. [5]. Three decision-makers, whose weight vector is (0.40, 0.35, 0.25), assessed the five alternatives:

- (i) A_1 is car company
- (ii) A_2 is food company
- (iii) A_3 is computer company
- (iv) A_4 is car company
- (v) A_5 is tire company

Under four criteria:

- Risk analysis (C_1)
- Growth analysis (C_2)
- Environmental impact analysis (C_3)
- Available space (C_4)

The weight vector of criterion is given as (0.35, 0.30, 0.20, 0.15). The assessment of each alternative corresponding to each criterion s given by the expert in the form of T2IFS is presented in Table 1.

Table 1 Assessment of alternatives by decision-makers in T2IFSs

		DM 1				DM2				DM3			
		PMF	SMF	PNMF	SNMF	PMF	SMF	PNMF	SNMF	PMF	SMF	PNMF	SNMF
C_1	A_1	1	0.7	0	0.2	0.9	1	0.1	0	0.9	0.4	0.1	0.5
	A_2	0.7	0.9	0.1	0.1	0.9	0.7	0	0.1	0.2	0.5	0.7	0.4
	A_3	0.5	0.4	0.4	0.5	0.9	0.5	0.1	0.4	0.2	0.4	0.7	0.5
	A_4	0.7	0.5	0.2	0.4	0.9	0.5	0.1	0.4	0.2	0	0.7	1
	A_5	0.7	0.5	0.2	0.4	0.7	0.4	0.2	0.5	0.9	0.7	0	0.1
C_2	A_1	1	0.9	0	0.1	0.9	0.4	0.1	0.5	0.9	0.7	0.1	0.2
	A_2	0.7	0.7	0.4	0.4	0.9	0.7	0.1	0.2	0.2	0.5	0.7	0.4
	A_3	0.9	1	0.1	0	0.7	0.4	0.2	0.5	0.4	0.7	0.5	0.2
	A_4	0.9	0.7	0.1	0.2	0.7	0.4	0.2	0.5	0.4	0.5	0.5	0.4
	A_5	1	0.9	0	0.1	0.5	0.5	0.4	0.4	0.4	0.7	0.5	0.2
C_3	A_1	1	0.7	0	0.2	0.9	1	0.1	0	0.9	0.9	0.1	0.1
	A_2	1	0.7	0	0.2	0.9	1	0.1	0	1	0.9	0	0.1
	A_3	1	0.9	0	0.1	0.9	0.9	0.1	0.1	0.4	0.7	0.5	0.2
	A_4	1	0.9	0	0.1	0.9	0.5	0.1	0.4	0.4	0.7	0.5	0.2
	A_5	0.9	0.4	0.1	0.5	0.7	0.4	0.2	0.5	0.4	0.7	0.5	0.2
C_4	A_1	1	0.9	0	0.1	0.9	1	0.1	0	0.9	0.5	0.1	0.4
	A_2	1	0.9	0	0.1	0.9	1	0.1	0	1	0.7	0	0.2
	A_3	1	0.4	0	0.5	0.7	0.9	0.2	0.1	0.4	0.4	0.5	0.5
	A_4	0.7	0.5	0.2	0.4	0.9	0.5	0.1	0.4	0.5	0.7	0.4	0.2
	A_5	0.9	0.7	0.1	0.4	0.7	0.4	0.2	0.5	0.5	0.7	0.4	0.2

Then, the decision of the problem using conversion and IVIF MCDM modeling approach is taken through the following steps.

Step 1: Convert the T2IF decision matrix into IVIF decision matrix using proposed method. The obtained IVIF decision matrix is given in Table 2.

Step 2: The distance matrix is computed and given in Table 3.

Step 3: The weighted decision matrix is calculated and presented in Table 4.

Step 4: The overall assessment of alternatives is given in Table 5.

Table 2 IVIF decision matrix after conversion

		DM 1				DM2				DM3			
		LMF	UMF	LNMF	UNMF	LMF	UMF	LNMF	UNMF	LMF	UMF	LNMF	UNMF
C ₁	A ₁	0.82	0.88	0.1	0.1	0.82	0.88	0.1	0.1	0.64	0.66	0.29	0.31
	A ₂	0.74	0.86	0.1	0.1	0.74	0.86	0.1	0.1	0.33	0.37	0.51	0.59
	A ₃	0.43	0.47	0.43	0.47	0.43	0.47	0.43	0.47	0.29	0.31	0.56	0.64
	A ₄	0.56	0.64	0.29	0.31	0.56	0.64	0.29	0.31	0.1	0.1	0.82	0.88
	A ₅	0.56	0.64	0.29	0.31	0.56	0.64	0.29	0.31	0.7	0.9	0.05	0.05
C ₂	A ₁	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.77	0.83	0.15	0.15
	A ₂	0.65	0.75	0.2	0.21	0.65	0.75	0.2	0.21	0.33	0.37	0.51	0.59
	A ₃	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.51	0.59	0.33	0.37
	A ₄	0.73	0.87	0.1	0.1	0.73	0.87	0.1	0.1	0.43	0.47	0.43	0.47
	A ₅	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.51	0.59	0.33	0.37
C ₃	A ₁	0.82	0.88	0.1	0.1	0.82	0.88	0.1	0.1	0.9	0.9	0.1	0.1
	A ₂	0.82	0.88	0.1	0.1	0.82	0.88	0.1	0.1	0.95	0.95	0.05	0.05
	A ₃	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.51	0.59	0.33	0.37
	A ₄	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.51	0.59	0.33	0.37
	A ₅	0.64	0.66	0.29	0.31	0.64	0.66	0.29	0.31	0.51	0.59	0.33	0.37
C ₄	A ₁	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.69	0.71	0.24	0.26
	A ₂	0.95	0.95	0.05	0.05	0.95	0.95	0.05	0.05	0.82	0.88	0.1	0.1
	A ₃	0.69	0.71	0.24	0.26	0.69	0.71	0.24	0.26	0.38	0.42	0.47	0.53
	A ₄	0.56	0.64	0.29	0.31	0.56	0.64	0.29	0.31	0.56	0.64	0.29	0.31
	A ₅	0.8	0.8	0.2	0.2	0.8	0.8	0.2	0.2	0.56	0.64	0.29	0.31

Table 3 Distance matrix

	DM1				DM2				DM3			
	C ₁	C ₂	C ₃	C ₄	C ₁	C ₂	C ₃	C ₄	C ₁	C ₂	C ₃	C ₄
A ₁	0.124	0.05	0.124	0.05	0.05	0.323	0.05	0.05	0.323	0.174	0.1	0.274
A ₂	0.146	0.245	0.124	0.05	0.117	0.174	0.05	0.05	0.590	0.590	0.05	0.124
A ₃	0.493	0.05	0.05	0.274	0.274	0.394	0.1	0.174	0.638	0.394	0.394	0.542
A ₄	0.344	0.145	0.05	0.344	0.274	0.394	0.274	0.274	0.865	0.493	0.394	0.344
A ₅	0.344	0.05	0.323	0.2	0.394	0.443	0.394	0.394	0.117	0.394	0.394	0.344

Table 4 Weighted decision matrix

	DM1	DM2	DM3
A ₁	0.0908	0.1321	0.2267
A ₂	0.1573	0.1107	0.4121
A ₃	0.2386	0.2603	0.5018
A ₄	0.2262	0.310	0.5814
A ₅	0.2305	0.4091	0.2898

Table 5 Overall performance

Alternatives	Overall performance
A ₁	0.139
A ₂	0.204
A ₃	0.312
A ₄	0.344
A ₅	0.307

Step 5: The ordering is given as $A_1 > A_2 > A_3 > A_5 > A_4$. Thus, the best alternative is A_1 .

4.3 Result Analysis in T2IF and IVIF Domains

From Ref. [5], it is found that the ordering of alternatives is $A_1 > A_3 > A_2 > A_5 > A_4$. It is observed that the result obtained after conversion is similar for best and worst alternatives even though the complete ordering is not the same. The best company for investing in is A_1 . The best alternative to the problem is not altered after converting the decision matrix of T2IFSs into IVIFSs. Therefore, the proposed conversion method can be suitably utilized for solving decision-making problems. The computational process for decision-making in T2IF environment is complex. Hence, using the proposed conversion, the complexity can be reduced to a trouble-free process.

5 Conclusion

In the context of available generalizations of fuzzy sets, it is natural to contemplate the correlation between them and methods to convert one type of fuzzy set into another. Thus, in this chapter, we are motivated to discuss a conversion method from T2IFS, which has great expressive power and appealing conceptually into IVIFS is intuitive and trouble-free in decision-making. In this framework, hesitancy is actively involved in the conversion. Furthermore, a random numerical example of T2IFS is taken to demonstrate the conversion. Moreover, the basic ideal set (best) and null set (worst) of T2IFSs are considered for the conversion and obtained the

basic ideal set and null set of IVIFS, validating the conversion process. T2IFSs have a significant benefit in handling imprecise and inadequate information. However, their wide applications are getting less attention. Though the methods of handling MCDM problems in the framework of T2IFS are inadequate, we can use abundant methods of IVIFS to obtain the decision through this conversion.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection, and data analysis were done by V. Anusha. The first draft of the manuscript was written by N. Annapurna. Review and editing were done by Sireesha V. All authors read and approved the final manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

References

1. Zadeh LA, Fuzzy sets Information and control, (1965), Vol. 8, 338–353.
2. Atanassov KT, Intuitionistic Fuzzy Sets, Fuzzy Sets Syst. (1986) Vol. 20, 87–96, [https://doi.org/10.1016/S0165-0114\(86\)80034-3](https://doi.org/10.1016/S0165-0114(86)80034-3).
3. Atanassov and Gargov, Interval-valued Intuitionistic fuzzy sets, Fuzzy Sets Syst. (1989) Vol. 31, 343–349.
4. Bui Cong Cuong, Some operations on Type-2 Intuitionistic Fuzzy sets, J. Comput. Sci. Cybern. (2012) Vol. 28, 274–283, <https://doi.org/10.15625/1813-9663/28/3/2607>.
5. Singh and Garg, Distance measures between Type-2 Intuitionistic Fuzzy sets and their Application to Multi-Criteria Decision-Making process, Appl. Intell. (2016), Vol. 46, 788–799. <https://doi.org/10.1007/s10489-016-0869-9>.
6. Dipak Kumar Jana, Novel arithmetic operations on Type-2 Intuitionistic fuzzy and its Applications to Transportation problem, Pac. Sci. Rev. A: Nat. Sci. Engg. (2016), Vol. 18, Issue 3, <https://doi.org/10.1016/j.pusra.2016.09.008>.
7. Singh and Garg, Algorithm for solving Group Decision-making problems based on the similarity measures under Type 2 Intuitionistic Fuzzy sets environment, Soft Comput. (2020), Vol. 24 (2), <https://doi.org/10.1007/s00500-019-04359-8>.
8. Kamesh Kumar and Pandey D, Discussion on the switching between Type-2 fuzzy sets and Intuitionistic fuzzy sets: An application in medical diagnosis, J. Inf. Opt. Sci. (2018), Vol. 39 (2), 427–444, <https://doi.org/10.1080/02522667.2017.1411014>.
9. Atanassov, Operators over interval valued Intuitionistic fuzzy sets, Fuzzy Sets Syst. (1994), Vol. 64, 159–174.
10. Tao Z, Zhu, Zhaou, Liu and Chen, Multi-attribute decision making with Pythagorean fuzzy sets via conversions to Intuitionistic fuzzy sets and ORESTE method, J Control. Decis. (2020), <https://doi.org/10.1080/23307706.2020.1830445>.
11. Bustince and Burillo, Vague sets are Intuitionistic fuzzy set, Fuzzy Sets Syst. (1996), Vol. 79, 403–405, [https://doi.org/10.1016/0165-0114\(95\)00154-9](https://doi.org/10.1016/0165-0114(95)00154-9).
12. Deschrijver G and Kerre E, On the relationship between some extensions of fuzzy set theory, Fuzzy Sets and Syst. (2003), Vol. 133 (2), 227–235, [https://doi.org/10.1016/s0165-0114\(02\)00127-6](https://doi.org/10.1016/s0165-0114(02)00127-6).
13. Anusha V and Sireesha V, Ranking Interval Valued Intuitionistic Fuzzy Sets by a New Distance Measure, Adv. Math., Sci. J. (2021), Vol. 10 (3), 1249–1258, <https://doi.org/10.37418/amsj.10.3.13>.

14. Manonmani A and Suganya M, An Examination in IFS and IVIFS based on Various Distance Measure, Int. J. Recent Technol. Eng. (2019), Vol. 8 (3S3), 15–18, <https://doi.org/10.35940/ijrte.C1024.1183S31>.
15. Klir, George J. and Yuan, Bo Fuzzy Sets and Fuzzy Logic Theory and Applications, (1995).

A Framework for Secure Database and Similarity Comparison in Android



Aishwarya Phadtare and Kishor Mane

Abstract There is an increase in utilization of mobile devices in today's time. As a result, there is significant development in mobile technologies. User data stored on mobile devices is not in encrypted format. All the user content can easily be manipulated if any person gets access to the database. Thus, it is necessary to protect user information in the database to shield it from various security threats. The chapter demonstrates a secured database framework that performs encryption and prevents attacks. The proposed system presents an architecture for creating an environment that secures database on the android operating system. The secure database framework manages the generation of keys, implements encryption/decryption on data, and then stores encrypted data in the database. The framework generates a key for data encryption and secures database operations by key decrypted after application authentication. The secure database framework also checks whether the newer application is a renewed version of the old one. Based on the metadata information of the application that is retrieved and compared by the secure database framework to verify the identity of the application by signature verification and similarity comparison. The secure database framework tested for SQL injection and colluding attacks. The SQL injection is prevented by specifying validation items and colluding attack is prevented by detecting similar applications by similarity comparison based on the similarity score. Both attacks are tested on the framework by checking various conditions and parameters.

Keywords Android OS · Secure database framework · Similarity comparison · Encryption · Colluding attack · SQL injection attack

A. Phadtare (✉) · K. Mane

D.Y. Patil College of Engineering and Technology, Kolhapur, Maharashtra, India

1 Introduction

Mobile devices are handy and used to perform daily activities. Due to digitalization, the use of smartphones has increased. All user data such as contact information, important notes, and documents are stored on mobile devices. Mobiles are small size devices, so these are easy to carry anywhere. The use of smartphones has increased, leading to growth in mobile technology. Mobile devices are used for communication and performing online activities such as shopping, banking, data transfer, and so on. As a result, mobile devices are important in today's time. There are different operating systems used in mobile devices but the most common is Android. In Android, there are two types of applications: one that comes in-built with the device and the other is third-party applications. SQLite database is used by Android devices to store data. SQLite database [1] is similar to the relational database that needs not to be configured as it comes built-in. In Android, no connection is needed to be created as the database is locally stored on the device. As only one user uses the device, the database is a single user and does not support multiple users. Thus, if anyone gets access to the data stored in the database, the user privacy is invaded. While the data in the mobile database is vulnerable as it is not encrypted, protection of the database is required in Android [2, 3]. Attacks such as colluding attacks, SQL injection, loss of device, network-related attacks are common due to which user data can be acquired and manipulated [4]. In the designed framework, the app and the database are separated and the data is encrypted and then stored in the database. Each database operation is performed after a security check. The secured database framework generates the key for data encryption and stores the metadata of the applications for application authentication. The identity of the applications is verified by signature verification, and similarity comparison. The secure database framework prevents SQL injection attack by validation and colluding attack by similarity comparison to detect whether two applications are similar or not.

2 Literature Review

In Android, secure database architecture [5] focuses on splitting the database and the app. The uid is used to permit access in Android OS. In order to check if the app is updated or old version, similarity is checked. The time required for database operation is longer. The different SQL operations time is measured. The overhead due to the number of operations has been checked. The updated version is examined to determine if it is the original application's modified version. A new, updated version is installed when this is checked. Mobile hazards, such as the user losing the device or a malware attack, might compromise personal data saved on the device [6]. The author [7] discusses several risks to mobile devices, the significance of protecting user data on the device, and outlines the security of the storage component. Security restrictions have resulted from updating an old application

and adding a few new features to an already existing application [8]. Therefore, similarity must be examined. The framework [9] is made to find similarities at various phases. To identify fraudulent applications created to obtain user data, Android applications can be compared based on their graphical user interfaces [10]. Since plagiarism is widespread today, it is possible to use applications to commit it, usually to make money or gain unauthorized access. Based on text and visual components, these applications are evaluated. The applications are compared to see if they are the same or different using various functions [11, 12] to calculate scores to check similarities. The Android database makes use of the SQLite database. This database only has one user. SQL injection attacks are widespread in android [13] and concentrate on a protection mechanism that uses specific validations to prevent SQL injection attacks. When protecting web applications, the size of input queries is typically compared to the actual query. A defense model is created to guard against attacks. In this model, access is only granted after the IP address has been verified. After access is granted, several parameters are checked to determine whether an input is valid or not. If the string is correct, the page loads; if not, the user is prevented from accessing it. Fraudulent apps gain access to a number of privileges during installation. Android apps consequently interact with one another, which results in app collusion. There is a chance of a colluding attack, which must be avoided. In order to check more privileged programs, a system [14] was created. These programs are given authorization to share identities. The system is safeguarded and rules are set to help identify fraud activity. This chapter [15] discusses the inappropriate use of the permissions that the user grants to the application. Before installing, the application asks the user for permission. For example, when a certain picture editing application is to be installed, it requests access to the mobile device's photo collection, leading to the user occasionally giving the application permission without realizing it. As a result, the program has access to the data that users have on their devices. Such security flaws make it simple for attackers to attack the OS. Following are some drawbacks connected to the security of Android operating systems that have been discovered through the literature review:

- Data stored in android databases must be encrypted.
- Data encryption calls for a more effective encryption technique.
- Database operations must be performed in the shortest possible session time.
- The applications' identities must be verified.
- Colluding attacks and SQL injection attack should be avoided.

3 Overview of Secure Database Framework

In the proposed system architecture, the secure database framework encrypts the data and stores in the database and performs authentication checks for secure database operations. The framework checks the identity of the application by simi-

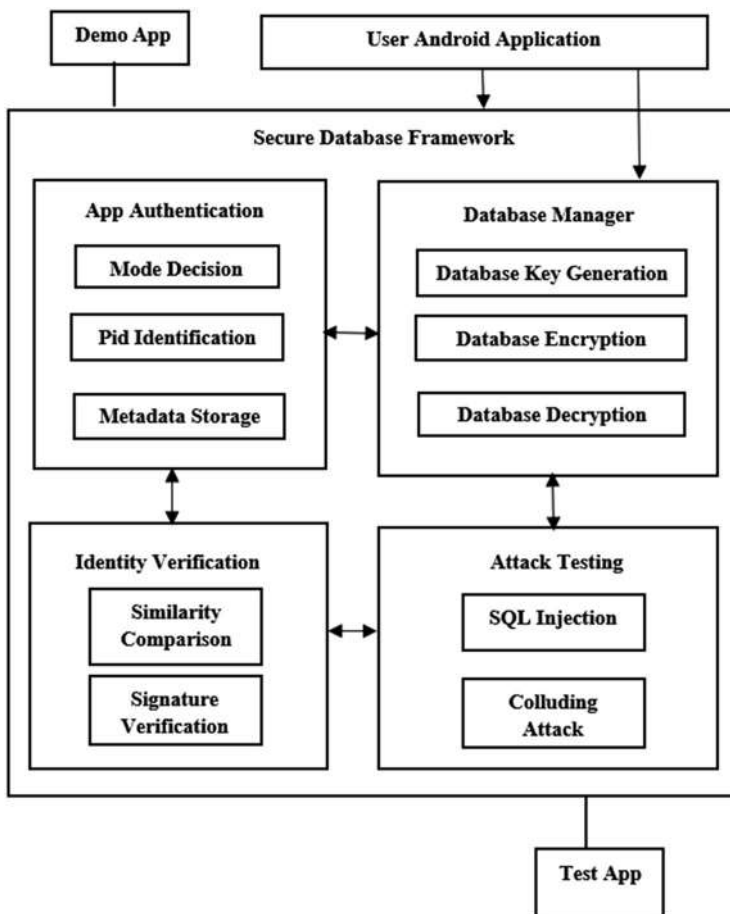


Fig. 1 System architecture of secure database framework

larity comparison along with attack tests and prevents attacks such as SQL injection and colluding attacks. Figure 1 shows the architecture for the secure database framework. The architecture of secure database framework consists of four modules application authentication, data encryption/decryption, identity verification, and attack testing.

1. *App authentication*: The secure database framework is built to provide a secure environment. On initial booting, the secured database framework generates key for authentication of data and database access operations [16, 17]. In app authentication, along with a uid, a one-to-one pairing is created between the application and its database. The secure database framework knows the process identification of the application thus determines files referred to by the process.

The secured database framework generates the metadata files, which is used to check authentication of apps.

2. *Database encryption and decryption:* The secure database generates a key on initial booting and this key is encrypted by a device-specific key. The database operations are also secured by the framework, and it generates the database key, which is encrypted by the secure database framework key. The AES-GCM algorithm [18] is used for encryption. This database key is generated on the first database operation and later used on each database operation. The secure database framework encrypts the data before storing it in the database.
3. *Identity verification:* The developer of android applications can share user_id between applications to share the data in the database. As a result, an attacker can get the main privileges of the applications [19, 20]. Hence, in signature verification, it is checked if the old and updated application is signed by the same key. In similarity comparison, two applications are compared to get a similarity score. The score is measured as 1 or 0. Applications are similar if score 1, else not similar. Hence, only genuine updated application of the old version is installed.
4. *Attack testing:* SQL database attacks are carried out by inserting code into the input variables, also some attackers insert destructive code into strings; therefore, validations are specified to prevent such attacks such as making no assumptions about the type, size, or content of data and checking the string variables' contents. It is necessary to confirm the input's size and value type. Applications are updated to provide new features and remove bugs from the old version. However, similar type of applications are designed by attackers to get privileges and to access database. This leads to colluding attack [21]. The secure database framework compares the metadata of the apps to check if the apps are similar or not; thus, preventing colluding attack. Figure 2 shows flowchart and pseudocode for secure database framework.

4 Result and Discussions

4.1 Experimental Setup

The proposed system is implemented and tested with following experimental setup as:

- The proposed system is developed using Android Studio.
- JDK 1.6 and above are used for implementation.
- The database used for data storage is the SQLite database.
- SQLite viewer is software to view data in the application database.

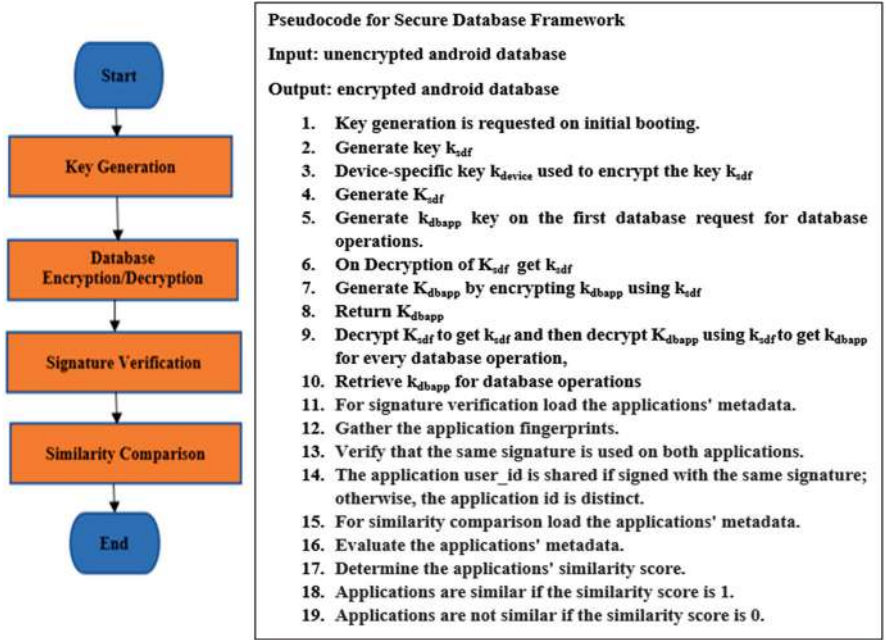


Fig. 2 Flowchart and pseudocode of secure database framework

Table 1 Time for database operation on plain data operation and encrypted data operation

Operation on database (MB)	Plain data time (Tp) (ms)	Encrypted data time on existing system using AES-CBC (ms)	Encrypted data time (Te) on proposed system using AES-GCM (ms)
5	236	239	238
20	237	241.5	240
50	240	243	242
100	246	250.7	250
200	255	262	260

4.2 Plain Data Operation Versus Encrypted Data Operation

The data in the database is encrypted and then stored in the database, so the time required for database operation on plain data and encrypted data is calculated. Table 1 shows the time required to perform database operations on plain and encrypted data using AES-CBC in existing system and AES-GCM in proposed system. The existing system used AES-CBC where data is EX-OR with plaintext, which can be known by attacker and hence is not that secure and prone to attack. The proposed system uses AES-GCM, which is more effective and secure as plain data is

Table 2 Average time on plain data versus encrypted data

Number of operation (<i>N</i>)	Database operations	Average time on plain data <i>Ts</i> (P) (ms)	Average time on encrypted data <i>Ts</i> (E) (ms)
1	CREATE	10	12.5
2	OPEN, INSERT	17.6	19
3	OPEN, SELECT, COMPARE	19	20.5

EX-OR with output from block cipher where the counter is maintained and current value of counter is sent through block cipher. The AES-GCM performs parallel computation; hence, it is slightly faster than AES-CBC and is more efficient and secure.

The database operations performed are if the database is already present open database else the database is created, insertion of user data and metadata of the applications, selecting the user data or metadata of applications and verifying the data by comparing, etc. The time to perform a database operation on plaintext data and encrypted data is measured. The time required is calculated in milliseconds (ms). Table 2 shows the average time required for database operations. The number of operations carried out in database sessions is considered where the database operations are open, create, insert, compare. The average time on plain data and encrypted data are calculated. First, the database will be created if it is already not created or else open. The metadata of the application will be inserted into the database for application authentication. To verify and compare the application, the metadata of the applications will be selected and then compared to determine its identity.

On database operation, the framework checks the authentication of the application and decrypts the data after security checks, so the time required is more. The percentage of time required for database operation on plain data is around 2% less than on encrypted data. Figure 3 shows the graphs of the processing time required for database operation on plaintext and encrypted data.

4.3 Similarity Comparison

Applications are updated to remove bugs. When bugs are removed, a modified version of the application is available; thus, there is a need to check whether an app is an updated version or the old version. This can be checked by using similarity measurements. If methods are similar, then we calculate it as 1 or else 0. The elapsed time for the similarity check will be checked based on the average time that was required to compare the similarity of two android applications. Figure 4 shows similarity comparison graph. Two applications are compared to calculate the similarity score. The data required for access authentication and database operation

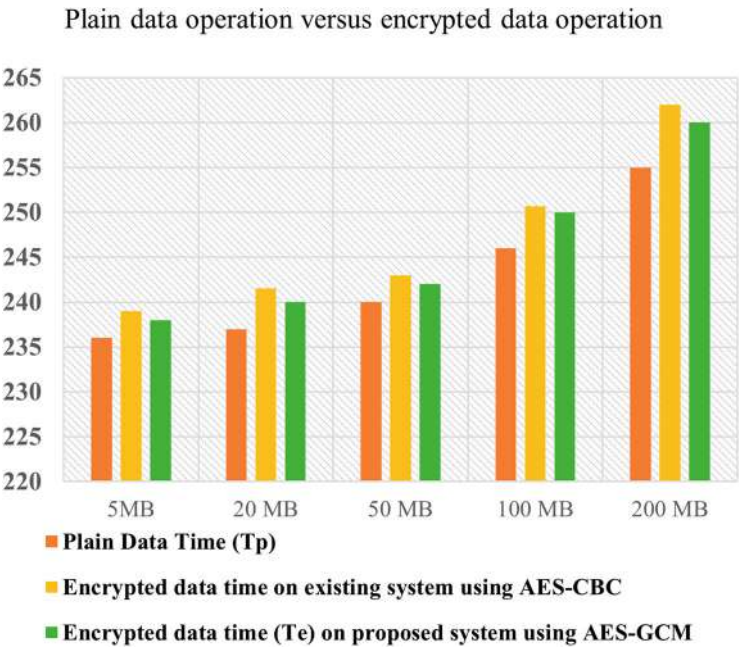


Fig. 3 Graph for processing time on plain data operation versus encrypted data operation

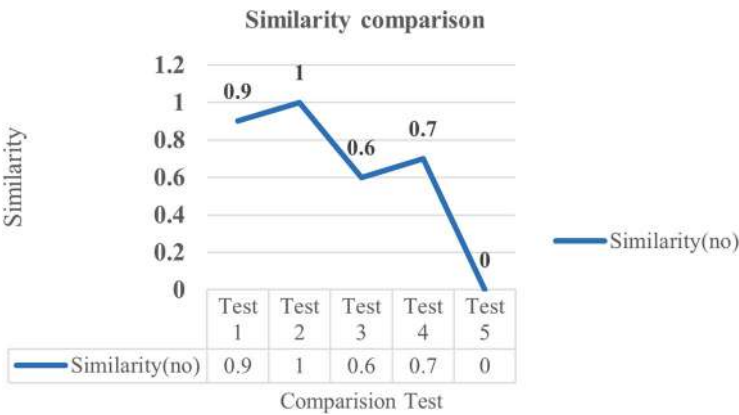


Fig. 4 Similarity comparison graph

is stored in the metadata storage. The secured database framework retrieves and stores the metadata information. The two applications' metadata fields are compared to get a similarity score. The score is less than 1 even if single parameter is different. Different tests cases are performed as follows:

- Test 1: Altering package and version
- Test 2: When the applications are similar
- Test 3: Altering only application fingerprint
- Test 4: Altering the cachesize, datasize
- Test 5: When the applications are dissimilar

5 Conclusion

In Android smartphones, user data is stored in the database in plaintext. Some attackers try to access user content to misuse the data or get excess privileges to make profit. Thus, there is need of protection of user's privacy. In the chapter, the design of the secure database framework has been implemented and tested for different attacks. The secure database framework performs application authentication, database encryption/decryption, identity verification, and attack testing. While the data is encrypted, there is slight overhead in database operation time, but the system is enhanced with minimum time overhead. The framework checks the identity of the application by signature verification and similarity comparison. Database attacks are common in android and the secure database framework prevents SQL injection attack and colluding attack. Thus, the secure database framework creates a safe environment for android user. In the future, the framework can be modified and designed to work on other platforms.

References

1. Mutti Simone, Enrico, Paraboschi, "SeSQLite: Security Enhanced SQLite: Mandatory Access Control for Android databases," In: 31st Annual Computer Security Applications Conference (2015).
2. Xin Li, Xudong, "Hack Android Application and Defence," In: 3rd International Conference on Computer Science and Network Technology (2013).
3. Min Jiangbo Guo, BaoShuai, Pengpeng, "The Design and Implementation of Security Defense System Based on Android," In: 8th IEEE International Conference on Software Engineering and Service Science (2017).
4. Bahman Rashidi and Carol Fung Virginia, "A Survey of Android Security Threats and Defenses," In: Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 6, number: 3, pp. 3–35 (2015).
5. J. H. Park, S. Yoo, I. S. Kim, and D. H. Lee, "Security Architecture for a Secure Database on Android," In: IEEE Transaction of multidisciplinary open access, vol. 6, pp. 11482–11501 (2018).
6. P. D. Meshram, Dr. R.C. Thool, "A Survey Paper on Vulnerabilities in Android OS and Security of Android Devices," In: IEEE Global Conference on Wireless Computing & Networking (GCWCN), Lonavala, pp. 174–178 (2014).

7. Poonguzhali P, Prajyot Dhanokar, M. K. Chaithanya, and Mahesh U. Patil, "Secure Storage of Data on Android Based Devices," In: IACSIT International Journal of Engineering and Technology, Vol. 8, No. 3 (2016).
8. Sujata Chakravarty, Ravi Kiran Verma, Santosh K, " Feature Selection and Evaluation of Permission-based Android Malware Detection," In: 4th International Conference on Trends in Electronics and Informatics (2020).
9. Tegawende F. Bissyand, Jacques Klein, " SimiDroid: Identifying and Explaining Similarities in Android Apps," In: IEEE Computer Society (2017)
10. Jiawei Zhu, Zhengang Wu, Zhi Guan, and Zhong Chen, "Appearance Similarity Evaluation for Android Applications," In: Seventh International Conference on Advanced Computational Intelligence (ICACI), Wuyi, pp. 323–328 (2015).
11. B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," In: Communication of the ACM, vol. 13, no. 7, pp. 422–426, Jul.(1970).
12. Y. Qiao, T. Li, and S. Chen, "Fast Bloom filters and their generalization," In: IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 1, pp. 93–103, Jan. (2014).
13. Li Qian, Zhenyuan Zhu, Jun Hu, and Shuying Liu, "Research of SQL injection attack and prevention technology," In: International Conference on Estimation, Detection, and Information Fusion (ICEDIF), Harbin, pp. 303–306 (2015).
14. Iman Kashefi, Maryam Kassiri, and Mazleena Salleh, "Preventing Collusion Attack in Android," In: The International Arab Journal of Information Technology, Vol. 12, No. 6A (2015).
15. Karthick S and Dr. Sumitra Binu, "Android Security Issues and Solutions," In: International Conference on Innovative Mechanisms for Industry Applications (2017).
16. T. Coolijmans, J. de Ruiter, and E. Poll, "Analysis of secure key storage solutions on Android," In: Proc. 4th ACM Workshop Security Privacy Smartphones Mobile Devices (2014).
17. Bacis, Simon, "AppPolicyModules: Mandatory Access Control for Third-Party Apps," In: 10th ACM Symposium on Information, Computer and Communications ACM (2015).
18. Alexander Uskov, Adam Byerly, Colleen Heinemann, "Advanced Encryption Standard Analysis with Multimedia Data on Intel AES-NI Architecture," In: International Journal of Computer Science and Applications, Technomathematics Research Foundation Vol. 13, No. 2, pp. 89–105 (2016).
19. Sen Chen, Feng We, Tianming Liu, "ATVHunter: Reliable Version Detection of Third-Party Libraries for Vulnerability Identification in Android Applications," In: ACM 43rd International Conference on Software Engineering, (2021).
20. Igor Zavalysyn, Nuno Santos, " Flowverine: Leveraging Dataflow Programming for Building Privacy-Sensitive Android Applications," In: 19th International Conference on Trust, Security and Privacy in Computing and Communications (2020).
21. Khokhlov, Leon Reznik, "Colluded Applications Vulnerabilities in Android Devices," In: IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing (2017).

A Comprehensive Review and a Conceptual Framework for Predicting the Position of the Mobile Sinks in Wireless Sensor Networks



Uma Maheswari Gali and Nagarjuna Karyemsetty

Abstract The purpose of this research is to review the existing literature on clustering protocols in wireless sensor networks (WSNs) and to analyze the methodologies used to anticipate the position of mobile sinks in WSNs, as well as their limitations. As they move in a traffic pattern and collect data from sensor nodes, movable sinks have become more common in WSNs. Accurate mobile sink location prediction is crucial for developing effective routing algorithms and optimizing the data collection process. This chapter proposes the construction of a framework with an upgraded LEACH (low-energy adaptive clustering hierarchy) algorithm to solve the constraints of the original LEACH algorithm, such as the assumption of a static network and the inability to handle mobile sinks. The suggested framework also addresses the challenges of shorter lifetime after a specific number of rounds, the stability period, and the emphasis on single mobile sink prediction. The importance of the suggested architecture for increasing the performance of WSNs in terms of data gathering and routing is emphasized in this chapter.

Keywords Wireless sensor networks · Clustering protocols · Mobile sinks · Position prediction · Routing algorithms · Data gathering · LEACH algorithm

1 Introduction

Clustering protocols have been extensively researched in wireless sensor networks (WSNs) to prolong network lifetime, reduce energy consumption, and enhance network scalability [1, 2]. However, with the emergence of mobile sinks, which

U. M. Gali

Research Scholar Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

N. Karyemsetty (✉)

Department of CSE, Koneru Lakshmaiah Foundation, Vaddeswaram, Andhra Pradesh, India
e-mail: nagarjunak@kluniversity.in

collect data by moving in a traffic pattern, existing clustering protocols face new challenges [3]. Accurate prediction of the position of mobile sinks is critical for optimizing the data gathering process and establishing effective routing algorithms [4, 5].

In this chapter, the literature reviews on clustering protocols in WSNs and analyzes the existing techniques for predicting the position of mobile sinks and their limitations. To emphasize the increasing prevalence of mobile sinks in WSNs and the significance of accurate mobile sink position prediction. It also present the development of a framework with an improved LEACH (low-energy adaptive clustering hierarchy) algorithm to address the limitations of the original LEACH algorithm, including the assumption of a static network and the inability to accommodate mobile sinks [6].

Deep learning has lately emerged as a potent method for tackling complicated issues across multiple areas [7–9]. Long short-term memory (LSTM) networks is considered as a procedure of the advanced recurrent neural network (RNN) and have been used successfully in a wide range of sequence prediction applications [10–12]. Because LSTM networks can manage long-standing dependences in serial data, they are well-suited to forecasting the location of mobile sinks in WSNs [13, 14].

To propose an enhanced LEACH algorithm with a deeper LSTM network with a regression layer in this research proposal, to forecast the position of mobile sinks moving in a traffic pattern. The suggested technique overcomes the constraints of the original LEACH algorithm by utilizing the capabilities of deep learning to precisely forecast the position of mobile sinks.

This framework aims to address the lower lifetime after a given number of rounds, the stability period, and the focus on single mobile sink prediction, which have been overlooked in previous studies. By addressing these issues, the proposed framework is expected to improve the performance of WSNs in terms of data gathering and routing.

2 Literature Review

Deng et al. [15] presented a primal–dual approach [15]. The approach can be computationally expensive, especially for large-scale WSNs. As the number of nodes in the network increases, the computational time and memory requirements of the primal–dual approach can become prohibitive.

Wen et al. [16] suggested a cooperative data collection algorithm (CDCA) for grouping sensor nodes and assigning a mobile sink for the various groups. However, it was observed that lifetime has been reduced after certain number of rounds.

In the chapter, Tabatabaei et al. [17] suggested a reliable routing protocol depending on clustering process as well as improved form of mobile synchronization in WSNs. This can result in increased energy consumption and a shorter lifespan for the network, particularly if the clustering and synchronization protocols

are not optimized or there are a large number of mobile sinks. In addition, accurate prediction of mobile sink locations can be difficult in practice, resulting in suboptimal routing decisions and potential network performance degradation.

Wang et al. [18] suggested a distance-aware routing algorithm. A distance-aware routing algorithm operates through three primary steps: neighbor discovery, route discovery, and data transmission. During neighbor discovery, each sensor node determines the proximity and identity of its neighbors. The node determines the most energy-efficient path to the destination node during route discovery. Due to factors such as noise, interference, and obstructions, accurate distance measurement is difficult to achieve in practice. Inaccurate measurement of distance can lead to suboptimal path selection and increased energy consumption.

Sethi [19] employs sinks node to be in mobility to decrease the attainable energy gap and in turn increase network life [19]. The use of multiple mobile sinks and route patterns could introduce additional overhead into the network, such as increased control traffic and routing messages.

Banimelhem et al. projected a process with principal component analysis (PCA) [20]. It works only on single mobile sink prediction. In a nutshell while PCA can be a powerful instrument for data analysis in WSNs, its constraints in terms of mobile sinks should be considered, particularly with regard to the assumption of stationary data, computational demands, and linear relationships.

Sharma et al. proposed an evolutionary strategy for WSNs based on clustering and an optimized invasive weed algorithm, which is modified in this study based on the advanced fuzzy modelling [21]. The algorithm's sluggish convergence rate prolongs the time required to reach the optimal solution.

3 Significance and Problem Statement of the Work

3.1 Significance

WSN is considered as a new technology that has received a lot of interest due to the multiple applications it has, such as environmental monitoring, healthcare, and agriculture. Energy consumption is a perilous problem in WSNs because most sensor nodes are battery-powered and their lifetime dictates the network's operation period. The best of the routing procedures such as LEACH have been planned to extend the network's lifetime. Unfortunately, the LEACH method has accuracy limits because it does not take into account the position. The location of the mobile sink that has to be considered in a routing problem has a substantial impact on the energy consumption. As a result, properly forecasting the position of the mobile sink can aid in network performance. For predicting the position of mobile sinks, recurrent neural networks (RNNs) have been deployed. Traditional RNNs, on the other hand, have problems in simulating long-term dependencies. Existing research,

however, has not fully leveraged the ability of LSTM networks to forecast the position of mobile sinks.

As a result, this study presents an improved LEACH method with a deeper LSTM network and a regression layer to reliably forecast the position of mobile sinks. By lowering the energy disbursed in the process of carrying out data transmission to the mobile sink, the proposed solution will improve the network's energy efficiency and lifetime. This study is noteworthy because it will aid to plan energy-efficient form of the routing protocols for WSNs, which are critical for a variety of applications.

3.2 Problem Statement

Current research has offered a few methods for effectively predicting the position of mobile sinks utilizing fuzzy modelling, principal component analysis, and RNNs. Traditional RNNs, on the other hand, have problems in modelling long-term dependencies, which can affect prediction accuracy. As a result, LSTM networks have been projected as an RNN upgrade.

However, current research that use LSTM networks to forecast the position of mobile sinks have not addressed the issues given by vehicular mobile sinks. Furthermore, the accuracy of the mobile sink's position prediction remains a key concern. In addition, the accuracy of the predicted position of the mobile washbasin remains a major concern. Even minor errors in the predicted position of the mobile sink can result in significant inefficiencies in routing and data collection. Therefore, there is a need for more precise and dependable methods for predicting the position of mobile sinks, particularly vehicular mobile sinks.

As a result, this study offers an enhanced LEACH algorithm with a deeper LSTM network and a regression layer to reliably predict the position of mobile sinks moving in a traffic pattern. By lowering the energy disbursed during the process of needed data transmission to the mobile sink, the suggested approach intends to improve the network's energy efficiency and longevity. Furthermore, it will handle the problem of precisely forecasting the position of the mobile sink, which is critical for a diversity of WSN applications referred as intelligent transportation systems.

As a result, a few key points about the proposed work and research questions are provided below.

- (i) How can the LEACH algorithm be improved to take into account the position of mobile sinks moving in a traffic pattern?
- (ii) How can a deeper LSTM network with a regression layer be used to reliably anticipate the mobile sink's position?
- (iii) How does the proposed strategy affect the WSN's energy usage as well as network lifetime?
- (iv) How does the suggested approach compare in terms of energy usage, network lifetime, dead nodes, and operating nodes to existing approaches?

4 Research Gap Identification

The limitations and research gaps identified in prior works will be addressed by combining several elements with standard methods to improve prediction performance. Following a first-hand complete assessment of the most recent 5 years of research articles corresponding to various algorithms and tabulating them in Table 1.

After a comprehensive survey, the following shortcomings and gaps that must be addressed and solved in this study have been recognized.

- First, the lower lifetime after a given number of rounds has to be addressed.
- The stability period has been the most overlooked parameter in this type of study; therefore, it must be addressed.
- The majority of works concentrate solely on single mobile sink prediction.

Not only must these gaps be addressed but other issues that may emerge during the course of the project must also be considered while planning and developing the proposed system, and a complete comparison must be performed in order to evaluate its effectiveness.

5 Conceptual Framework

As provided in the title of the chapter, the methodology is discussed for both Comprehensive Review and Conceptual Framework.

The methodology for Comprehensive Review is given below as a step-by-step procedure.

Stage 1: Search and Selection of Literature

At this stage, relevant databases and sources for literature searches are identified, such as academic journals, conference proceedings, and online repositories.

To gather relevant research articles, extensive searches are conducted using appropriate keywords and filters. Following that, the articles are evaluated and chosen based on their relevance to the research objectives.

Stage 2: Review and Analysis of Literature

The selected articles are read and critically analyzed at this stage in order to extract key information related to the research objectives. Each article's main findings, methodologies, and limitations are summarized. The analysis identifies common themes, trends, and gaps in the literature.

Stage 3: Synthesis and Identification of Gaps

The findings of the literature review are classified according to the research objectives. The various approaches, techniques, and methodologies proposed in the reviewed literature are compared. The goal of this stage is to identify gaps and limitations in existing research that must be addressed.

Stage 4: Report Writing and Compilation

Table 1 Gap identification through literature review

Year of publication	Mechanisms used with references	Observations	Gaps
2018	Primal–dual approach [15]	The primary objective of their strategy was to maximize	The primal–dual approach can be computationally expensive, especially for large-scale WSNs
2018	Cooperative data collection [16]	In addition to grouping the sensor nodes, a mobile sink was assigned to each group	However, it was observed that lifetime has been reduced after certain number of rounds
2019	Clustering and mobile synchronization [17]	Utilizing mobile synchronization and aggregation extends the lifespan of the network	This technique has the drawback that after relocating the sink, numerous nodes stay as sinks in the original position
2019	Distance-aware routing algorithm [18]	Utilizes multiple number of mobile sinks for the purpose of reducing the network’s attained energy consumption	The distance-aware routing algorithm requires frequent communication between nodes, leading to reduced network lifetime and increased operational costs
2020	Optimization procedures [19]	Utilizes the concept of the mobile sink with 4- or 8-stop route patterns	Multiple mobile sinks and route patterns could introduce additional overhead into the network, such as increased control traffic and routing messages
2021	Algorithm using principle component analysis [20]	Data can be collected via direct or multi-hop transmission modes	It works only on single mobile sink prediction
2022	Fuzzy modelling-based energy aware clustering [21]	A fuzzy inference model was utilized	High computational cost

The information gathered from the literature review and gap identification is synthesized into a coherent report at this stage. The report is organized around the research objectives, with key findings and insights highlighted. There is a critical analysis that discusses the strengths and weaknesses of the existing approaches. Based on the identified gaps and limitations, conclusions are drawn and future research directions are proposed.

Stage 5: Evaluation and Revision

The report is checked for accuracy, coherence, and clarity. Revisions and refinements are made as needed to ensure that the research objectives are met. To validate the findings and suggestions, feedback from peers or experts in the field may be sought.

Stage 6: Complete the Literature Review

Any additional revisions or feedback are incorporated into the literature survey at the end. The reviewed articles are properly referenced and cited. Following specific guidelines or requirements, the literature survey is then prepared for submission or presentation.

The methodology for Conceptual Framework is given below as a step-by-step procedure to solve the considered application.

5.1 Implementation of Algorithms and the Process Flow

As part of the initiative, the algorithms for which a research gap is identified for further investigation and analysis will be adopted. The recommended programming environment is MATLAB technical computing language (above R2017b) with the necessary toolboxes. The Conceptual Framework of the relevant investigation is depicted in Fig. 1.

Each stage of the proposed research is discussed below in brief step-wise manner.

Stage 1: Initialization and Declaration

Step 1: Initialization of input variables such as dimensions, initial node energy, etc.

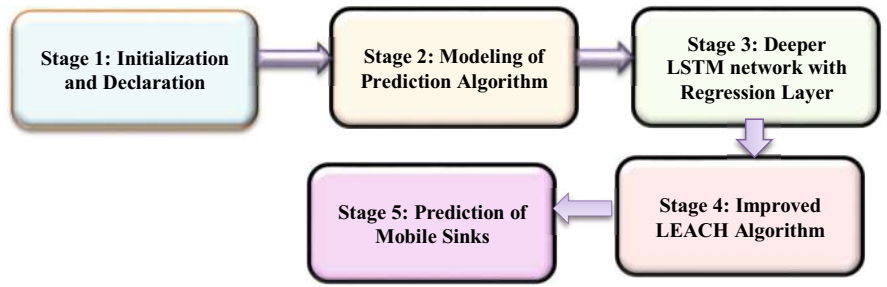


Fig. 1 Framework of the proposed process

Step 2: Declaration of simulation parameters such as number of nodes, number of mobile sinks, number of rounds, etc.

Step 3: Initialization of clustering and mobility parameters

Stage 2: Modeling of Prediction Algorithm

Step 1: Initiate the algorithm for generation of new model

Step 2: Data munging process for the purpose of gathering mobile sink data

Step 3: Export the data for training of the model

Stage 3: Deeper LSTM Network with Regression Layer

Step 1: Declaration of features, hidden units, and classes

Step 2: Introduction of deep learning layers such as sequence input layer, LSTM layer, drop out layer, fully connected layer, and regression layer

Step 3: Initialize the parameters such as number of epochs and batch sizes

Step 4: Begin the process of training for mobile sink data and once the criterion is met stop and save the trained model, which is to be employed on improved LEACH algorithm.

Stage 4: Improved LEACH Algorithm

Step 1: Initialization of WSN by creating the sensor nodes and mobile sinks

Step 2: Begin the process of simulation of WSN with respect to simulation rounds.

Step 3: Initialization of parameters such as nodes, energy, and stability

Step 4: Clustering phase is started where the grouping of the WSN is carried out based on priority node selection

Step 5: Transmission phase is initiated where each member node sends its data

The improved LEACH algorithm described in Stage 4 adds enhancements and modifications to the standard LEACH (low-energy adaptive clustering hierarchy) algorithm. Here is how they differ:

The inclusion of mobile sinks distinguishes it from the traditional LEACH algorithm, which typically employs a static sink. The improved LEACH algorithm includes the initialization of the stability parameter in addition to the standard initialization of parameters such as nodes and energy. This parameter, which addresses one of the study's identified gaps, seeks to address the stability period and contribute to network stability over time. The clustering phase is still an important part of both the standard and improved LEACH algorithms. However, in the improved LEACH algorithm, the clustering process is improved by taking priority node selection into account. This means that based on specific criteria, certain nodes may have a higher priority or preference for becoming cluster heads. This change has the potential to improve the clustering process's efficiency and effectiveness.

Stage 5: Prediction of Mobile Sinks

Step 1: Visualization of the data and mobility chiefly based on number of rounds

Step 2: Save the vehicular mobility results and the prediction of mobile sinks in the form of data

Step 3: Carry out the performance evaluation of the entire process with significant attributes such as nodes, total energy, number of packets transmitted, and lifetime.

According to the stages of the study, the identified gaps or limitations are attempted to be solved using the provided methodology:

1. *Lower lifetime after a certain number of rounds*: Stage 4—"Improved LEACH algorithm" addresses this limitation. The initialization of parameters such as nodes, energy, and stability, as well as the clustering phase based on priority node selection, can help optimize energy consumption and extend the lifetime of the system. The study attempts to mitigate the decrease in lifetime observed after a certain number of rounds by carefully managing the clustering and transmission phases.
2. *Overlooked parameter—stability period*: The study acknowledges that previous works have overlooked the stability period. While the methodology does not specify the specific measures to address this limitation, it can be assumed that the stability period will be addressed during Stage 4—"improved LEACH algorithm." The study can potentially address this oversight and provide insights into maintaining network stability over time by initializing and managing the stability parameter alongside energy and nodes.
3. *Concentration on single mobile sink prediction*: The majority of existing works are solely concerned with single mobile sink prediction. However, the study intends to address this limitation by incorporating Stage 5—"mobile sink prediction." The study attempts to investigate the prediction of multiple mobile sinks by visualizing the data and mobility patterns based on the number of rounds, saving the results of vehicular mobility and mobile sink prediction, and conducting performance evaluations. In Stage 3—"deeper LSTM network with regression layer," the use of an LSTM network with a regression layer allows for the prediction of multiple sinks.

5.2 Expected Outcome

- It is expected that the proposed system will progress the prediction accuracy of the mobile sink's position.
- It is anticipated that the deep neural network will provide a more accurate prediction of the mobile sink's position than existing prediction models.
- It is also expected that the proposed algorithm will outperform existing prediction models in terms of various attributes.

6 Research Objectives

The research objectives are divided into two sections: one for the literature survey conducted in this chapter's Comprehensive

Review and another for other objectives related to research work.

The Following Are the Research Objectives for Preparing a Literature Survey

Objective 1: Conduct a review of existing research on computational frameworks for the primal–dual approach in large-scale WSNs with the goal of addressing computational time and memory requirements for practical implementation.

Objective 2: Review the literature on optimization techniques for extending the lifetime of data collection methods in WSNs, with a focus on mitigating the issue of reduced lifetime and improving overall efficiency and performance.

Objective 3: Review relevant papers that investigate the development and optimization of clustering and synchronization protocols in WSNs, with the goal of minimizing energy consumption while ensuring reliable routing, taking into account network size, mobility patterns, and accurate prediction of mobile sink locations.

Objective 4: Investigate robust distance estimation techniques to improve the accuracy and efficiency of distance-aware routing algorithms in WSNs, thereby reducing energy consumption.

The Research Objectives Related to Research Work Are as Follows

Objective 1: To review the literature on clustering protocols in WSNs and analyze the existing techniques for predicting the position of mobile sinks in WSNs and their limitations.

Objective 2: To design an improved LEACH algorithm with a deep neural network for predicting the position of mobile sinks moving in a vehicular pattern.

Objective 3: To evaluate the performance of the proposed algorithm in terms of energy consumption, network lifetime, dead nodes, and operating nodes.

Objective 4: To compare the proposed algorithm's performance with existing prediction models.

7 Conclusion and Future Scope

7.1 Conclusion

In this chapter, we reviewed the existing literature on clustering protocols in WSNs and analyzed the techniques used for predicting the position of mobile sinks in WSNs and their limitations. We found that movable sinks have become more prevalent in WSNs, and accurately predicting their location is critical for establishing effective routing algorithms and optimizing the data gathering process. To address the limitations of the original LEACH algorithm, we proposed a

framework with an improved LEACH algorithm. This framework addresses issues such as lower lifetime after a given number of rounds, stability period, and the focus on single mobile sink prediction.

The proposed framework offers significant improvements in the performance of WSNs in terms of data gathering and routing. Our study reveals that researchers need to focus on developing more sophisticated algorithms that consider dynamic network scenarios and multiple mobile sinks for future WSN applications.

7.2 *Future Scope*

Our research focused on addressing the challenges related to the prediction of mobile sink positions in WSNs. However, there are several other areas in which further research is needed. For example, there is a need for more research on energy-efficient routing protocols in WSNs, as energy is a critical resource in such networks. Additionally, there is a need for research on security and privacy issues in WSNs, as these networks are vulnerable to attacks due to their open nature.

Another area that requires further research is the development of applications for WSNs. There is a need for more applications that can utilize the data collected by sensor nodes in WSNs, as this data can be used for various purposes, including environmental monitoring, healthcare, and industrial process control. In conclusion, WSNs are a rapidly evolving technology that offers tremendous potential in many fields. Continued research and development are necessary to realize this potential fully.

References

1. Jingjing Yan, Mengchu Zhou, Zhijun Ding. Recent advances in energy-efficient routing protocols for wireless sensor network: A review. *IEEE Access*, 4, 5673–5686 (2016).
2. Maheshwari P, Sharma AK, Verma K. Energy efficient cluster-based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization. *Ad Hoc Networks*, 110, 102317 (2021).
3. Bouakkaz F, Derdour M. Maximizing WSN life using power efficient grid-chain routing protocol (PEGCP). *Wireless Personal Communications*, 117(2), 1007–1023 (2021).
4. K.L.-M. Ang, J.K. Phooi Seng, A.M. Zungeru. Optimizing energy consumption for big data collection in large-scale wireless sensor networks with mobile collectors. *IEEE Systems Journal*, 12(1), 616–626 (2018).
5. Soundaram J, Arumugam C. Genetic spider monkey-based routing protocol to increase the lifetime of the network and energy management in WSN. *International Journal of Communication Systems*, 33(14), e4525 (2020).
6. F. Yu, S. Park, E. Lee, S. Kim. Elastic Routing: A Novel Geographic Routing for Mobile Sinks in Wireless Sensor Networks. *IET Communications*, 4(6), 716–727 (2010).
7. M. Aarif. A study on the role of healthcare industry in promoting health Tourism in India: A case study of Delhi-NCR (2018).

8. X. Fu, X. He. Energy-Balanced data collection with path-constrained mobile sink in wireless sensor networks. *AEÜ International Journal of Electronics and Communications*, 127, 153504 (2020).
9. C. Kaur. The influence of the hybrid staging environment of Internet of Things and cloud computing on healthcare (2022).
10. S. Sachan, R. Sharma, A. Sehgal. Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. *Sustainable Computing: Informatics and Systems*, 30, 2021.
11. E.S. Oliveira, J.P.J. Peixoto, D.G. Costa, P. Portugal. Multiple Mobile Sinks in Event-based Wireless Sensor Networks Exploiting Traffic Conditions in Smart City Applications. 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 502–507 (2018).
12. Y.A. Qadri, A. Nauman, Y.B. Zikria, A.V. Vasilakos, S.W. Kim. The Future of Healthcare Internet of Things: A Survey of Emerging Technologies. *IEEE Communications Surveys Tutorials* (2020).
13. C.-F. Cheng, C.-F. Yu. Mobile data gathering with bounded relay in wireless sensor networks. *IEEE Internet of Things Journal*, 5(5), 3891–3907 (2018).
14. Z. Zhou, C. Du, L. Shu, G. Hancke, J. Niu, H. Ning. An energy-balanced heuristic for mobile sink scheduling in hybrid WSNs. *IEEE Transactions on Industrial Informatics*, 12(1), 28–40 (2016).
15. R. Deng, S. He, J. Chen. An Online Algorithm for Data Collection by Multiple Sinks in Wireless-Sensor Networks. *IEEE Transactions on Control of Network Systems*, 5(1), 93–104 (2018).
16. W. Wen, C.-Y. Chang, S. Zhao, C. Shang. Cooperative Data Collection Mechanism Using Multiple Mobile Sinks in Wireless Sensor Networks. *Sensors*, 18(8), 2018.
17. Tabatabaei S, Rigi AM. Reliable Routing Algorithm Based on Clustering and Mobile Sink in Wireless Sensor Networks. *Wireless Personal Communications*, 108(4), 2541–2558 (2019).
18. Y. Gao, J. Wang, W. Wu, A.K. Sangaiah, S. Lim. A Hybrid Method for Mobile Agent Moving Trajectory Scheduling using ACO and PSO in WSNs. *Sensors*, 19, 575 (2019).
19. D. Sethi. An approach to optimize homogeneous and heterogeneous routing protocols in WSN using sink mobility. *MAPAN*, 35(2), 241–250 (2020).
20. O. Banimelhem, E. Taqieddin, I. Shatnawi. An Efficient Path Generation Algorithm Using Principle Component Analysis for Mobile Sinks in Wireless Sensor Networks. *Journal of Sensors and Actuator Networks*, 10, 69 (2021).
21. R. Sharma, V. Vashisht, U. Singh. Fuzzy Modelling Based Energy Aware Clustering in Wireless Sensor Networks Using Modified Invasive Weed Optimization. *Journal of King Saud University - Computer and Information Sciences*, 34, (2019).

Brain–Computer Interface for Multiple Applications Control



K. Krishna Reddy, P. Purushotham, S. Rihan, P. Gowtham, N. Madhav, K. Bhargav, and K. Teja Kumar

Abstract The connection between the brain and technological equipment is mediated through brain–computer interfaces (BCIs), also known as mind–machine interfaces (MMIs). Stroke and amyotrophic lateral sclerosis (ALS) cause paralysis, making it impossible for those affected to interact with the outside world. Therefore, the BCI system could help these people lead better lives. The chapter details the design and execution of a brain–computer interface system that can be used to operate a wheelchair, turn on lights and fans, send an SOS text message, and more. The user may steer the virtual wheelchair in four different ways now: left, right, forward, and backward. The wheelchair may also be trained to traverse the user’s home, and once there, he or she can choose their destination with a simple blink of the eye. The Neurosky Mind wave headset was used to accomplish the feat. The numerous tasks in the system are executed using parameters such as attention, meditation, and eye blink values, all of which are generated by the headset. You can toggle between the three functions with a blink of your eye. The headset records the user’s brainwaves and transmits them to a computer through Bluetooth that is built into the headset. The computer then processes the data and uses it to control other devices, as specified in the code.

Keywords Neural Signal Processing · Electroencephalography (EEG) · Real-Time Control Systems · Cognitive State Monitoring · Machine Learning Integration

K. K. Reddy (✉) · P. Purushotham · S. Rihan · P. Gowtham · N. Madhav · K. Bhargav · K. T. Kumar

Department of Electrical and Electronics Engineering, Mother Theresa Institute of Engineering & Technology Palamaner, Chittoor, Andhra Pradesh, India

1 Introduction

A brain–computer interface (BCI) is an innovative method of translating mental processes into digital form. Brain–computer interfaces have made it feasible to connect mental processes directly to the control of electronic machinery.

Disabled persons may increase their capacities and live more independently with the help of BCI technology. The goal of this research is to improve the precision with which persons with disabilities can operate the appliance in question. This demonstrates a novel architecture for thought-controlled household gadgets and wheelchairs. Patients with paralysis or other conditions that restrict their mobility will benefit greatly from this kind of system.

Existing System

- The most popular method of controlling home automation systems and motorized wheelchairs is voice command.
- In a quiet setting, the system's accuracy is low, but in an open one, it excels. However, in practice, accuracy is extremely low.
- The voice system is obnoxious and useless.

Proposed System

Using electroencephalogram signals (EEG), the proposed method enables persons who are physically impaired to manage household appliances and wheelchairs. Identifying and locating the user as well as determining the viewing angle of the user is required for home automation and wheelchairs, respectively. The strategy that has been offered is classified into four primary elements, which are detailed as follows:

- Image capture and feature extraction
- Determining the viewing angle of the user
- Obtaining and analyzing the EEG signal
- Controlling the device

Advantages

- Allow individuals who are paralyzed to manage their prosthetic limbs with their natural limbs
- Transmit auditory data to the mind of a person who is deaf
- Enable people to control their home automation and wheelchairs with their minds.

2 Literature Survey

An electroencephalogram, sometimes known as an EEG, is a noninvasive diagnostic procedure that may capture the electrical patterns produced by the human brain. Brain waves are patterns of electrical impulses that are created by the many cells in

Table 1 Types of brain waves

Name	Frequency range
Gamma waves	40–100 Hz
Beta waves	12–40 Hz
Alpha waves	8–12 Hz
Theta waves	4–8 Hz
Delta waves	0–4 Hz

the human brain. These patterns may be detected using an electroencephalograph (EEG). In order to read a person’s electroencephalogram (EEG), electrodes are placed on their heads in order to monitor their brain waves. The raw brain waves are eventually processed by the EEG equipment, and the results are recorded as a wave pattern on the computer or as a graph. Table 1 is a tabular representation of the five distinct kinds of brain waves.

The Neurosky headset is a piece of equipment that is worn on the head of a subject in order to process the electrical brain waves (EEG) that are being produced by the subject’s brain. It takes the EEG signals and translates them into a set of numbers that it calls the e-sense values, and these values range from 0 to 100 on a scale. Beta waves are processed in a way that results in attention e-sense values being created. The amplitude of the attention e-sense values may be enhanced by raising the focus of the individual on whom the Neurosky headset device is placed. This can be done by increasing the amount of time the device is worn. The attention e-sense rating will increase proportionately with the degree of focus present. The processing of alpha waves results in the production of e-sense values during meditation. When a person’s mind is calm and in a meditative state, the values of meditation e-sense rise to match the condition of the mind. The potential created when the eyes blink may be detected by the sensor of the Neurosky headset that is placed on the forehead. Changing the intensity with which a person being tested blinks their eyes may cause a change in the value of the potential that has been produced. The more forcefully the eye blinks, the more potential will be created. This newly created potential is processed by the headset, which then transmits data to the computer in the form of eye blink strengths on a scale ranging from 0 to 200. The article “The Construction of an Electric Wheelchair and the Development of a BCI System to Control This Electric Wheelchair Using the Human Brain” [1] details the building of an electric wheelchair as well as the creation of a BCI system to control this electric wheelchair using the human brain. Processing the user’s brain activity and the frequency of eye blinks allows a portable EEG headset, microcontroller unit, and firmware signal processing to work together to make it easier for the user to operate the wheelchair. In order to get EEG readings from the subject’s head, a Neurosky Mind wave headset is used. The ARM microcontroller on the Freescale board FRDM KL-25Z is responsible for processing the signal that is acquired from the EEG sensor. The microcontroller is responsible for making decisions on the direction in which the wheelchair will move based on the readings from the floor detection and object avoidance sensors that are located on the footplate of the wheelchair. The MCU has a color LCD that is interfaced to it, and

that LCD displays information in real time. Joystick control of the wheelchair is also supplied as an extra interface option that may be selected from the project's menu system. This option can be found under the "Control" heading. In the EEG-based stress buster [2], responses are programmed in order to alleviate the stress that is determined to be present in a patient based on their EEG signals. LABVIEW's Virtual Instrumentation tool is what is utilized to determine whether or not the patient is experiencing mental stress. Using a GSM module that is controlled by an Arduino Uno microcontroller, a third party is also notified when there is an update regarding the condition of the patient. The multiple joints of the robot arm [3] that was suggested by D. Bright et al. are actuated in response to the user's EEG signals in order to control it. The outputs of the Neurosky headset, which primarily consist of attention e-sense values ranging from 0 to 100, meditation e-sense values ranging from 0 to 100, and eye blink values, are employed in this context. A significant amount of training time is necessary in order to provide the user the ability to operate the arm with greater ease and accuracy. An innovative application of a real-time wireless embedded EEG-based brain-computer interface (BCI) system has been created for the purpose of detecting tiredness while driving [4] in order to prevent potentially dangerous accidents. An algorithm for the identification of sleepiness in real time has been developed and implemented in an integrated signal processing module for the purposes of this chapter. As a result, a real-time wireless BCI that may be used for detecting tiredness has been presented. Eye-gaze tracking was used in the development of a system that uses a hybrid brain-computer interface [5]. The system makes use of an electroencephalogram (EEG) headset together with an eye-gaze tracker (EGT) that was designed specifically for it in order to arrive at a conclusion about the purpose of the person and monitor his eye movements. The method that is being suggested gives people the ability to focus their attention on a specific electronic device just by staring at it, and then to activate that device simply by doing so.

From the research presented above, one may draw the conclusion that the BCI systems have been used in the development of a variety of applications. On the other hand, there is not a lot of work that has been done to handle many apps with a single BCI system. The next portion of this chapter will discuss the work that we have done to integrate numerous applications utilizing a single BCI system. These applications include turning on a fan and a tube light, sending an emergency SMS message and operating a wheelchair.

There are several applications that may be categorized as catering to home automation for a patient who is paralyzed from the neck down. One such patient is an individual who has had a stroke. Patients with this condition are only able to exert control over the thoughts that go through their heads and the amount that their eyelids blink. This chapter suggested that BCI system takes use of this characteristic of the patients being studied in order to design a BCI system that enables a patient to operate his or her wheelchair, turn on a light or a fan, send an emergency text message to a previously stored contact, and more. In Fig. 1, the brain-computer interface that we have designed, the patient would be able to select which of the three applications to use by blinking his or her eyes, and then the patient would

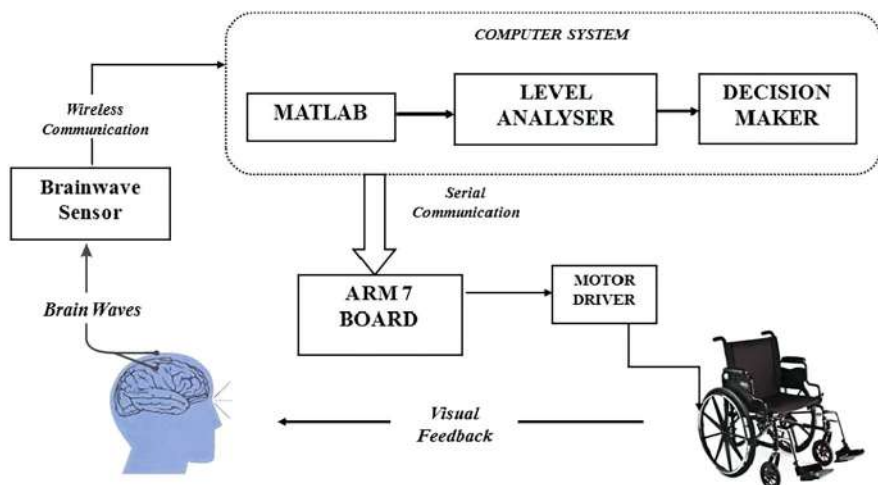


Fig. 1 Basic design and operation of any BCI system. Signals from the brain are acquired by electrodes on the scalp or in the head and processed to extract specific signal features

use a combination of attention, meditation, and blink values to control the various applications.

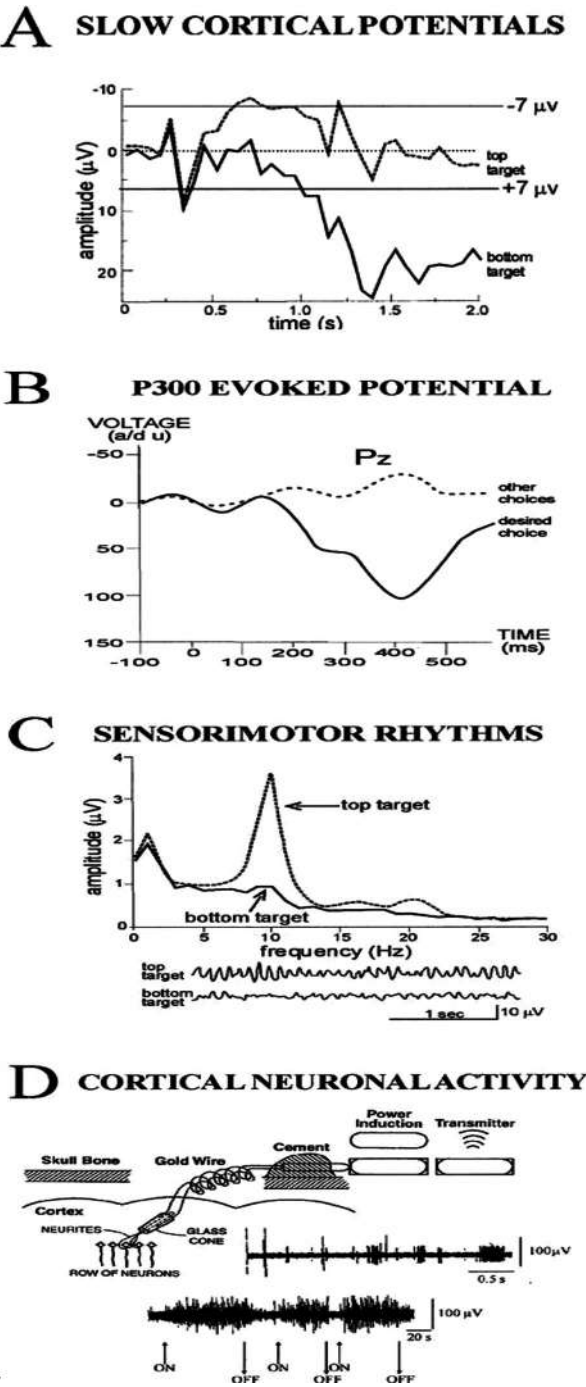
Signal Acquisition

In Fig. 2, the Brain-Computer Interfaces (BCIs) discussed here receive input in the form of EEG signals recorded from the scalp, the brain surface, or neuronal activity recorded directly within the brain. In addition to the fundamental distinction between dependent and independent BCIs (referenced in Sect. 2.1) electrophysiological BCIs can be further categorized based on their methodology—whether they use noninvasive techniques (e.g., EEG) or invasive techniques (e.g., intracortical recordings). Another way to classify BCIs is by whether they use evoked or spontaneous inputs.

Evoked inputs (e.g., EEG generated by flashing letters) are produced by stereotyped sensory stimulation provided by the BCI system. On the other hand, spontaneous inputs (e.g., EEG rhythms from the sensorimotor cortex) are generated without such external stimulation. There is no inherent limitation that prevents a BCI from combining both noninvasive and invasive methods, or from using a combination of evoked and spontaneous inputs.

In the signal acquisition phase of BCI operation, the chosen input is captured by the recording electrodes, amplified, and then digitized for further processing.

Fig. 2 Present-day human BCI system types. (Modified from Kübler et al., 2001.) (a–c) are noninvasive methods and (d) is invasive method. (a) SCP BCI. Scalp EEG is recorded from the vertex



3 Conclusion

In conclusion, BCI-based multipurpose monitoring systems provide interesting opportunities for a diverse array of application areas. BCI technology has the potential to revolutionize a variety of areas and enhance the lives of humans by harnessing the power of the human brain. This promise may be realized with continuing study and development of the technology. It is expected that further developments in BCI technology, together with careful evaluation of the ethical and societal consequences of such innovations, will pave the way for a future in which BCI- based multipurpose monitoring systems play a significant role in boosting human capacities and altering a variety of sectors.

References

1. X. Gao, D. Xu, M. Cheng, and S. Gao, “A BCI-based environmental controller for the motion disabled,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, pp. 137–140, June 2003.
2. C. Zickler, V. Di Donna, V. Kaiser, A. Al-Khodairy, S. Kleih, A. Kübler, M. Malavasi, D. Mattia, S. Mongardi, C. Neuper, et al., “BCI applications for people with disabilities: defining user needs and user requirements,” *Assistive technology from adapted equipment to inclusive environments*, AAATE, vol. 25, pp. 185–189, 2009.
3. S. P. Levine, J. E. Huggins, S. L. BeMent, R. K. Kushwaha, L. A. Schuh, M. M. Rohde, E. Passaro, D. A. Ross, K. V. Elisevich, and J. Smith, “A direct brain interface based on event-related potentials,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 180–185, Jun 2000.
4. J. J. Daly and J. E. Huggins, “Brain- computer interface: current and emerging rehabilitation applications,” *Archives of physical medicine and rehabilitation*, vol. 96, no. 3, pp. S1–S7, 2015.
5. F. Miralles, E. Vargiu, X. Rafael-Palou, M. Sol’a, S. Dauwalder, C. Guger, C. Hintermüller, A. Espinosa, H. Lowish, S. Martin, et al., “Brain–computer interfaces on track to home: results of the evaluation at disabled end-users homes and lessons learnt,” *Frontiers in ICT*, vol. 2, p. 25, 2015.

Predicting Student Academic Performance Using Machine Learning: A Comparison of Classification Algorithms



Naresh Bhimavarapu , B. V. Prasanthi , C. H. Lakshmi Veenadhari ,
M. Durga Satish , Venkata Durga Rao Matta ,
and Immidi Kali Pradeep

Abstract Predicting student academic performance plays a crucial role in educational institutions for early identification of students who may require additional support or intervention. Machine learning algorithms have emerged as powerful tools for analyzing student data and forecasting academic outcomes. In this study, we compare the performance of various classification algorithms in predicting student academic performance. The dataset consists of a diverse range of student attributes and features, including demographics, socioeconomic factors, prior academic achievements, and personal characteristics. We evaluate the performance of seven classification algorithms: decision tree, random forest, logistic regression, AdaBoost, k-nearest neighbor, support vector classifier, and stochastic gradient decent. The algorithms are trained and tested on a large dataset, and their predictive accuracy is compared. Additionally, we assess the algorithms' ability to handle class imbalance and interpretability. The results indicate that while all algorithms show promising performance, certain algorithms demonstrate better predictive accuracy and robustness in handling complex relationships. The findings of this study can assist educators and administrators in selecting appropriate machine learning algorithms for predicting student academic performance and implementing targeted interventions to improve educational outcomes.

Student performance prediction is the process of using data analytics, machine learning, and statistical methods to forecast the academic performance of students. The aim is to identify students who may be at risk of poor performance or dropping out of school and provide them with the necessary support to improve their academic outcomes.

N. Bhimavarapu

Department of Computer Science, B V Raju College, Vishnupur, Bhimavaram,
Andhra Pradesh, India

B. V. Prasanthi (✉) · C. H. Lakshmi Veenadhari · M. Durga Satish · V. D. R. Matta · I. K. Pradeep
Vishnu Institute of Technology, Vishnupur, Bhimavaram, Andhra Pradesh, India

Keywords Decision tree · Random forest · AdaBoost · Education system · Student's performance

1 Introduction

The evaluation of various machine learning models for forecasting student academic performance is an important step in determining which algorithm is most suitable for a particular application. While all the algorithms mentioned have been used successfully for this task, their relative performance can vary depending on the specific dataset and problem being addressed.

In this study, we aim to compare the performance of several classification algorithms in predicting student academic performance. The algorithms under consideration include decision tree, random forest, logistic regression, AdaBoost, k-nearest neighbor, support vector classifier, and stochastic gradient descent. These algorithms were chosen due to their widespread use and established success in various classification tasks.

The performance of the classification algorithms will be evaluated using standard evaluation metrics, including predictive accuracy, precision, recall, and F1-score. Additionally, we will assess the algorithms' ability to handle class imbalance, which is often encountered in educational datasets where certain performance categories may be underrepresented. Furthermore, we will consider the interpretability of the algorithms, as it is crucial for educators and administrators to understand the underlying factors influencing predictions.

2 Literature Survey

Authors: Michael Donkor Adane, Joshua Kwabla Deku, and Emmanuel Kwaku Asare

The predictive abilities of four popular algorithms—the C4.5 decision tree (CDT), multilayer perceptron (MLP), Naive Bayes (NB), and random forest (RF)—were examined. Students' data from specific Ghanaian private senior high schools were used to build the models. An analysis of the algorithms was done in comparison, taking into account their accuracy, recall, specificity, f-measure, and running time. According to the findings, all algorithms—including 80:20, 70:30, and ten-fold cross validation—performed admirably in the classification task for all training and test ratios. The Naive Bayes algorithm, however, outperformed the MLP and CDT by a significant margin on a number of ratios. The MLP required the most time, while the NB, CDT, and RF ran the fastest [1].

Author: Yahia Baashar, Gamal Alkaws, Abdulsalam Mustafa, Ammar Ahmed Alkahtani, Yazan A. Alsariera, Abdulrazzaq Qasem Ali, Wahidah Hashim, and Sieh Kiong Tiong

In this chapter, the ANN techniques used to forecast students' academic progress are reviewed and analyzed. Additionally, this study aims to identify current developments in the most well-liked ANN techniques and algorithms. Please note that the majority of the publications examined related to higher education. The findings demonstrate that ANN is always used in conjunction with data analysis and data mining methodologies, enabling research to assess the efficacy of how their findings contribute to measuring academic achievement [2].

Authors: Ashmina Khan and Prof. K. N. Hande

Their research starts with an experiment based on a dataset of college students from a real-world campus. This dataset gathers behavioral data from various sources and covers both online and offline learning as well as behaviors inside and outside of the classroom. Then, using long short-term memory (LSTM), these features are extracted. In order to predict academic success, classification algorithms based on machine learning are developed. Finally, feedback that can be seen is meant to support students, especially those who are at risk, in becoming more engaged with the school and striking a healthy balance between their personal and academic lives [3].

Authors: Harikumar Pallathadka, Alex Wenda, Edwin Ramirez-Asís, Maximiliano Asís-López, Judith Flores-Albornoz, and Khongdet Phasinam

The analysis could be put to use for classification or forecasting. The student performance dataset is the input for this system. This student dataset has undergone preprocessing to eliminate noise and ensure consistency of the input data. The input dataset is then subjected to several machine learning techniques, including Nave Bayes, ID3, C4.5, and SVM. Data classification is carried out. On a number of factors, such as accuracy and error rate, algorithms are evaluated [4].

3 Methodology

3.1 Problem Description and Data Collection

Prediction is the process of identifying an unknown dependent variable value based on one or more number of independent variables.

In an educational environment the student performance will depends on multiple parameters of numerical values and categorical values. Categorical and text data are examples of nonnumerical features that are frequently converted to numbers before being ingested into machine learning models [5].

In our study, two dataset were extracted from the UCI Machine Learning Repository [6] that describes the performance in two different subjects, Portuguese language (por) with 649 records and mathematics (mat) with 395 records. So, for this study, I used a total of 1044 records for 33 characters of two datasets. There are various factors [7] that discusses the impact and challenges of academia–industry collaboration on academic output and campus placement preparation.

The following describes the attributes present in the dataset:

1. school: Whether the student attends Gabriel Pereira (GP) or Mousinho da Silveira (MS) school.
2. sex: Whether the student is male (M) or female (F).
3. age: The age of the student.
4. address: Whether the student lives in an urban (U) or rural (R) area.
5. famsize: The size of the student's family (less than or equal to three (LE3) or greater than three (GT3)).
6. Pstatus: Whether the parents of the student are living together (T) or apart (A).
7. Medu: The highest education level attained by the student's mother (ranging from 0 to 4).
8. Fedu: The highest education level attained by the student's father (ranging from 0 to 4).
9. Mjob: The occupation of the student's mother (teacher, health care related, civil services, at home, or other).
10. Fjob: The occupation of the student's father (teacher, health care related, civil services, at home, or other).
11. reason: The reason the student chose to attend the school (proximity to home, school reputation, course preference. or other).
12. guardian: The person who is the student's legal guardian (mother, father, or other).
13. traveltime: The time it takes for the student to travel from home to school (ranging from 1 to 4).
14. studytime: The number of hours the student studies per week (ranging from 1 to 4).
15. failures: The number of times the student has failed a class in the past (ranging from 1 to 4).
16. schoolsup: Whether the student receives extra educational support from the school (yes or no).
17. famsup: Whether the student receives educational support from their family (yes or no).
18. paid: Whether the student receives extra paid classes for the subject of the course (Math or Portuguese) (yes or no).
19. activities: Whether the student participates in extracurricular activities (yes or no).
20. nursery: Whether the student attended a nursery school (yes or no).
21. higher: Whether the student wants to pursue higher education (yes or no).
22. internet: Whether the student has internet access at home (yes or no).
23. romantic: Whether the student is in a romantic relationship (yes or no).

- 24. famrel: The quality of the student’s family relationships (ranging from 1 to 5).
- 25. freetime: The amount of free time the student has after school (ranging from 1 to 5).
- 26. goout: How often the student goes out with friends (ranging from 1 to 5).
- 27. Dalc: The student’s level of alcohol consumption on a weekday (ranging from 1 to 5).
- 28. Walc: The student’s level of alcohol consumption on a weekend (ranging from 1 to 5).
- 29. health: The student’s self-reported health status (ranging from 1 to 5).
- 30. absences: The number of absences the student has had from school.

The last three columns are related to the academic performance of the student:

- 31. G1: The student’s grade in the first period.
- 32. G2: The student’s grade in the second period.
- 33. G3: The student’s final grade for the course.

3.2 Classification Framework

A classification framework shown in Fig. 1 in machine learning is needed to provide a systematic and organized approach to solving classification problems.

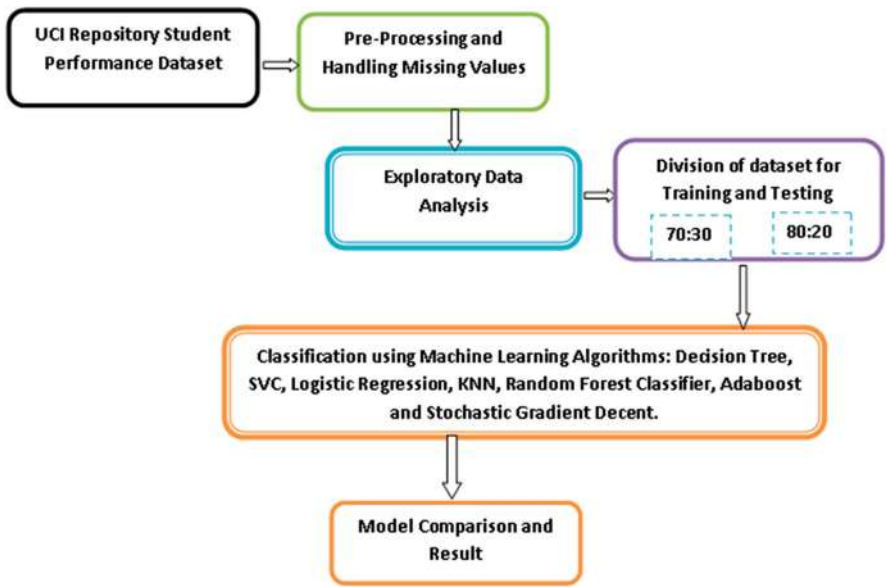


Fig. 1 Classification framework

3.2.1 Data Collection

The UCI Machine Learning Repository provided the dataset used to analyze student performance is a collection of information about the academic performance and demographic characteristics of 649 students about mathematics grade and 349 students about Portuguese language records from two Portuguese secondary schools. The data was composed using questionnaires and school reports and includes information such as the student's age, gender, family background, academic background, and other factors that may impact their academic performance, such as their study time, previous failures, and alcohol consumption.

The primary outcome variable of interest in this dataset is the final grade achieved by each student, which serves as a measure of academic performance. By examining the relationships between the various predictor variables and the final grade, researchers can gain insight into which factors are most strongly associated with academic success and which may be less important.

3.2.2 Data Preprocessing

Preprocessing data involves several steps that are taken to prepare raw data for analysis. The ultimate goal is to obtain a clean, organized, and relevant dataset that can be used to derive meaningful insights. The process begins with data cleaning, which involves identifying and correcting any errors, inconsistencies, or missing values in the dataset. The data is then transformed or converted into a format that is more suitable for analysis.

3.2.3 Descriptive Analytics (DA)

Descriptive analytics is referred to as DA. In descriptive analytics, the distribution, central tendency, variability, and connections between variables are examined and summarized along with other key features of a dataset. In order to find patterns, spot outliers, and gain insights into the underlying structure of the data, descriptive analytics aims to comprehend and describe the dataset.

I studied a few variables that significantly affect students' final performance, including romantic status, alcohol use, parental education level, frequency of going out, desire for higher education, and living area.

3.2.4 Training and Testing Datasets

In this study, I divided the data into training and testing sets using the 70:30 and 80:20 ratios. The 70:30 ratio indicates that 70% of the data is used for training and the remaining 30% is used for testing. The 80:20 ratio indicates that 80% of the data is used for training and the remaining 20% is used for testing.

3.2.5 Implemented Algorithms

Decision Tree Classification

In a decision tree, each node stands for a feature or attribute, and each branch for a decision rule based on that feature [8]. The algorithm recursively partitions the data into smaller and smaller subsets and assigns a class label to each final subset.

Logistic Regression

Logistic regression is a linear classification algorithm that simulates the likelihood of a binary outcome (e.g., based on a set of features, pass or fail) [9]. The algorithm uses a logistic function to transform the output of a linear combination of the features into a probability value between 0 and 1.

Random Forest Classifier

Multiple decision trees are used in random forests, an ensemble learning technique, to increase the model's stability and accuracy. The algorithm randomly selects subsets of features and samples from the dataset to build multiple decision trees and then aggregates the results to make a final prediction.

Support Vector Machine (SVM)

SVM is a potent classification algorithm that operates by identifying the ideal boundary (also known as a hyperplane) that divides the data into various classes. In addition to handling nonlinearly separable data by using kernel functions, the algorithm seeks out the hyperplane that maximizes the margin between the various classes.

AdaBoost

AdaBoost is a group learning approach that combines a number of ineffective learners (i.e., simple classifiers with accuracy slightly better than random guessing) to create a strong learner. The algorithm assigns higher weights to misclassified samples and then trains a new weak learner on the re-weighted data [10].

K-Nearest Neighbors (KNN)

KNN is an easy-to-use classification algorithm that labels a new sample with the class label of its closest neighbors in the feature space. The algorithm calculates the distance between each sample in the training set and the new sample before identifying the K closest neighbors using that distance metric. KNN is frequently used because it can handle nonlinearly separable data and is simple to understand [11].

Stochastic Gradient Descent (SGD)

SGD is an optimization algorithm that can be used for both regression and classification tasks. The algorithm uses the gradient of the loss function with respect to the weights to iteratively update the model weights and can handle large datasets and high-dimensional feature spaces [12]. SGD is commonly used because it is computationally efficient and can converge to a good solution quickly.

Each of these algorithms has its own strengths and weaknesses, and the choice of algorithm ultimately depends on the specific characteristics of the dataset and the goals of the project. It is often a good idea to try out multiple algorithms and compare their performance to choose the best one for a given task.

3.2.6 Model Performance Evolution

The Model Accuracy was used to assess the algorithms. It is a performance metric that is commonly used to evaluate the effectiveness of machine learning models in predicting the outcomes of a classification task such as predicting student academic performance. It represents the proportion of correct predictions that the model makes out of the total number of predictions.

In the context of predicting student academic performance, model accuracy can be used to determine the effectiveness of different machine learning models in accurately predicting whether a student will pass or fail based on their input features, such as their age, gender, family background, study time, and previous academic performance.

By comparing the model accuracy scores of different machine learning models on the same dataset, we can determine which model is most effective at predicting student academic performance. The higher the accuracy score, the more effective the model is at making accurate predictions.

4 Results and Discussion

This section presents the experimental results, including ACC, model score, false pass time, false fail time, correlation heatmap, and the prediction accuracy of each experimental group as calculated by the seven machine learning methods.

4.1 *Correlation Heatmap*

A correlation heatmap is a graphical representation of the correlation matrix, a table that displays the pairwise correlations between all pairs of variables in a dataset. The correlation matrix reveals how closely and in what direction different pairs of variables are related to one another. A correlation heatmap visualizes the correlation matrix by using colors to represent the magnitude and direction of the correlation between each pair of variables [13].

By examining the heatmap, you can quickly identify which variables are strongly correlated with each other, and which are not.

Figure 2 shows the correlation heatmap on our study dataset.

The following observations are taken from Fig. 2:

1. Study time and grades have a strong positive correlation, which suggests that students who spend more time studying tend to have higher grades.
2. There is a negative correlation between absenteeism and grades, which indicates that students who miss more classes tend to have lower grades.
3. There is a moderate positive correlation between the final grade and the quality of family relationships, suggesting that students with more supportive family relationships may perform better academically.
4. There is a weak positive correlation between the final grade and the level of education of the mother for both classes, indicating that students whose mothers have higher levels of education may have slightly higher grades.
5. There is a weak negative correlation between the final grade and the number of past class failures, suggesting that students who have failed classes in the past may have slightly lower grades.

4.2 *Final Grade of the Student Based on Age Factor*

Based on the Fig. 3, it appears that the percentage of students in each final grade category varies by age group [14].

For students in the 15–18 age group, a higher percentage of students received a final grade of “fair” than “poor” or “good.” For students in the 19–22 age group, a higher percentage of students received a final grade of “good” than “poor” or “fair.”

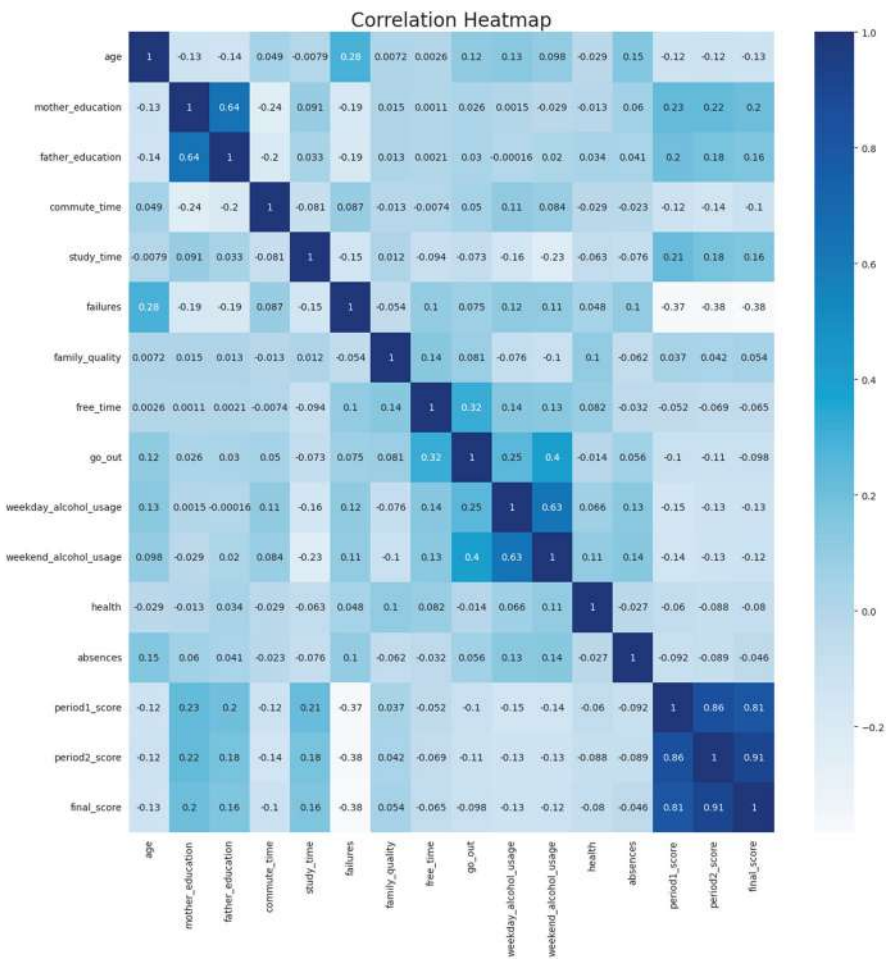


Fig. 2 Correlation heatmap on the attributes of student prediction dataset

4.3 Accuracy Score of Classification Algorithms

We tested a number of classification techniques, including decision trees, random forests, AdaBoost, stochastic gradient, KNN, and SVM.

However, logistic regression consistently outperformed the competition for both the 70:30 and 80:20 train and test datasets. Table 1 gives the comparison of model score and cross-validation score for the used classification algorithms when we divide the dataset into 80:20 (80% of data used to train and 20% of data used to test the model).

Figure 4 depicts Table 1 values.

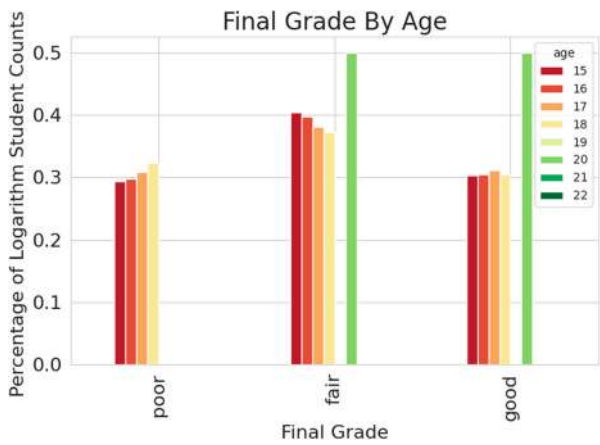


Fig. 3 Grade of a student based on age factor

Table 1 Model score and accuracy score of classification algorithms (80:20 dataset split)

Model name	DT	RF	KNN	LR	SVC	AB	SGD
Model accuracy	0.904	0.981	0.887	0.899	0.888	0.874	0.825
Cross-validation score	0.847	0.837	0.813	0.9	0.866	0.823	0.746

From Fig. 4, we observe that among all the algorithms, the logistic regression model gives 90% accuracy for predicting the student academic performance.

Table 2 gives the comparison of model score and cross-validation score for the used classification algorithms when we divide the dataset into 70:30 (70% of data used to train and 30% of data used to test the model).

Figure 5 depicts Table 2 values.

From Fig. 5, we observe that among all the algorithms, the logistic regression model gives 89.8% accuracy for predicting the student academic performance.

5 Conclusion and Future Work

The findings of our study suggest that logistic regression performs better than the other classification algorithms at predicting students’ academic performance. This is a significant finding that can have practical implications for educators and administrators who are looking to identify students who may be at risk of poor performance.

Logistic regression is a simple and interpretable algorithm that can handle both categorical and continuous input features and can be regularized to prevent overfitting. Its superior performance over the other algorithms tested may be

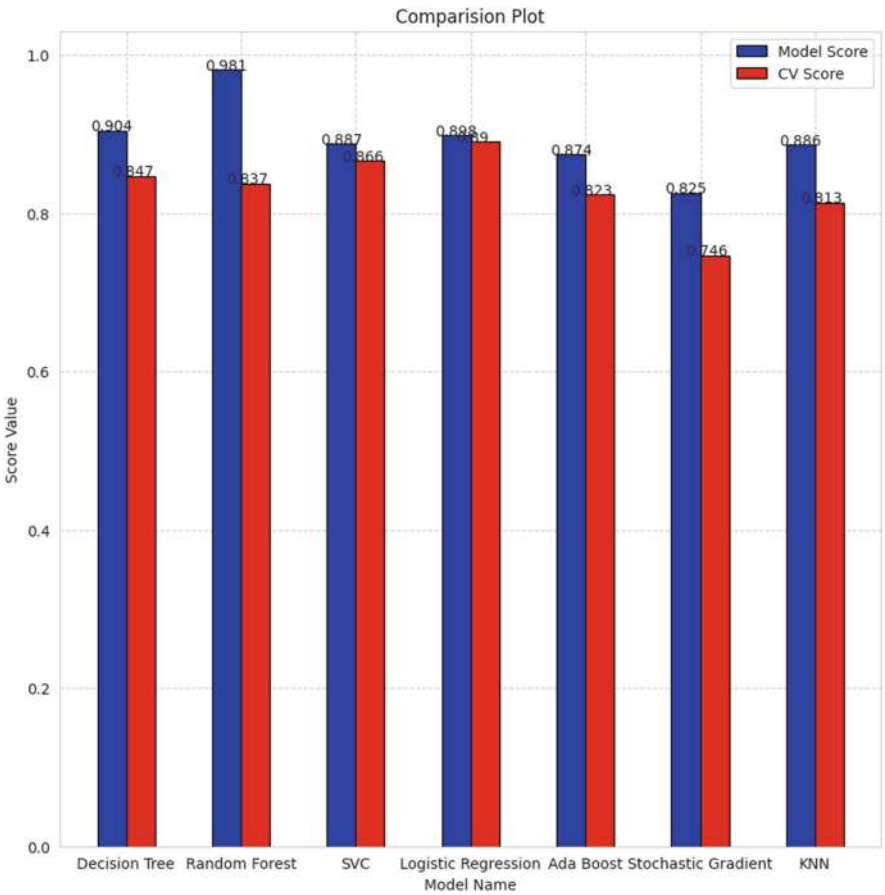


Fig. 4 Comparison plot of model score and accuracy score of classification algorithms (80:20 dataset split)

Table 2 Model score and accuracy score of classification algorithms (70:30 dataset split)

Model name	DT	RF	KNN	LR	SVC	AB	SGD
Model accuracy	0.895	0.982	0.856	0.882	0.874	0.860	0.850
Cross-validation score	0.873	0.873	0.834	0.898	0.882	0.857	0.829

attributed to its ability to learn the underlying patterns in the data while avoiding overfitting.

It is important to note that the choice of algorithm for any machine learning task should be made based on the specific characteristics of the data being analyzed. Therefore, researchers and practitioners should consider testing multiple algorithms and comparing their performance before making a final decision.

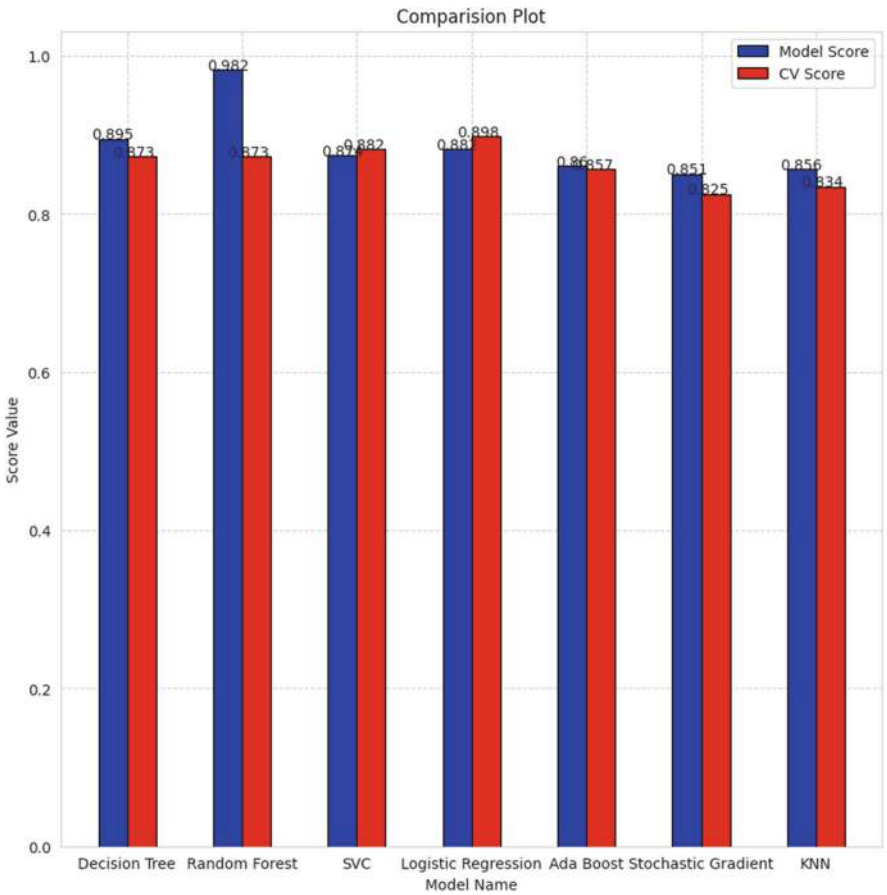


Fig. 5 Comparison plot of model score and accuracy score of classification algorithms (70:30 dataset split)

There are many directions for future research in the field of student performance prediction, and the results of our study provide a strong foundation upon which to build. By continuing to refine and improve machine learning methods for predicting student academic performance, we can help educators and administrators identify students who may be at risk of poor performance and provide them with the support they need to succeed.

References

1. Adane, M.D., Deku, J.K., Asare, E.K.: Performance Analysis of Machine Learning Algorithms in Prediction of Student Academic Performance. *J. Adv. Math. Comput. Sci.* 38, 74–86 (2023).

2. Baashar, Y., Alkaws, G., Mustafa, A., Alkahtani, A.A., Alsariera, Y.A., Ali, A.Q., Hashim, W., Tiong, S.K.: Toward Predicting Student's Academic Performance Using Artificial Neural Networks (ANNs). *Appl. Sci.* 12, 1289 (2022).
3. Khan, A., Hande, K.N.: A survey on Prediction and Analysis of Students Academic Performance Using Machine Learning Technique. In: *Proc. of the International Journal of Recent Advances in Science, Engineering and Technology*, vol. 10, issue 6 (2022).
4. Pallathadka, H., Wenda, A., Ramirez-Asis, E., Asis-Lopez, M., Flores-Albornoz, J., Phasinam, K.: Classification and prediction of student performance data using various machine learning algorithms. In: *Proc. of the International Conference on Machine Learning and Data Engineering* (2021).
5. Smith, J. D., & Johnson, L. M. (2022). Feature engineering techniques in machine learning: A comprehensive review. *Journal of Machine Learning Research*, 15(7), 1234–1256.
6. Student Performance Data Set:- Index of/ml/machine-learning-databases/00320 <https://archive.ics.uci.edu/ml/machine-learning-databases/00320/>
7. Sridharan, S. A., Ramrao, N., Murugan, K., Prasad, V., Krishna, R., & Nair, R. (2020). The Significant Contextual Predictors of Industrial Academic Collaboration in Engineering. *Journal of Engineering Education Transformations*, 34(Special Issue).
8. Kushwah, J. S., Kumar, A., Patel, S., Soni, R., Gawande, A., & Gupta, S. (2022). Comparative study of regressor and classifier with decision tree using modern tools. *Materials Today: Proceedings*, 56, 3571–3576.
9. LaValley, M. P. (2008). Logistic regression. *Circulation*, 117(18), 2395–2399.
10. Popular Machine Learning Algorithms https://www.simplilearn.com/10-algorithms-machine-learning-engineers-need-to-know-article#:~:text=This%20is%20what%20linear%20regression,%3D%20a%20*X%20%2B%20b.
11. Reddy, V. R., Yakobu, D., Prasad, S. S., & Vardhini, P. H. (2022, March). Clustering Student Learners based on performance using K-Means Algorithm. In *2022 International Mobile and Embedded Technology Conference (MECON)* (pp. 302–306). IEEE.
12. Sharma, D., & Kumar, N. (2017). A review on machine learning algorithms, tasks and applications. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 6(10), 2278–2323.
13. Jain, A., & Solanki, S. (2019, July). An efficient approach for multiclass student performance prediction based upon machine learning. In *2019 International Conference on Communication and Electronics Systems (ICCES)* (pp. 1457–1462). IEEE.
14. Hussain, A. A., & Dimililer, K. (2021). Student grade prediction using machine learning in Iot era. In *Forthcoming Networks and Sustainability in the IoT Era: First EAI International Conference, FoNeS–IoT 2020, Virtual Event, October 1–2, 2020, Proceedings 1* (pp. 65–81). Springer International Publishing.

A Novel Approach in Machine Learning for Solar Energy Prediction System



Supriya Vaddi, Fathima Sumreen, Vaishnavi Narayanam,
Akshara Patlan-nagari, and Peruru Ananya

Abstract India, a rapidly growing economy with a population exceeding one billion, faces a significant demand for energy. Despite the increasing population, power generation in the country has also seen a rise. To address India's escalating energy requirements, harnessing solar energy emerges as the most suitable solution, taking advantage of the country's favorable geographical location. However, effectively utilizing the solar energy generated presents challenges. While organizations invest effort in generating solar energy, they often neglect the importance of utilizing it optimally. This oversight can result in financial losses for organizations, as they invest substantial amounts in setting up solar energy systems but fail to monitor their utilization. The main objective of this study is to analyze the usage of solar energy and predict the amount of energy that our institution, G. Narayanamma Institute of Technology and Sciences, would generate. The study uses four distinct algorithms, namely multiple linear regression, decision trees, gradient boost, and XGBoost. We have evaluated these algorithms and identified the most accurate model based on their performance.

Keywords Machine learning · Solar energy · Multiple linear regression · Decision trees · Gradient boost · XGBoost

1 Introduction

Solar power is a renewable and environmentally friendly energy source that holds significant potential for combating climate change and achieving sustainable energy practices. Accurate forecasting of solar energy production is crucial as it plays a vital role in optimizing the integration of solar power into the power grid and ensuring a reliable energy supply. Precise forecasts enable effective management

S. Vaddi (✉) · F. Sumreen · V. Narayanam · A. Patlan-nagari · P. Ananya
Department of Information Technology, GNITS, Hyderabad, Telangana, India
e-mail: supriya-vaddi@gnits.ac.in

of energy generation, storage, and distribution, thereby enhancing grid stability and reducing reliance on traditional energy sources. Utilities need to plan effectively to maintain grid stability as solar power penetration increases. This study focuses on forecasting solar energy production for a period of 14 days or 2 weeks. Our research specifically examines the prediction of solar energy using carefully selected small-to-medium-sized solar panels installed at our institution, G. Narayanamma Institute of Technology and Science. We analyze the energy generated by Polycrystalline solar panels that were reinstalled in 2019.

The combined capacity of 81 solar panels is 30 kWh, with each panel connected to a 160-kW inverter. Rooftop solar energy is gaining popularity, and precise forecasting is essential to fully leverage the benefits it offers. Our study employs machine learning techniques that utilize publicly available weather predictions to forecast solar intensity. By utilizing these predicted sun intensity values, we can estimate the amount of solar energy generated based on the assumption that intensity directly correlates with energy production. This approach offers several sustainable economic advantages, including potential subsidies on electricity bills through surplus energy supply to the grid. Such benefits contribute to the economic viability, sustainability, and growth of the energy sector.

2 Literature Survey

The paper titled “A Hybrid Approach of Solar Power Forecasting Using Machine Learning” [1] investigates solar power forecasting by employing multiple machine learning models and analyzing various weather parameters. The accuracy of these models is evaluated using historical weather data. This approach combines pipelining, clustering, classification, and regression algorithms to develop a unique photovoltaic (PV) forecasting model. By grouping data points with similar weather conditions and utilizing segmented regression, the model leverages the next-day weather forecast to enhance prediction accuracy. The hybrid model, which integrates clustering, classification, and regression techniques, surpasses the performance of the random forest regression model, demonstrating improved forecasting capabilities under specific weather conditions.

In another study [2], a mathematical approach is introduced to predict solar energy generation for the upcoming 7 days. This model takes into account factors such as ambient temperature, solar irradiance, panel efficiency, and plant component efficiency by considering weather data and plant specifications. The analysis provides insights into how temperature, irradiance, plant design, and efficiency affect the output of a photovoltaic system. However, the model’s vulnerability to uncertainties arising from weather conditions is a concern.

Additionally, a research paper [3] presents an artificial neural network model for solar power forecasting. The study conducts a sensitivity analysis of different input variables to determine the optimal selection and compares the model’s performance with multiple linear regression and persistence models. The objective is to generate

hourly solar power forecasts for a month-long timeframe. The European Center for Medium-Range Weather Forecasts (ECMWF) serves as the source for 12 independent variables used to predict the target variable, solar power.

3 Dataset

Data from the college portal “Solaris” was collected between May 2022 and January 2023 to monitor the generation of solar energy. The solar energy produced by the panels was recorded at 5-min intervals on a daily basis, along with various attributes such as voltage, current, total power, frequency, and power grid total apparent power. The collected data was then aggregated on an hourly basis. In conjunction with the solar energy dataset, weather data from the same time intervals was obtained using the Visual Crossing OpenWeather data API.

4 Methodology

The proposed system aims to forecast solar energy production for the next 14 days. To build the system, various algorithms including multiple linear regression, decision tree, gradient boost, and XGBoost were compared.

Figure 1 illustrates the system’s architecture, which involves the use of solar panels to generate energy that is then sent to an inverter. Data related to the solar panels is stored by a data logger and sent to the solar energy prediction system along with weather data obtained from an external website. The system creates a machine learning model and displays the forecasted solar energy for the selected date on the user interface (UI). Users interact with the system through the UI.

Several assumptions were made for the analysis. Since the college collects data only from one building with solar panels, it is assumed that the energy generated by panels in other buildings is similar, as they were installed at the same time. The accuracy of the weather data collected from the Visual Crossing Weather API is also crucial for the accuracy of the outcomes.

In the data preprocessing phase, irrelevant attributes such as inverter status code, inverter serial number, data logger serial number, alert details, alert code, and grid power factor were removed. The 5-min time intervals in the solar dataset were converted to 1-h intervals to align with the weather data. Outlier analysis was conducted to identify and exclude days with fewer than the average number of recorded time intervals in a day. The common time intervals during which solar energy was generated, specifically from 6:00:00 to 19:00:00, were identified. Finally, the weather dataset was merged with the solar dataset.

Feature Engineering In our analysis, we utilized both correlation matrix and mutual information score to identify the relevant attributes. The correlation matrix

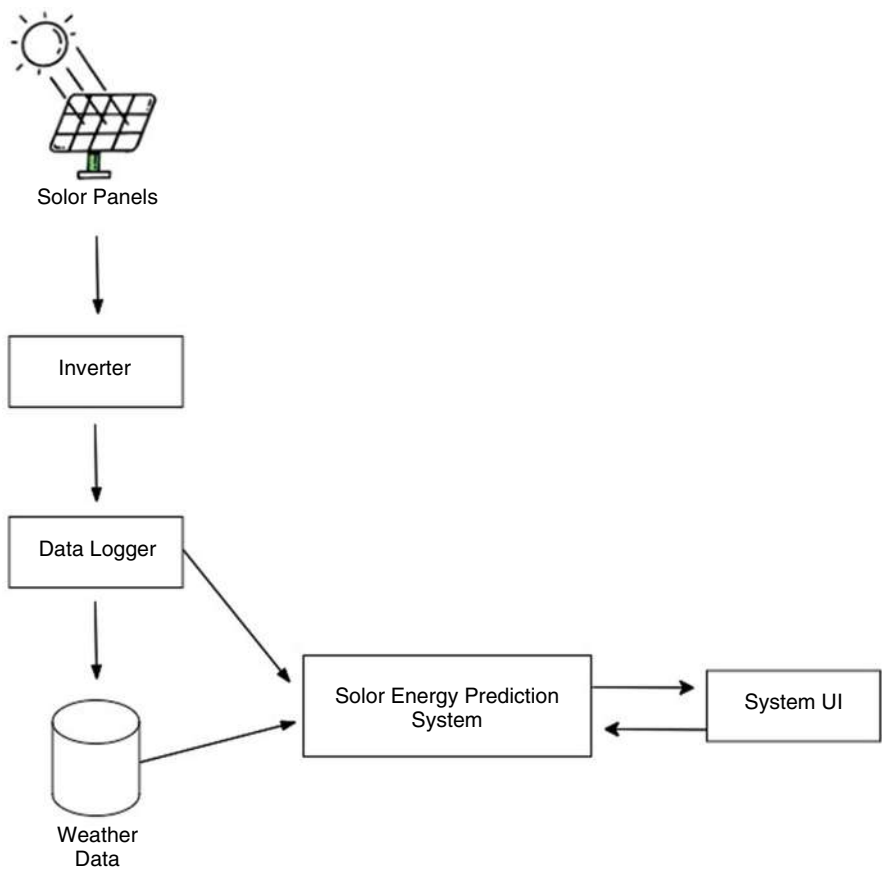


Fig. 1 Architecture for solar energy prediction system

allowed us to measure the linear relationships between each attribute, while the mutual information score helped us identify nonlinear relationships. After conducting these analyses, we determined that the important attributes for our study included time, temperature, dew, humidity, wind direction, cloud cover, visibility, and solar radiation.

Model Selection To predict the daily solar power generation, we employed various machine learning models, namely multiple linear regression, decision tree, gradient boost, and XGBoost. These models were selected based on their compatibility with the dataset and their ability to forecast solar energy up to 14 days in advance. Ultimately, we chose XGBoost as our preferred model due to its highest accuracy, reaching approximately 81%.

Performance Metric We evaluated the performance of our models using two commonly used evaluation metrics for regression models: R-squared and mean-

squared error. These metrics allowed us to assess the goodness of fit and the accuracy of predictions. After considering these metrics, we selected the model with the highest R- squared value and the lowest mean-squared error.

Prediction Our model is designed to forecast the daily solar power generation for the next 14 days, starting from the current day. This prediction is based on obtaining forecasted weather data through an API call. Additionally, we developed a user interface using HTML, CSS, and Flask to effectively display the forecasted results.

5 Results and Discussion

To assess the models’ performance, we initially divided the dataset into 75% for training and 25% for testing. We then utilized the R-squared value and mean-squared error to evaluate the models’ performance. The R-squared value, as defined by Eq. (1), quantifies the proportion of variance in the dependent variable that can be attributed to the independent variable.

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \tag{1}$$

Mean-squared error was additionally used to compare the models further. The mathematical formula of mean-squared error is given below by Eq. (2). Mean-squared error assesses the average squared difference between the observed and predicted values.

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2 \tag{2}$$

Table 1 summarizes our obtained result. According to the provided R-squared values, the XGBoost algorithm demonstrates the highest R-squared value, indicating the best fit to the data compared to the other algorithms considered. The decision tree algorithm performs the second best, followed by multiple linear regression and gradient boost. The superior performance

Table 1 Result from multiple algorithms

S. no.	Algorithm	R-squared value	Mean-squared error
1	Multiple linear regression	0.7704	393.6623
2	Decision tree	0.7881	363.3103
3	Gradient boost	0.7721	390.7736
4	XGBoost	0.8168	314.4238

of XGBoost can be attributed to its capability to enhance weaker learners through effective parallel computing methods.

6 Conclusion and Future Scope

The chapter is aimed at predicting the solar energy generated using machine learning algorithms. We found out that XGBoost gives the best performance with an accuracy of 81%. However, this study emphasizes on the different weather attributes and their importance in solar energy generation, and one of the vital limitations of the study is that it depends upon the weather forecasted by the third-party website. Hence, more the reliability of the weather data, the higher is the accuracy of the study. The major limitation of this study is the limited dataset. Accuracy can be further improved with the help of larger dataset. This will also ensure better prediction results. The study can be extended by using a hybrid model to categorize similar weather patterns. We also believe that model fitting upon the clusters of data can lead to a better and higher accuracy in performance. Automating the whole system can make it effective in real time.

References

1. A. Bajpai and M. Duchon, "A Hybrid Approach of Solar Power Forecasting Using Machine Learning," 2019 3rd International Conference on Smart Grid and Smart Cities (ICSGSC), 2019.
2. D. Solanki, U. Upadhyay, S. Patel, R. Chauhan and S. Desai, "Solar Energy Prediction using Meteorological Variables," 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE), 2018.
3. M. Abuella and B. Chowdhury, "Solar power forecasting using artificial neural networks," 2015 North American Power Symposium (NAPS), 2015.
4. M. Yesilbudak, M. Çolak and R. Bayindir, "A review of data mining and solar power prediction," 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, UK, 2016, pp. 1117-1121, <https://doi.org/10.1109/ICRERA.2016.7884507>.
5. K. A. Dotche, A. A. Salami, K. M. Kodjo, Y. P. C. D. Blu and Y. E. J. Diabo, "Evaluating Solar Energy Harvesting using Artificial Neural Networks: A Case study in Togo," 2019 II International Conference on High Technology for Sustainable Development (HiTech), Sofia, Bulgaria, 2019, pp. 1-5, <https://doi.org/10.1109/HiTech48507.2019.9128285>.
6. A. Dhage, A. Kakade, G. Nahar, M. Pingale, S. Sonawane and A. Ghotkar, "Recommendation and Prediction of Solar energy consumption for smart homes using machine learning algorithms," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1-5, <https://doi.org/10.1109/AIMV53313.2021.9670909>.
7. Rahul, A. Gupta, A. Bansal and K. Roy, "Solar Energy Prediction using Decision Tree Regressor," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 489-495, <https://doi.org/10.1109/ICICCS51141.2021.9432322>.

8. A. Tuohy et al., "Solar Forecasting: Methods, Challenges, and Performance," in *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 50-59, Nov.-Dec. 2015, <https://doi.org/10.1109/MPE.2015.2461351>.
9. K. Pun, S. M. Basnet and W. Jewell, "Solar Power Prediction in Different Forecasting Horizons Using Machine Learning and Time Series Techniques," 2021 IEEE Conference on Technologies for Sustainability (SusTech), Irvine, CA, USA, 2021, pp. 1-7, <https://doi.org/10.1109/SusTech51236.2021.9467464>.
10. Kim, Seul-Gi, Jae-Yoon Jung, and Min Kyu Sim. 2019. "A Two-Step Approach to Solar Power Generation Prediction Based on Weather Data Using Machine Learning" *Sustainability* 11, no. 5: 1501.
11. Mohammed, Azhar & Aung, Zeyar. (2016). Ensemble Learning Approach for Probabilistic Forecasting of Solar Power Generation. *Energies*. 9. 1017. <https://doi.org/10.3390/en9121017>.
12. Gensler-Janosch, A., et al. "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks." 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.
13. Hammer, A., et al. "Short-Term forecasting of solar radiation: a statistical approach using satellite data." *Solar Energy*, vol. 67, no. 1-3, 1999, pp. 139-150, [https://doi.org/10.1016/S0038-092X\(00\)00038-4](https://doi.org/10.1016/S0038-092X(00)00038-4).
14. Mellit, Adel. "Artificial Intelligence technique for modelling and forecasting of solar radiation data: a review." *International Journal of Artificial Intelligence and Soft Computing*, vol. 1, no. 1, 2008, p. 52., <https://doi.org/10.1504/ijaisc.2008.021264>.
15. Perera, Kasun S., et al. "Machine Learning Techniques for Supporting Renewable Energy Generation and Integration: A Survey." *Data Analytics for Renewable Energy Integration Lecture Notes in Computer Science*, 2014, pp. 81-96., <https://doi.org/10.1007/978-3-319-13290-77>
16. Srivistava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *Journal of Machine Learning Research* 15 (2014).

Real-Time Tomato Leaf Disease Detection and Diagnosis Using Deep Learning-Based Computer Vision Techniques



S. Kanakaprabha, P. Arulprakash, G. Ganesh Kumar, T. Udhayakumar, and R. Janani

Abstract Tomato plants are vulnerable to numerous diseases, which can have significant impact their growth and yield. Early discovery and accurate analysis of these diseases are critical for implementing timely and operative control measures. In modern years, deep learning-based computer visualization methods have emerged as powerful tools for real-time tomato leaf disease recognition and diagnosis. The system utilizes a convolutional neural network (CNN) model to repeatedly categorize tomato leaves as healthy or diseased based on images taken in the field. The model is skilled on a dataset of labeled imageries of tomato leaves, with each image containing a specific disease or being healthy. The system is implemented on a portable device for real-time use in the field, allowing for rapid and accurate diagnosis of tomato leaf diseases. The investigational results show that the suggested system achieved high accuracy in disease classification, with an average precision of 98.11%. Furthermore, the system also successfully detects multiple diseases present in the same leaf, which is a significant improvement

S. Kanakaprabha (✉) · T. Udhayakumar

Department of Computer Science and Engineering, Rathinam Technical Campus, Anna University, Coimbatore, India

P. Arulprakash

Department of Computer Science and Engineering, Rathinam Technical Campus, Anna University, Coimbatore, India

Department of Information Technology, Rathinam Technical Campus, Anna University, Coimbatore, India

e-mail: rtchod.cse@rathinam.in; ganeshkumar.it@rathinam.in

G. Ganesh Kumar

Department of Information Technology, Rathinam Technical Campus, Anna University, Coimbatore, India

e-mail: rtchod.cse@rathinam.in; ganeshkumar.it@rathinam.in

R. Janani

Department of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

725

A. Patel et al. (eds.), *Advances in Machine Learning and Big Data Analytics I*,

Springer Proceedings in Mathematics & Statistics 441,

https://doi.org/10.1007/978-3-031-51338-1_58

over traditional methods that often fail to detect multiple diseases. The proposed real-time tomato leaf disease detection and diagnosis system using deep learning-based computer vision procedures can effectively detect and diagnose multiple diseases in tomato leaves in real time. This technology can be used to improve the effectiveness and correctness of disease discovery, ultimately reducing the use of harmful pesticides and increasing crop yields. This chapter highlights the potential impact of real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision performances. Early detection and diagnosis can facilitate timely interventions, leading to improved disease management, increased crop yield, and reduced reliance on chemical treatments.

Keywords Automatically · Deep learning · Diagnosis · Leaf disease · Tomato · Yields

1 Introduction

Traffic tomato is one of the most extensively grown vegetables in the world, and it is an essential source of nutrition for millions of people. However, tomato crops are vulnerable to various diseases, which can result in significant losses in yield and quality. To address this challenge, computer vision techniques based on deep learning have been projected for real-time tomato leaf disease detection and diagnosis. These techniques use high-resolution imageries of tomato leaves to identify signs of disease and provide accurate diagnosis. The deep learning models are trained on a huge dataset of images of diseased and healthy tomato leaves, allowing them to learn the visual features that distinguish between healthy and diseased plants. The models can then be used to categorize new images of tomato leaves, helping farmers and researchers to identify disease outbreaks and implement appropriate control measures quickly and accurately. The use of deep learning-based computer visualization methods for real-time tomato leaf disease detection and diagnosis has the probable to transfigure the way that growers and researchers accomplish tomato crops and protect against disease. Tomato crop sicknesses can have a devastating impression on small farmers and their livelihoods. Early discovery and analysis of these diseases are serious for minimizing their impact and preserving the health of the crops. However, traditional methods of disease diagnosis, such as manual inspection and laboratory analysis, can be time-consuming, labor-intensive, and often unreliable. This is where the use of deep learning-based computer vision techniques can provide significant benefits. These techniques can speedily and precisely diagnose diseases based on images of the tomato leaves, reducing the time and effort required to diagnose diseases compared to traditional methods.

Therefore, the development and implementation of these techniques has the probable to play an important role in reducing poverty and improving the livelihoods of small farmers, as well as improving the agricultural economy as a whole [1]. In

this research, we aim to advance a system for real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques.

2 Literature Survey

A. Lakshmanarao et al. [2]: The use of ConvNets for vegetal disease detection and classification has proven to be a successful approach, as demonstrated by your results. Achieving an correctness of 98.3% for potato plant disease detection, 98.5% for pepper plant disease detection, and 95% for tomato plant disease detection is a remarkable accomplishment. The ability to accurately diagnose plant diseases using images of the leaves is critical for reducing the spread of diseases and improving crop yields. The use of deep learning-based computer vision techniques has the potential to transform the way that plant diseases are diagnosed and managed, making it easier and more efficient for farmers and researchers to protect their crops. The potential to make significant impact on the agriculture sector, helping to reduce poverty and improve the livelihoods of small farmers by reducing crop losses and improving the quality and yield of crops.

G. Irmak et al. [3]: Deep learning-based computer image techniques for plant disease detection and organization has the potential to be extended to other agricultural crops, not just tomatoes. This would greatly increase the potential impact of this research on the agriculture sector and help to reduce crop losses and improve the quality and yield of a wider range of crops. This type of robot could transfigure the way that plant diseases are diagnosed and managed, making it easier and more efficient for farmers and researchers to protect their crops. The potential of deep learning-based computer vision techniques for educating the accuracy and effectiveness of plant disease diagnosis. The development of such techniques can greatly contribute to reducing poverty and improving the livelihoods of small farmers by reducing crop losses and improving the quality and yield of crops.

S. Ashok et al. [4]: Early uncovering of plant leaf diseases is indeed crucial for an agricultural economy like India, where both agricultural production and population size place a high demand on food security. By using duplicate processing techniques, such as image subdivision, clustering, and open-source procedures, to identify tomato plant leaf diseases, the proposed system has the potential to provide a consistent, harmless, and accurate method of disease diagnosis. The potential to make a significant impact on the agricultural sector in India, helping to progress food security by dropping crop sufferers and cultivating the quality and yield of crops. The development of a reliable, safe, and accurate system for diagnosing tomato plant leaf diseases can help farmers and researchers to better protect their crops and contribute to a sustainable and thriving agricultural economy.

S. V. Militante et al. [5]: The advancements in computer apparition and deep learning have unlocked up new possibilities for the discovery and diagnosis of plant diseases. By using deep learning models qualified on a large dataset of images of healthy and diseased plant leaves, the system is able to achieve a high

accuracy rate of 96.5% in perceiving and knowing the plant variation and type of disease the plant is infected with. The ability of the system to detect and recognize multiple plant diversities and ailments has the potential to greatly simplify the process of diagnosing plant diseases, making it easier and more efficient for farmers and researchers to protect their crops. The potential of computer vision and deep learning techniques for improving the accuracy and efficiency of plant disease diagnosis. The development of such techniques can greatly contribute to reducing crop losses and educating the superiority and yield of crops, and help to confirm food security for growing populations like in India.

H. E. David et al. [6]: Early detection of diseases in tomato crops is crucial for preventing economic losses and improving crop production. The traditional disease detection methods have limitations in terms of accuracy and speed, and therefore, the adoption of processor vision-based knowledge and deep learning techniques can be a game-changer in this regard. Deep learning algorithms can procedure large quantities of data rapidly and correctly, making them well suited for the discovery of plant diseases. This early uncovering of ailments can help farmers take prompt action to avoid the spread of diseases and diminish their impact on crop production. The implementation of computer vision-based technology and deep learning techniques can also reduce the manual effort required for disease detection, making the process faster and more efficient. The tropical climate of many countries makes them ideal for tomato cultivation, and the implementation of these techniques can help to further improve the growth and productivity of tomato crops in these regions.

K. Karthik et al. [7]: The planned system aims to provide an explanation for early detection of tomato plant leaf diseases using image processing and deep learning techniques. By using these techniques, the system can accurately identify the plant leaf diseases and provide preventive measures, reducing economic loss and improving crop production. The use of computer vision-based knowledge in agriculture is still in its early stages, but it has the probable to revolutionize the way farmers monitor their crops and detect diseases. The proposed system not only helps farmers but also provides a common knowledge base for people interested in agriculture and helps to improve the overall efficiency of crop production.

3 Methodology

3.1 Dataset Description

The dataset consists of ten types of diseases that commonly affect tomato plants, with each disease class having a uniform distribution of samples. The ten diseases included in the dataset are Bacterial Spot, Healthy, Early Blight, Leaf Mold, Late Blight, Septoria Leaf Spot, Mosaic Virus, Spider Mites, Target Spot, and Yellow Leaf Curl Virus. The dataset contains a total of 1000 images, with 100 images

for each disease class. The images are captured under different lighting conditions and angles, and may contain single or multiple instances of the disease in Fig. 1. Researchers can use this dataset to train machine learning models to classify different tomato diseases accurately, which can be beneficial for plant disease diagnosis and management.

3.2 Preprocessing

Tomato detection is a critical step in the real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques shows in Fig. 2. The preprocessing stage involves various steps to prepare the input data for efficient processing by the deep learning models. The first step is to capture high-quality images of tomato leaves using a digital camera or a smartphone camera. These images are then resized and normalized to a standard size to ensure consistency across all images. Next, image enhancement techniques are applied to improve the quality of the images, such as adjusting brightness and contrast, removing noise, and

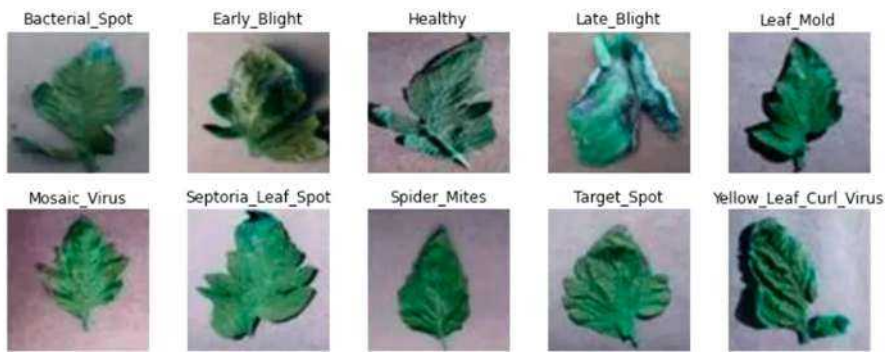


Fig. 1 Dataset images for ten disease class

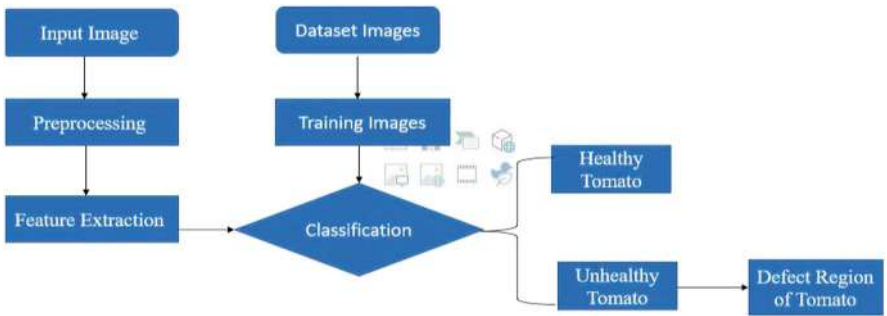


Fig. 2 Proposed methodology of architecture

sharpening edges. This helps to improve the accuracy of the deep learning models in detecting and diagnosing tomato leaf diseases.

3.3 Feature Extraction

In the real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques, feature extraction refers to the process of identifying and extracting relevant visual features from tomato leaves to enable disease detection. These visual features may include color, texture, shape, and size characteristics of the leaves, which are analyzed using various deep learning models and computer vision techniques to identify signs of disease. By accurately extracting and analyzing these features, the system can quickly and accurately detect and diagnose tomato leaf diseases, potentially enabling more efficient and effective crop management practices.

3.4 Classification

Tomato detection using deep learning-based computer vision techniques has become a popular approach for identifying normal and abnormal tomatoes in real time. By analyzing images of tomato plants, these techniques can distinguish between healthy and diseased plants, which can help farmers to identify and treat plant diseases before they spread. The process involves capturing images of tomato plants using cameras, which are then processed by deep learning algorithms. These algorithms can detect patterns and features in the images that are indicative of normal or abnormal tomato plants. For example, healthy tomato plants have green leaves and smooth surfaces, while diseased plants may have yellow or brown patches and deformed leaves. By accurately detecting abnormal tomatoes, farmers can take action to prevent the spread of diseases and increase their crop yield. This technology has the potential to revolutionize the agricultural industry by allowing farmers to quickly and easily identify and diagnose plant diseases, ultimately leading to more efficient and sustainable farming practices.

3.5 Defect Region Classified

The defect region can help in identifying the specific type of disease, which can aid in the selection of appropriate measures for disease control. Computer vision-based techniques can be used to detect the defect region in tomato leaves. These techniques involve segmenting the tomato leaf image into different regions, each representing a different part of the leaf. The segmented regions can then be classified as normal

or abnormal based on their color, texture, and other features. Deep learning-based computer vision techniques, such as convolutional neural networks (CNNs), can be trained on a large dataset of tomato leaf images to accurately detect the defect region in the leaves. Once the defect region is identified, it can be classified into different categories, depending on the type of disease present. This information can be used to select appropriate measures for disease control.

4 Result and Analysis

Real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques has shown promising results in the accurate and efficient identification of tomato leaf diseases. The use of deep learning-based approaches such as CNNs has allowed for the development of highly accurate and efficient models for disease detection.

Early detection of diseases can help in the selection of appropriate measures for disease control, reducing the need for harmful chemicals and pesticides. The use of deep learning-based computer vision techniques for real-time tomato leaf disease detection and diagnosis has shown promising results in accurately and efficiently identifying tomato leaf diseases. The high accuracy of these models can help in the early detection and management of diseases, leading to increased crop yield and sustainability.

The use of deep learning-based computer vision techniques, shown in Fig. 3, for the real-time tomato leaf disease detection and diagnosis has shown promising results. In the training phase, a large dataset of tomato leaf images is used to train the CNNs to accurately detect and classify different types of diseases. The trained CNNs can then be used in the testing phase to accurately detect and diagnose diseases in real time.

The results of the training and testing phases are typically presented in the form of graphs or tables. These graphs or tables show the accuracy of the CNNs

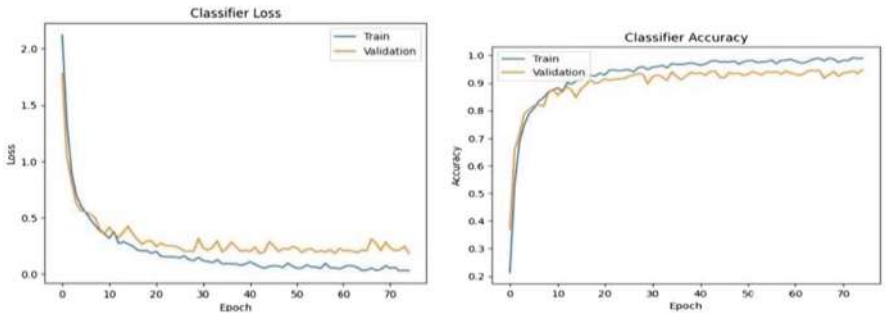


Fig. 3 Training and testing of tomato detection

in detecting and classifying different types of diseases. The graphs typically show an increasing trend in accuracy with increasing epochs or iterations of the training process. The accuracy of the CNNs can be further improved by using data augmentation techniques such as rotation, translation, and flipping of the images.

The analysis of the results of real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques can provide valuable insights into the effectiveness of these techniques.

The confusion matrix shown in Fig. 4 represents the evaluation of the deep learning-based computer vision technique used for real-time tomato leaf disease detection and diagnosis. The matrix provides a visual representation of the performance of the model in predicting the presence or absence of tomato diseases. The rows of the matrix correspond to the true labels or the actual disease status of the tomatoes, while the columns represent the predicted labels or the disease classification made by the model. The matrix is divided into the following four quadrants:

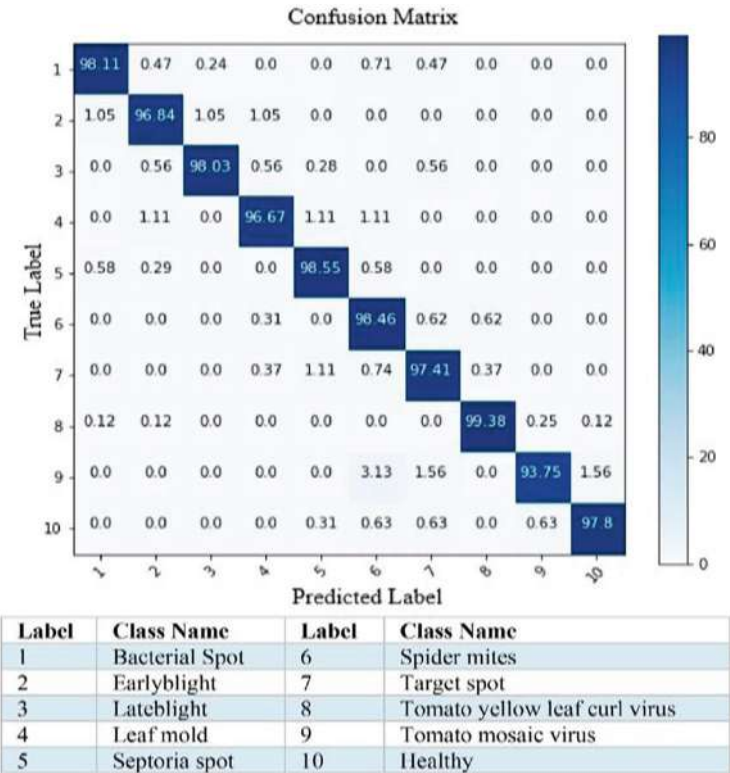


Fig. 4 Confusion matrix for tomato disease detection

True Positive (TP): This quadrant represents the correctly classified instances where the model correctly detects and diagnoses the presence of a disease in the tomatoes.

True Negative (TN): This quadrant represents the correctly classified instances where the model accurately identifies and diagnoses healthy tomatoes without any disease.

False Positive (FP): Also known as Type I error, this quadrant represents the instances where the model incorrectly predicts the presence of a disease in the tomatoes, even though they are healthy.

False Negative (FN): Also known as Type II error, this quadrant represents the instances where the model fails to detect and diagnose the disease in the tomatoes, incorrectly classifying them as healthy.

5 Conclusion

Real-time tomato leaf disease detection and diagnosis using deep learning-based computer vision techniques have the potential to significantly improve crop management practices. This approach enables farmers and agricultural experts to identify plant diseases quickly and accurately, thereby reducing crop loss and improving overall crop yield. By using deep learning algorithms, the computer vision system can learn to recognize specific patterns and features associated with various plant diseases. This approach has been shown to achieve high accuracy rates in detecting and diagnosing tomato leaf diseases in real time. Furthermore, this technology can be easily integrated with existing agricultural equipment and can operate in real time, allowing farmers to take immediate action to prevent further spread of diseases. Additionally, the system can be trained to detect multiple diseases simultaneously, providing a comprehensive diagnosis of plant health.

References

1. Trivedi, Naresh K., Vinay Gautam, Abhineet Anand, Hani Moaiteq Aljahdali, Santos Gracia Villar, Divya Anand, Nitin Goyal, and Seifedine Kadry. 2021. "Early Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network" *Sensors* 21, no. 23: 7987. <https://doi.org/10.3390/s21237987>.
2. A. Lakshmanarao, M. R. Babu and T. S. R. Kiran, "Plant Disease Prediction and classification using Deep Learning ConvNets," *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, Gandhinagar, India, 2021, pp. 1–6, doi: <https://doi.org/10.1109/AIMV53313.2021.9670918>.
3. G. IRMAK and A. SAYGILI, "Tomato Leaf Disease Detection and Classification using Convolutional Neural Networks," *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, 2020, pp. 1–5, doi: <https://doi.org/10.1109/ASYU50717.2020.9259832>.

4. S. Ashok, G. Kishore, V. Rajesh, S. Suchitra, S. G. G. Sophia and B. Pavithra, "Tomato Leaf Disease Detection Using Deep Learning Techniques," *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2020, pp. 979–983, doi: <https://doi.org/10.1109/ICCES48766.2020.9137986>.
5. S. V. Militante, B. D. Gerardo and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," *2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, Yunlin, Taiwan, 2019, pp. 579–582, doi: <https://doi.org/10.1109/ECICE47484.2019.8942686>.
6. H. E. David, K. Ramalakshmi, H. Gunasekaran and R. Venkatesan, "Literature Review of Disease Detection in Tomato Leaf using Deep Learning Techniques," *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2021, pp. 274–278, doi: <https://doi.org/10.1109/ICACCS51430.2021.9441714>.
7. K. Karthik, S. Rajaprakash, S. Nazeem Ahmed, R. Perincheeri and C. R. Alexander, "Tomato And Potato Leaf Disease Prediction With Health Benefits Using Deep Learning Techniques," *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, 2021, pp. 1–3, doi: <https://doi.org/10.1109/I-SMAC52330.2021.9640765>.
8. H. Hong, J. Lin and F. Huang, "Tomato Disease Detection and Classification by Deep Learning," *2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, Fuzhou, China, 2020, pp. 25–29, doi: <https://doi.org/10.1109/ICBAIE49996.2020.00012>.
9. A. S. Chakravarthy and S. Raman, "Early Blight Identification in Tomato Leaves using Deep Learning," *2020 International Conference on Contemporary Computing and Applications (IC3A)*, Lucknow, India, 2020, pp. 154–158, doi: <https://doi.org/10.1109/IC3A48958.2020.233288>.
10. A. Elhassouny and F. Smarandache, "Smart mobile application to recognize tomato leaf diseases using Convolutional Neural Networks," *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)*, Agadir, Morocco, 2019, pp. 1–4, doi: <https://doi.org/10.1109/ICCSRE.2019.8807737>.
11. E. Suryawati, R. Sustika, R. S. Yuwana, A. Subekti and H. F. Pardede, "Deep Structured Convolutional Neural Network for Tomato Diseases Detection," *2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, Yogyakarta, Indonesia, 2018, pp. 385–390, doi: <https://doi.org/10.1109/ICACSIS.2018.8618169>.
12. H. I. Peyal, S. M. Shahriar, A. Sultana, I. Jahan and M. H. Mondol, "Detection of Tomato Leaf Diseases Using Transfer Learning Architectures: A Comparative Analysis," *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, Rajshahi, Bangladesh, 2021, pp. 1–6, doi: <https://doi.org/10.1109/ACMI53878.2021.9528199>.

Index

A

ACARS, 566
 Accuracy, 22, 29, 50, 58, 74, 88, 113, 126, 162, 181, 186, 214, 243, 263, 286, 305, 317, 326, 339, 348, 358, 369, 381, 394, 407, 419, 432, 453, 458, 498, 507, 528, 542, 566, 578, 602, 623, 632, 646, 685, 696, 704, 718, 727
 Active power, 459, 461, 472, 474, 475
 Adaptive boosting (ADaBoost), 242, 243, 245, 246, 250–251, 255, 256, 309–311, 433, 438–441, 603, 704, 709, 712
 Adaptive thresholding, 101, 102
 ADC pin, 506
 Ad hoc On-Demand Distance Vector (AODV), 134–136, 143–148, 153, 155–157, 159
 Advanced ML approach, 417–426, 642
 Adverse drug reactions (ADRs), 210, 211, 214
 Agriculture, 48, 49, 55, 58, 120, 338, 408, 498–500, 616, 645, 685, 727, 728
 Aircraft, 277, 566, 568, 569
 Air quality index (AQI), 260, 261, 263–266, 268, 270
 Algorithms, 2, 14, 28, 50, 62, 74, 88, 103, 120, 134, 152, 163, 174, 186, 215, 228, 242, 260, 286, 309, 316, 326, 338, 348, 368, 381, 392, 406, 420, 431, 446, 468, 483, 498, 505, 528, 542, 566, 586, 603, 618, 633, 646, 677, 684, 698, 704, 718, 728
 A Mobile Ad-hoc Network (MANET), 134–150, 152–154, 159
 Android OS, 674, 675
 ANSYS, 277

Antenna, 273–282
 Arduino, 499, 506, 508, 698
 Artificial intelligence (AI), 74, 190–192, 195, 247, 248, 285–300, 315, 321, 349, 476, 530, 557, 560, 566, 578, 580, 619, 632
 Artificial neural network (ANN), 74, 242, 243, 245–247, 249–250, 255, 256, 288, 317, 318, 327, 359, 393, 407, 412, 433, 437–438, 440, 441, 469, 476, 532, 535, 538, 539, 603, 705, 718
 Atrial septal defect (ASD), 75, 76, 79, 80, 84
 Attention, 27, 76–83, 121, 123, 124, 176, 202, 358, 370, 542, 552–557, 559, 562, 657, 671, 697–699
 Attention mechanism, 73–84, 112, 123, 124, 369, 370, 552, 553, 555, 559, 610
 Attribute-based encryption (ABE), 226, 227
 Attrition, 429–442
 Auscultation, 74, 76
 Authentication system, 492, 494
 Author name disambiguation, 89
 Automatically, 28, 32, 48, 50, 51, 88, 142, 163, 174, 242, 248, 252, 286, 318, 358, 359, 380, 394, 446, 458, 471, 500, 505, 534, 542, 615, 646
 Automatic generation of medical reports, 544
 Average precision (AP), 62, 64

B

Bibliometric analysis, 90, 91
 BiLSTM, 75, 77–79, 83

Biological features, 242
 BLOSUM62 matrix, 2, 4, 10
 Bone fracture, 325–336

C

Camera, 300, 340, 345, 408, 446–448, 450, 451, 492, 495, 531, 604, 729, 730
 Causal language modeling (CLM), 370, 372, 374
 Citation analysis, 89, 90
 Classification, 9, 52, 62, 74, 89, 106, 121, 141, 162, 175, 190, 215, 243, 260, 287, 304, 318, 326, 340, 350, 358, 370, 383, 393, 407, 425, 433, 473, 505, 533, 542, 567, 578, 602, 619, 633, 646, 704, 718, 727
 Cloud security, 226
 Cloud storage, 226, 227, 229, 230
 Clustering, 3, 52, 250, 320, 369, 373, 407, 623, 647, 684, 685, 688, 690–692, 718, 727
 Clustering protocols, 683, 684, 692
 Cognitive state monitoring, 695
 Collaborative filtering, 305, 308, 311, 319
 Colluding attack, 674–677, 681
 Computational power, 315
 Computer, 23, 27, 48, 49, 198, 201, 202, 216, 226, 243, 287, 349, 381, 384, 421, 423–425, 445–448, 454, 491, 492, 509, 580, 614, 615, 632, 668, 697, 726, 727
 Congenital heart disease, 73–84
 Content-based filtering, 305, 307, 308
 Context, 32, 48–51, 103, 109, 110, 113, 152, 164, 266, 289, 304, 321, 329, 334, 341, 369, 370, 372, 411, 432, 439, 490, 555, 580, 603, 618, 632, 670, 698, 710
 Control algorithms, 467–477
 Conversion, 247, 469, 472, 476, 512, 657–671
 Convolution(al) neural networks (CNNs), 31, 32, 50, 75, 77, 83, 106, 175, 176, 179–182, 187, 288, 295, 317–319, 326, 327, 329–336, 339, 340, 357–364, 381, 382, 385, 389, 393, 407, 408, 410, 412, 528, 530, 536, 537, 543, 558–561, 602, 603, 619, 633, 647, 731, 732
 Correlation, 75, 77, 244, 245, 266, 267, 393, 436, 535, 582, 620, 670, 711, 712, 719
 COVID-19, 212, 631–642

D

Data confidentiality, 236

Data gathering, 152, 189, 410, 619, 628, 684, 692, 693
 Data mining, 2, 185–192, 420, 705
 Data privacy, 316, 619
 DC-DC converter, 515
 Decision-making, 58, 219, 316, 319, 430, 580, 585, 657–671
 Decision tree (DT), 53, 55, 90, 91, 94, 98, 175, 186, 188, 192–193, 214, 217, 218, 248, 250, 251, 309, 317, 330, 350, 399, 409, 420, 422, 424–426, 432, 439, 530, 532–534, 538, 539, 619, 620, 623, 635, 636, 640, 642, 650, 704, 709, 712, 719–721
 Decision tree classifier, 266, 267, 269, 270, 330
 Deep convolutional neural network (DCNN), 28, 58, 121, 605
 Deep convolution generative adversarial networks, 27–43
 Deep learning (DL), 28, 50, 58, 74, 106, 121, 163, 174, 247, 288, 315, 330, 338, 361, 380, 394, 406, 528, 542, 566, 578, 602, 642, 647, 684, 726
 Defense-in-depth, 197
 Demographic filtering, 304, 308, 311
 DenseNet, 360–363, 560, 582
 DFACTS devices, 470, 472, 473
 DH, 565
 DHT11, 497
 Diabetic, 103, 173–182
 Diabetic retinopathy (DR), 103–105, 114, 116, 174–176, 179, 182, 320
 Diagnosis, 15, 49, 58, 59, 74, 82, 84, 114, 116, 181, 241, 243, 256, 316–321, 326, 339, 358, 392, 394, 411, 504, 542, 544, 558, 560, 578, 619, 725–733
 Digital image processing, 50, 109
 Digital library, 88
 Discrete Fractional Fourier Transform (DFrFT), 15
 Disease, 58, 74, 105, 126, 174, 186, 211, 241, 316, 338, 358, 408, 528, 542, 577, 602, 628, 633, 645, 726
 DistilBERT, 164–165, 167, 169–171
 DL algorithm, 173–182, 320, 331, 381, 527–529, 542, 543, 728, 730, 733
 DL techniques, 15, 182, 316–321, 325–336, 338, 340, 380, 389, 406, 532, 543, 647, 728
 Donation, 618–622, 628
 Donors, 617–629
 Drug reactions, 210

E

EASA, 565
 Education system, 704
 Electric grid, 512–514
 Electric vehicle (EV), 511–523
 Electroencephalography (EEG), 410, 696–700
 Electronic health records (EHRs), 316, 319–320, 579, 629
 Embedded system, 446, 613–615
 Encoder-decoder, 543, 549–553, 559
 Encryption, 23, 226, 229–231, 674–677, 681
 Energy model, 154, 159
 Ensemble model, 74, 251, 357–364
 Explainable AI (XAI), 321, 577–586
 Eyeball movement, 446–448, 451

F

Factors, 3, 102, 103, 111, 126–130, 136, 152, 174, 186, 210, 211, 226, 244, 245, 278, 304, 305, 318, 326, 328, 339, 348, 350, 355, 358, 364, 410, 420, 430, 442, 460, 461, 468, 471, 474, 500, 507, 513, 544, 566, 578, 581, 582, 618, 620, 632, 633, 685, 705, 706, 711, 713, 718, 719
 Feature extraction, 74, 75, 80, 112, 116, 177–178, 189–190, 249, 288–290, 328–330, 332, 333, 359, 393, 394, 531, 543, 549, 550, 561, 602, 603, 619, 646, 696, 730
 Feature selection, 91, 247, 252, 303–312, 396–397, 434, 622, 628
 Fertility, 613
 First section, 484
 Flask, 38, 39, 401, 721
 Forecasting, 50, 250, 254, 259–271, 348, 350, 420, 438, 500, 529, 578, 684–686, 704, 705, 717, 718
 F1-score, 58, 65, 68, 70, 79–83, 92, 94, 96, 98, 163, 249, 253, 266, 295, 333, 334, 353, 354, 394, 398, 399, 440, 530, 559, 566, 579, 608, 623, 638–641, 646, 647, 649–651, 653, 704
 Functional murmur, 74, 75, 79, 80, 82
 Fundus, 101–116, 173–182, 578, 580–582, 585

G

GAN-CLS, 27–43
 GATE clearance prediction, 417–426
 GATE CS marks prediction, 418
 GATE score prediction, 418, 421, 425

Generative adversarial networks (GANs), 27–43, 121, 122, 407, 559
 Genes, 211, 241–256, 320, 586
 Geofencing, 490
 Glaucoma, 577–586
 Gradient boost, 719–721
 Gradient matrix, 631
 Graphical representations, 1–11, 138, 711
 Greenhouse, 47–55
 Green IoT, 151–159
 Green networks, 152

H

Haar-cascade method, 289, 293
 Hate speech, 161–171
 Health, 103, 119, 185–195, 211, 220, 242, 260, 262, 271, 339, 355, 408, 411, 413, 528, 529, 532, 542, 558, 579, 611, 618, 619, 624, 629, 632, 634, 706, 707, 726, 733
 Healthcare management, 315–321
 Healthcare providers, 220, 318, 321
 Heat conduction, 590
 Hedging, 348
 Heterogeneous water heaters, 589–599
 High-Frequency Structure Simulator (HFSS), 277, 280
 Hog-face detection, 293
 Human stress, 527–539
 Humidity, 48, 49, 51, 55, 260, 498, 500, 615, 720
 Hybrid routing protocols, 134

I

Image captioning, 379–389
 Image classification, 363, 535, 542, 543, 605
 Image processing, 15, 23, 50, 52, 103, 106–111, 120, 177, 289, 290, 341, 413, 483, 530, 603, 649, 728
 Inception V3, 332, 601–611
 Integration, 110, 316, 320, 321, 383, 400–401, 408, 467–477, 490, 505, 513, 610, 629, 717
 Intelligent Traffic Control, 484
 Internet of Things (IoT), 48, 59, 151–159, 204, 207, 497–500, 503–509, 615
 Interpretable machine learning (IML), 577–586
 Interval-valued intuitionistic fuzzy sets (IVIFSs), 657–671
 Irrigation, 49, 468, 497–500, 616
 ISB, 14
 IWT, 15, 16

K

K-nearest neighbor (KNN), 83, 174, 175, 179, 181, 188, 192, 214, 215, 220, 266, 267, 269, 270, 288, 350, 431, 432, 534, 536–539, 566, 620, 647, 650, 704, 710, 712–714
 KNN classifier, 192, 266, 647

L

LEACH algorithm, 684, 686, 690–693
 Leaf diseases, 50, 52, 59, 339, 340, 645–653, 725–733
 Least significant bit (LSB), 14–17, 19–24
 Li battery, 512
 Lights, 105–107, 179, 358, 470, 483–487, 512, 515, 578, 616, 698
 Linear regression, 4, 248, 265, 266, 348, 418, 420, 422–426, 634, 636, 640
 Long short-term memory (LSTM), 318–320, 332, 348, 382, 384, 385, 387, 389, 393, 394, 551, 552, 559, 566–568, 571–574, 684, 686, 690, 691, 705
 Loss, 31, 34, 36, 37, 41, 42, 78, 105, 121, 126–128, 148, 169, 181, 182, 249, 251, 252, 254, 288, 295, 326, 338–340, 345, 348, 369, 372, 374, 375, 393, 430, 458, 460, 470, 476, 499, 504, 515, 535, 536, 552, 555, 561, 562, 581–583, 608–609, 639, 646, 674, 710, 726–728, 733
 Low-resolution (LR), 120–126

M

Machine learning (ML), 31, 49, 58, 74, 88, 120, 162, 174, 186, 211, 240, 242, 260, 286, 309, 315, 326, 348, 360, 372, 386, 392, 406, 418, 430, 500, 529, 544, 566, 578, 603, 618, 632, 646, 704, 718, 729
 Mango fruit classification, 119–131
 Markovian process, 591, 593
 Medical, 15, 16, 84, 176, 186–188, 194, 210–223, 235, 243, 276, 316–319, 326, 327, 331, 332, 358, 364, 406, 408, 411, 413, 542–562, 578–580, 585, 590, 602, 618–621, 624, 628, 629, 632, 633
 Medical decision support, 629
 Medical diagnosis, 413, 542–562
 Medical documentation, 541
 Medical EMG, 405
 Medical imaging analysis, 316–319, 326, 332, 364
 Medical Trails, 209
 Medicinal plants, 339, 601–611

Metasploit, 199, 204, 205, 207
 Microaneurysms, 102–116, 175, 177
 Microstrip, 274–282
 ML algorithms, 120, 163, 186, 215, 219, 242, 243, 248, 255, 256, 318, 320, 326, 327, 348–355, 392–402, 406, 422, 426, 437, 439, 528–539, 544, 568, 579, 603, 618, 621, 623, 636, 647
 MLSB, 14
 ML techniques, 88–98, 121, 179, 211, 242, 243, 260, 326–336, 340, 394, 406, 420, 438, 439, 530, 532, 534, 535, 603, 618, 632, 633, 635, 646, 649, 705, 718
 Mobile sinks, 683–693
 Model, 2, 28, 49, 58, 74, 88, 103, 121, 138, 153, 162, 174, 186, 204, 212, 226, 245, 260, 277, 286, 305, 316, 326, 338, 348, 358, 369, 381, 392, 406, 418, 431, 447, 458, 482, 505, 517, 530, 542, 566, 578, 590, 602, 618, 632, 646, 675, 688, 704, 718, 726
 Mode seeking generative adversarial network (MSGAN), 28, 31, 34–35, 43
 Modified Ad hoc On-Demand Distance Vector (MAODV), 134–150
 Morphological image processing, 103, 107–111
 Mortality, 74, 246, 319, 631–642
 Mouse, 7, 8, 445–454
 MSC, 13
 Multilayer feed-forward neural network model, 287
 Multiple linear regression, 718–721

N

Natural language processing (NLP), 3, 162, 163, 165, 171, 179, 248, 251, 355, 369, 370, 381, 406, 413, 442, 535, 542, 543, 553, 559
 Netdiscover, 205
 Network transmission, 504, 507
 Neural network, 28, 30, 58, 111, 124, 247–249, 288, 289, 292, 330, 332, 361, 373, 385, 387, 409, 412, 413, 420, 438, 530, 619, 621, 623, 634, 636
 Neural Signal Processing, 695
 Next item prediction, 369, 373
 Nmap, 199, 202–205
 NODEMCU, 503, 613, 615, 616
 Node MCU, 615, 616
 Non-sleep mode, 528–539
 Non-steroidal anti-inflammatory drugs (NSAID), 210–223

NS3, 151
 NS2 simulator, 134, 144
 Numerical characterization, 5–6

O

Object detection (OD), 58, 59, 61, 62, 65, 107, 338, 412
 Open CV, 445
 Optimized distributed gradient boosting library (XGBoost), 243, 245, 246, 251–252, 255, 256, 350, 351, 354, 433, 439–441, 532–534, 719–722

P

Pandemic, 631, 632, 641, 642
 Parasitic element, 274, 275
 Patient, 15, 16, 76, 84, 173–182, 186, 187, 195, 211, 212, 219–223, 226, 235, 243–245, 256, 315–321, 392, 408, 413, 447, 542, 544–548, 557, 558, 560, 562, 578, 582, 584, 585, 619, 621, 624–626, 628, 629, 632–642, 696, 698
 Patient outcome prediction, 319, 560
 Pattern recognition, 50, 406–413
 Peak signal-to-noise ratio (PSNR), 15–17, 126, 130
 People analytics, 430
 Personalized medicine, 242, 317, 319–321, 560
 Phishing attack, 489
 Phonocardiogram, 75
 Phylogenetic tree, 6, 7, 9, 10
 Physicians, 585, 640, 642
 PIC, 505
 Pilot super resolution network (PSRN), 119–131
 Pixel density, 581
 Plant Biology, 405
 Plant disease, 50, 58–60, 65, 338–341, 345, 408, 727–730, 733
 Plant diseases detection, 57–70, 338–345, 727
 PlantDoc, 58, 59, 61, 64, 65, 341, 344
 Pollution, 152, 259–271, 457, 459, 512
 Position prediction, 683–693
 Power quality, 457–465, 467–477, 513
 Precision, 16, 58, 59, 62, 64, 65, 67–70, 78–83, 90, 92, 94, 96, 98, 186, 194, 216, 243, 244, 246, 249, 252, 261, 266, 295, 333–335, 353, 380, 398, 399, 407, 431–433, 440, 448, 499–501, 507, 530, 607, 623, 638–642, 646–649, 653, 696, 704

Prediction, 43, 62, 64, 66, 79, 80, 112, 124, 165, 174, 185–195, 210, 211, 215, 217, 219, 245, 247–252, 255, 265, 266, 268, 289, 290, 295, 304, 305, 308, 317–320, 326–336, 340, 348, 350, 351, 360, 362, 363, 370, 371, 373, 385, 406, 409, 418–426, 430–432, 437–439, 442, 499, 528, 534, 542–562, 566, 568, 578, 585, 618, 622, 634–636, 642, 646, 648, 684–688, 690–693, 704, 705, 709–712, 715, 717–722
 Predictive analytics, 248, 432, 618–629
 Pressure, 102, 103, 528, 529, 578
 Protein sequences, 1–11
 PV array, 511–523

Q

QAR, 566
 QR code, 490, 492–495
 Queuing Probabilities, 589

R

Radiation pattern, 275, 280, 281
 Random forest, 53, 55, 94, 96, 98, 185–195, 269, 288, 305, 309–311, 317, 331, 340, 348, 350, 354, 355, 393, 397, 399, 409, 412, 431, 432, 498, 529, 530, 536–538, 619, 621, 623, 640, 642, 650, 704, 709, 718
 Random forest classifier, 260, 266, 267, 269, 270, 288, 331, 532, 533, 538, 539, 709
 Ranking, 89, 91, 243, 244, 375, 629, 661, 662, 667
 Real time, 28, 50, 58, 59, 62, 107, 144, 149, 165, 286, 287, 290, 296, 338, 340, 368, 406, 408, 442, 446, 449, 483, 490, 498, 504, 505, 507, 523, 603, 611, 615, 629, 698, 722, 726–733
 Real-Time Control Systems, 695
 Real time image processing, 58
 Recall, 58, 59, 65, 67, 69, 78, 80, 81, 90, 92, 94, 96, 98, 243, 246, 249, 252, 266, 295, 333–335, 353, 374, 375, 398, 399, 440, 530, 607, 623, 638–641, 646–649, 653, 704
 Recipient, 14, 619, 624–626
 Recommendation system, 89, 248, 304–312, 368–376, 413
 Recurrent neural networks (RNNs), 28, 32, 318–320, 332–336, 369, 381, 382, 385, 407, 410, 530, 543, 551, 559, 566, 605, 610, 619, 622, 633, 647, 684–686

Region of interest, 542
 Renewable energy, 467–477, 512, 513, 717
 Resilient backpropagation (Rprop), 103,
 111–112, 116
 ResNet, 37, 63, 65, 329, 331, 340, 360, 362,
 364, 381, 393
 Reviewer assignment problem (RAP), 88–93,
 95
 R-FTG, 566
 Risk, 48, 104, 114, 163, 182, 186, 202, 204,
 210–212, 219, 220, 222, 316, 318, 319,
 321, 338, 340, 348, 354, 413, 420, 499,
 504, 528, 560, 566, 578, 581, 582, 585,
 618–629, 631–642, 668, 674, 705, 713,
 715
 Root Tkinter, 185
 Routing, 134–150, 153–156, 159, 684–686,
 688, 692, 693
 Routing algorithms, 684, 685, 688, 692

S
 Search, 174, 204, 227, 229–233, 235, 236, 238,
 239, 316, 380, 381, 552, 556, 566, 687
 Searchable encryption, 225–240
 Secure database framework, 674–681
 Security, 13–25, 58, 140, 142, 149, 197–204,
 207, 226, 230, 286, 321, 338, 339, 413,
 489–491, 493, 495, 618, 674, 675, 679,
 693, 727, 728
 Sensors, 48, 49, 52, 105, 152, 409, 413, 447,
 498–500, 506–508, 529, 530, 614–616,
 684, 685, 688, 690, 693, 697
 SEPIC converter, 515–518, 520, 521
 Sequential recommendation, 369
 Session-based recommendation, 368–370, 373,
 376
 Shunt hybrid power filter (SHPF), 458, 459,
 461–463
 Similarity comparison, 674–681
 Skin disease, 358–364
 Sleep mode, 528–539
 Smart Health, 185–195
 Soft computing, 469, 471, 475–477
 Softmax, 78, 249, 251, 288, 326, 362, 372,
 373, 580
 Soil moisture, 498–500
 Solar energy, 512, 514, 717–722
 S-Parameter, 277, 280
 Speech disorder, 392
 Spike mind association (SNN), 581, 582
 SQL injection attack, 674–676, 681
 Steganography, 13–25

Stress detection, 528–539
 Stress levels, 529–533, 535, 536, 539
 Student's performance, 422, 704–715
 Stuttering, 392–402
 Sub-modular pick local interpretable model-
 agonistic explanation (SP-LIME), 580,
 581, 584, 585
 Super resolution, 120, 124, 126, 129, 131
 Support vector machines (SVMs), 50, 53, 55,
 74, 75, 83, 88, 90, 93, 95, 169, 170,
 174, 175, 179, 181, 248, 288, 317–320,
 327, 330, 333, 335, 336, 339, 350, 354,
 393, 408, 410, 412, 431, 432, 476, 529,
 530, 532, 535–539, 566, 567, 570, 571,
 574, 602, 603, 619, 621, 623, 646, 647,
 649, 705, 709, 712

T

Temperature, 48, 49, 51, 55, 260, 475,
 498–500, 507, 531, 533, 539, 590, 615,
 718, 720
 TensorFlow, 58, 78, 294, 361
 Text analytics, 88
 Thing Speak, 497
 Tomato, 61, 646, 726–733
 Total harmonic distortion (THD), 462–465
 Tough touchdown (TD), 566
 Toxic comment, 162–171
 Traffic management system, 482–484
 Transfer learning, 84, 165, 358–364, 383, 550,
 560, 603
 Transformer, 164, 165, 169, 171, 176, 369,
 370, 372–374, 458, 504–506, 508, 515,
 580
 Type-2 intuitionistic fuzzy sets (T2IFSs),
 657–671

U

User interface (UI), 40, 190–191, 286, 289,
 292–294, 296, 298, 394, 400–401, 532,
 535–536, 675, 719
 User interface integration, 400–401

V

Ventricular septal defect (VSD), 75, 76, 79, 80,
 84
 Vision, 105, 179, 412, 440, 485, 577
 Vulnerability assessments and penetration
 testing (VAPT), 197–207

W

Web service, 48

Wi-Fi, 155, 507, 615

Wireless sensor networks (WSNs), 49, 153,
500, 683–693

Word2Vec, 28, 30, 42, 164, 171, 369

X

XLNet, 369–373, 376

X-ray images, 327, 333, 334, 543–545, 547,
549, 553, 558, 560, 561

Y

Yields, 48, 58, 59, 110, 169, 181, 246, 339,
345, 384, 399, 499, 648, 726–728, 730,
731, 733

YOLOv7, 57–70, 337–345

YOLOv8, 57–70