

Devendra Kumar Sharma
Sheng-Lung Peng
Rohit Sharma
Gwanggil Jeon *Editors*

Micro-Electronics and Telecommunication Engineering

Proceedings of 7th ICMETE 2023

Lecture Notes in Networks and Systems

Volume 894

Series Editor

Janusz Kacprzyk , Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye

Derong Liu, Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA

Institute of Automation, Chinese Academy of Sciences, Beijing, China

Witold Pedrycz, Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Marios M. Polycarpou, Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary

Jun Wang, Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Devendra Kumar Sharma · Sheng-Lung Peng ·
Rohit Sharma · Gwanggil Jeon
Editors

Micro-Electronics and Telecommunication Engineering

Proceedings of 7th ICMETE 2023



Springer

Editors

Devendra Kumar Sharma
Department of Electronics
and Communication Engineering
ABES Engineering College
Ghaziabad, Uttar Pradesh, India

Rohit Sharma
Department of Electronics
and Communication Engineering
ABES Engineering College
Ghaziabad, Uttar Pradesh, India

Sheng-Lung Peng
National Taipei University of Business
Taoyuan, Taiwan

Gwanggil Jeon
Department of Embedded Systems
Engineering
Incheon National University
Incheon, Korea (Republic of)

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-99-9561-5

ISBN 978-981-99-9562-2 (eBook)

<https://doi.org/10.1007/978-981-99-9562-2>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024, corrected publication 2024

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 Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721,
Singapore

Paper in this product is recyclable.

Preface

The book presents high-quality papers from the Fourth International Conference on Microelectronics and Telecommunication Engineering (ICMTE 2023). It discusses the latest technological trends and advances in major research areas such as microelectronics, wireless communications, optical communication, signal processing, image processing, big data, cloud computing, artificial intelligence, and sensor network applications. This book includes the contributions of national/international scientists, researchers, and engineers from both academia and the industry. The contents of this volume will be useful to researchers, professionals, and students alike.

Ghaziabad, India

Taoyuan, Taiwan

Ghaziabad, India

Incheon, Korea (Republic of)

Devendra Kumar Sharma

Sheng-Lung Peng

Rohit Sharma

Gwanggil Jeon

Contents

Transportation in IoT-SDN Using Vertical Handoff Scheme	1
Jyoti Maini and Shalli Rani	
MLP-Based Speech Emotion Recognition for Audio and Visual Features	13
G. Kothai, Prabhas Bhanu Boora, S. Muzammil, L. Venkata Subhash, and B. Naga Raju	
Drain Current and Transconductance Analysis of Double-Gate Vertical Doped Layer TFET	29
Mandeep Singh, Nakkina Sai Teja, Tarun Chaudhary, Balwinder Raj, and Deepti Kakkar	
OpenFace Tracker and GoogleNet: To Track and Detect Emotional States for People with Asperger Syndrome	43
Mays Ali Shaker and Amina Atiya Dawood	
Vehicle Classification and License Number Plate Detection Using Deep Learning	57
Kaushal Kishor, Ankit Shukla, and Anubhav Thakur	
Car Price Prediction Model Using ML	67
Kaushal Kishor, Akash Kumar, and Kabir Choudhary	
Effects of Material Deformation on U-shaped Optical Fiber Sensor	75
Mohd Ashraf, Mainuddin, Mirza Tariq Beg, Ananta Sekia, and Sanjai K. Dwivedi	
Classification of DNA Sequence for Diabetes Mellitus Type Using Machine Learning Methods	87
Lena Abed AL Raheem Hamza, Hussein Attia Lafta, and Sura Zaki Al Rashid	

Unveiling the Future: A Review of Financial Fraud Detection Using Artificial Intelligence Techniques	103
Sankalp Goel and Abha Kiran Rajpoot	
Remodeling E-Commerce Through Decentralization: A Study of Trust, Security and Efficiency	113
Adnan Shakeel Ahmed, Danish Raza Rizvi, Dinesh Prasad, and Amber Khan	
Estimation of Wildfire Conditions via Perimeter and Surface Area Optimization Using Convolutional Neural Network	125
R. Mythili, K. Abinav, Sourav Kumar Singh, and S. Suresh Krishna	
A Framework Provides Authorized Personnel with Secure Access to Their Electronic Health Records	137
Kanneboina Ashok and S. Gopikrishnan	
Explainable Artificial Intelligence for Deep Learning Models in Diagnosing Brain Tumor Disorder	149
Kamini Lamba and Shalli Rani	
Pioneering a New Era of Global Transactions: Decentralized Overseas Transactions on the Blockchain	161
Khadeer Dudekula and Annapurani Panaiyappan K.	
A Perspective Review of Generative Adversarial Network in Medical Image Denoising	173
S. P. Porkodi and V. Sarada	
Osteoporosis Detection Based on X-Ray Using Deep Convolutional Neural Network	183
Abulkareem Z. Mohammed and Loay E. George	
Fault Prediction and Diagnosis of Bearing Assembly	197
Chirag Agarwal, Aman Agarwal, Anmol Tyagi, Dev Tyagi, Mohini Preetam Singh, and Rahul Singh	
Bearing Fault Diagnosis Using Machine Learning Models	219
Shagun Chandrvanshi, Shivam Sharma, Mohini Preetam Singh, and Rahul Singh	
A High-Payload Image Steganography Based on Shamir's Secret Sharing Scheme	235
Sanjive Tyagi, Maysara Mazin Alsaad, and Sharvan Kumar Garg	
Design and Comparison of Various Parameters of T-Shaped TFET of Variable Gate Lengths and Materials	249
Jyoti Upadhyay, Tarun Chaudhary, Ramesh Kumar Sunkaria, and Mandeep Singh	

Experiment to Find Out Suitable Machine Learning Algorithm for Enzyme Subclass Classification	263
Amitav Saran, Partha Sarathi Ghosh, Umasankar Das, and Thiagarajan Chenga Kalvinathan	
Iris Recognition Method for Non-cooperative Images	275
Zainab Ghayyib Abdul Hasan	
An Exploration: Deep Learning-Based Hybrid Model for Automated Diagnosis and Classification of Brain Tumor Disorder	289
Kamini Lamba and Shalli Rani	
Recognition of Apple Leaves Infection Using DenseNet121 with Additional Layers	297
Shubham Nain, Neha Mittal, and Ayushi Jain	
Techniques for Digital Image Watermarking: A Review	309
Bipasha Shukla, Kalpana Singh, and Kavita Chaudhary	
Improved Traffic Sign Recognition System for Driver Safety Using Dimensionality Reduction Techniques	319
Manisha Vashisht and Vipul Vashisht	
Detection of Fake Reviews in Yelp Dataset Using Machine Learning and Chain Classifier Approach	331
Lina Shugaa Abdulzahra and Ahmed J. Obaid	
Data Governance Framework for Industrial Internet of Things	347
Mohammed Alaa Al-Hamami and Ahmed Alaa Al-Hamami	
IOT-Based Water Level Management System	357
N. C. A. Boovarahan, S. Lakshmi, K. Umapathy, T. Dinesh Kumar, M. A. Archana, K. Saraswathi, S. Omkumar, and Ahmed Hussein Alkhayyat	
A Review on Privacy Preservation in Cloud Computing and Recent Trends	365
Srutipragyan Swain, Prasant Kumar Pattnaik, and Banchhanidhi Dash	
EEECT-IOT-HWSN: The Energy Efficient-Based Enhanced Clustering Technique Using IOT-Based Heterogeneous Wireless Sensor Networks	377
Mustafa Dh. Hassib, Mohammed Joudah Zaiter, and Wasan Hashim Al Masoody	
IoT-Based Smart System for Fire Detection in Forests	389
M. A. Archana, T. Dinesh Kumar, K. Umapathy, S. Omkumar, S. Prabakaran, N. C. A. Boovarahan, C. Parthasarathy, and Ahmed Hussein Alkhayyat	

Machine Learning Approach to Lung Cancer Survivability Analysis	397
Srichandana Abbineni, K. Eswara Rao, Rella Usha Rani, P. Ila Chandana Kumari, and S. Swarajya Lakshmi	
Application of Analytical Network Processing (ANP) Method in Ranking Cybersecurity Metrics	409
Seema Gupta Bhol, Jnyana Ranjan Mohanty, and Prasant Kumar Patnaik	
Advanced Real-Time Monitoring System for Marine Net Pens: Integrating Sensors, GPRS, GPS, and IoT with Embedded Systems	419
Sayantan Panda, R. Narayananamoorthi, and Samiappan Dhanalakshmi	
Harnessing Machine Learning to Optimize Customer Relations: A Data-Driven Approach	437
Santosh Kumar, Priti Verma, Dhaarna Singh Rathore, Richa Pandey, and Gunjan Chhabra	
Immersive Learning Using Metaverse: Transforming the Education Industry Through Extended Reality	447
Gayathri Karthick, B. Rebecca Jeyavadhanams, Soonleh Ling, Anum Kiyani, and Nalinda Somasiri	
Internet of Things Heart Disease Detection with Machine Learning and EfficientNet-B0	457
D. Akila, M. Thyagaraj, D. Senthil, Saurav Adhikari, and K. Kavitha	
Deep Learning in Distance Awareness Using Deep Learning Method	469
Raghad I. Hussein and Ameer N. Onaizah	
Analysis of Improving Sales Process Efficiency with Salesforce Industries CPQ in CRM	481
Pritesh Pathak, Souvik Pal, Saikat Maity, S. Jeyalaksshmi, Saurabh Adhikari, and D. Akila	
Analyze and Compare the Public Cloud Provider Pricing Model and the Impact on Corporate Financial	497
Jaideep Singh, Souvik Pal, Bikramjit Sarkar, H. Selvi, Saurabh Adhikari, K. Madhumathi, and D. Akila	
A Data-Driven Analytical Approach on Digital Adoption and Digital Policy for Pharmaceutical Industry in India	509
Anup Rana, Bikramjit Sarkar, Raj Kumar Parida, Saurabh Adhikari, R. Anandha Lakshmi, D. Akila, and Souvik Pal	
Framework for Reverse Supply Chain Using Sustainable Return Policy	523
Tridha Bajaj, Snigdha Parashar, Tanupriya Choudhury, and Ketan Kotecha	

Sentiment Analysis Survey Using Deep Learning Techniques	539
Neha Singh, Umesh Chandra Jaiswal, and Jyoti Srivastava	
Identifying Multiple Diseases on a Single Citrus Leaf Using Deep Learning Techniques	549
Ayushi Gupta, Anuradha Chug, and Amit Prakash Singh	
IoT-Based Health Monitoring System for Heartbeat—Analysis	561
B. Mary Havilah Haque and K. Martin Sagayam	
A Study and Comparison of Cryptographic Mechanisms on Data Communication in Internet of Things (IoT) Network and Devices	571
Abhinav Vidwans and Manoj Ramaiya	
Fake News Detection Using Data Science Approaches	585
Lina Shugaa Abdulzahra and Ahmed J. Obaid	
Reversible Data-Hiding Scheme Using Color Coding for Ownership Authentication	593
Anuj Kumar Singh, Sandeep Kumar, and Vineet Kumar Singh	
Comprehensive Approach for Image Noise Analysis: Detection, Classification, Estimation, and Denoising	601
Rusul A. Al Mudhafar and Nidhal K. El Abbadi	
Optimal Path Selection Algorithm for Energy and Lifetime Maximization in Mobile Ad Hoc Networks Using Deep Learning	617
Jyoti Srivastava and Jay Prakash	
Automated Air Pollution Monitoring System	631
G. Poornima, S. Lakshmi, D. Muthukumaran, T. Dinesh Kumar, K. Umapathy, N. C. A. Boovarahan, M. A. Archana, and Ahmed Hussein Alkhayyat	
Simulation and Implementation of Solar Charge Controller by MPPT Algorithm	641
D. Vanitha, V. Malathi, and K. Umapathy	
Nanoscale Multi-gate Graded Channel DG-MOSFET for Reduced Short Channel Effects	653
Ashutosh Pandey, Kousik Midya, Divya Sharma, and Seema Garg	
Performance Enhancement and Scheduling in Communication Networks—A Review into Various Approaches	661
Priya Kumari and Nitin Jain	
A Survey of Network Protocols for Performance Enhancement in Wireless Sensor Networks	673
Abhishek Gupta, Devendra Kumar Sharma, and D. N. Sahai	

Airbnb Price Prediction Using Advanced Regression Techniques and Deployment Using Streamlit	685
Ayan Sar, Tanupriya Choudhury, Tridha Bajaj, Ketan Kotecha, and Mirtha Silvana Garat de Marin	
Tiger Community Analysis in the Sundarbans	699
Richa Choudhary, Tanupriya Choudhury, and Susheela Dahiya	
An Overview of the Use of Deep Learning Algorithms to Predict Bankruptcy	715
Kamred Udhamp Singh, Ankit Kumar, Gaurav Kumar, Teekam Singh, Tanupriya Choudhury, and Ketan Kotecha	
An Analytical Study of Improved Machine Learning Approaches for Predicting Mode of Delivery	727
Vaishali Bhargava and Sharvan Kumar Garg	
Comparative Study of Different Document Similarity Measures and Models	737
Anshika Singh and Sharvan Kumar Garg	
Schmitt Trigger Leakage Reduction Using MT莫斯 Technique at 45-Nm Technology	747
Deepak Garg and Devendra Kumar Sharma	
A Review of Survey and Assessment of Facial Emotion Recognition (FER) by Convolutional Neural Networks	755
Sanyam Agarwal, Veer Daksh Agarwal, Ishaan Agarwal, Vipin Mittal, Lakshay Singla, and Ahmed Hussein Alkhayyat	
A New Compact-Data Encryption Standard (NC-DES) Algorithm Security and Privacy in Smart City	769
Abdullah J. Alzahrani	
Design, Development, and Mathematical Modelling of Hexacopter	785
Vishwas Mishra, Priyank Sharma, Abhishek Kumar, and Shyam Akashe	
A Keypoint-Based Technique for Detecting the Copy Move Forgery in Digital Images	797
Kaleemur Rehman and Saiful Islam	
Correction to: Machine Learning Approach to Lung Cancer Survivability Analysis	C1
Srichandana Abbineni, K. Eswara Rao, Rella Usha Rani, P. Ila Chandana Kumari, and S. Swarajya Lakshmi	
Author Index	813

Editors and Contributors

About the Editors

Devendra Kumar Sharma received his B.E. degree in Electronics Engineering from Motilal Nehru National Institute of Technology, Allahabad, in 1989, his M.E. degree from Indian Institute of Technology Roorkee, Roorkee, in 1992, and his Ph.D. degree from National Institute of Technology, Kurukshetra, India, in 2016. He is Professor and Dean of SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India. He has authored many papers in several international journals and conferences of repute. His research interests include VLSI interconnects, electronic circuits, digital design, testing, and signal processing. He is Life Member of ISTE, Fellow of IETE, and Senior Member of IEEE.

Sheng-Lung Peng is a professor of the Department of Creative Technologies and Product Design in National Taipei University of Business Taiwan, an honorary professor in Beijing Information Science and Technology University, and a visiting professor in Ningxia Institute of Science and Technology, China. He received a B.S. degree in Mathematics from National Tsing Hua University, and the M.S. and Ph.D. degrees in Computer Science from the National Chung Cheng University and National Tsing Hua University, Taiwan, respectively. He is an honorary professor of Beijing Information Science and Technology University, China, and a visiting professor of the Ningxia Institute of Science and Technology, China. He serves as the secretary-general of the ACM-ICPC Contest Council for Taiwan and the regional director of the ICPC Asia Taipei-Hsinchu site. He is a director of the Institute of Information and Computing Machinery, of Information Service Association of Chinese Colleges and Taiwan Association of Cloud Computing.

Rohit Sharma is currently Associate Professor and Head of the Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India. He is Editorial Board Member and Reviewer of more than 12 international journals and conferences. He serves as Book Editor for

16 different titles His research interests are data networks, data security, data mining, environment and pollution trend analysis, Big Data, IoT, etc. He is Member of ISTE, ICS, IAENG, IACSIT and Senior Member of IEEE.

Prof. Gwanggil Jeon received the B.S., M.S., and Ph.D. (summa cum laude) degrees from the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea, in 2003, 2005, and 2008, respectively. He is currently Full Professor at Incheon National University, Incheon, Korea. He has published more than 65 papers in journals and 25 conference proceedings. He served as Member of the Technical Program Committee at many international conferences in Poland, India, China, Iran, Romania, and Bulgaria. His area of interest includes IT in Business, IoT in Business, and Education Technology.

Contributors

Nidhal K. El Abbadi Computer Techniques Engineering Department, Al-Mustaqlal University, Babylon, Iraq

Srichandana Abbineni Department of CSE (DS), CVR College of Engineering, Hyderabad, India

Lina Shugaa Abdulzahra Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

K. Abinav Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India

Saurabh Adhikari School of Engineering, Swami Vivekananda University, Kolkata, India

Saurav Adhikari School of Engineering, Swami Vivekananda University, Kolkata, India

Aman Agarwal Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Chirag Agarwal Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Ishaan Agarwal Department of Computer Science Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

Sanyam Agarwal Department of Electronics and Communication Engineering, ACE College of Engineering and Management, Agray, India

Veer Daksh Agarwal Department of Computer Science Engineering, Thapar Institute of Engineering and Technology, Patialay, India

Adnan Shakeel Ahmed Department of Electronics and Communications Engineering, Faculty of Engineering and Technology, New Delhi, India

Shyam Akashe Department of ECE, ITM University, Gwalior, India

D. Akila Department of Computer Applications, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed University, Chennai, India

Wasan Hashim Al Masoody Electrical Engineering Department, College of Engineering, Babylon, Iraq

Rusul A. Al Mudhafar Computer Science Department, Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq

Lena Abed AL Raheim Hamza College of Science for Women, University of Babylon, Babylon, Iraq

Sura Zaki Al Rashid College of Information Technology, University of Babylon, Babylon, Iraq

Ahmed Alaa Al-Hamami Bedfordshire University, Luton, UK

Mohammed Alaa Al-Hamami Applied Science University, Eker, Kingdom of Bahrain

Ahmed Hussein Alkhayyat Scientific Research Centre of the Islamic University, The Islamic University, Najaf, Iraq

Maysara Mazin Alsaad Department of Computer Science, Gujarat University, Ahmedabad, India

Abdullah J. Alzahrani Computer Engineering Department, College of Computer Science and Engineering, University of Ha'il, Hail, Saudi Arabia

R. Anandha Lakshmi Department of Computer Application, Anna Adarsh College for Women, Chennai, India

M. A. Archana SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

Kanneboina Ashok School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Mohd Ashraf Department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi, India

Tridha Bajaj Informatics Cluster, School of Computer Science (SoCS), University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

Mirza Tariq Beg Department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi, India

Vaishali Bhargava Swami Vivekanand Subharti University, Meerut, UP, India

Seema Gupta Bhol KIIT University, Bhubaneshwar, India

Prabhas Bhanu Boora Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar, India

N. C. A. Boovarahan SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

Shagun Chandrvanshi Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Kavita Chaudhary Deptartment of ECED, MIET, Meerut, India

Tarun Chaudhary Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Gunjan Chhabra Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun, Uttarakhand, India

Kabir Choudhary ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Richa Choudhary School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

Tanupriya Choudhury CSE Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, India;
SoCS, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India;
CSE Department, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

Anuradha Chug University School of Information, Communication and Technology, GGSIPU, New Delhi, India

Susheela Dahiya School of Computer Science, Graphic Era Hill University, Dehradun, Uttarakhand, India

Umasankar Das LS Discovery, Cognizant, Chennai, India

Banchhanidhi Dash School of Computer Engineering, KIIT Deemed to Be University, Bhubaneswar, India

Amina Atiya Dawood Computer Science Department, College of Science for Women, University of Babylon, Babylon, Iraq

Mirtha Silvana Garat de Marin Engineering Research and Innovation Group, Universidad Europea del Atlántico, Santander, Spain;
Department of Project Management, Universidade Internacional do Cuanza, Cuito-Bié, Angola

Samiappan Dhanalakshmi Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203 India

T. Dinesh Kumar SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

Khadeer Dudekula Department of Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, India

Sanjai K. Dwivedi DOP, DRDO, Delhi, India

Deepak Garg Department of Electronics and Communication Engineering, SRMIST, Delhi-NCR Campus, Ghaziabad, India

Seema Garg Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh, India

Sharvan Kumar Garg Department of Computer Science and Engineering, Subharti Institute of Technology and Engineering, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India

Loay E. George Informatics Institute for Postgraduate Studies, Baghdad, Iraq

Partha Sarathi Ghosh LS Discovery, Cognizant, Chennai, India

Sankalp Goel Department of Computer Science Engineering, Sharda University, Greater Noida, India

S. Gopikrishnan School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Abhishek Gupta Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Ghaziabad, India

Ayushi Gupta University School of Information, Communication and Technology, GGSIPU, New Delhi, India

Zainab Ghayyib Abdul Hasan Department of Computer Science, Faculty of Education for Girls, University of Kufa, Kufa, Iraq

Mustafa Dh. Hassib Department of Communications Engineering, University of Technology, Baghdad, Iraq

Raghad I. Hussein Faculty of Pharmacy, University of Kufa, Najaf, Iraq

Saiful Islam ZHCET, AMU, Aligarh, UP, India

Ayushi Jain Excel Geomatics Pvt Ltd, Noida, India

Nitin Jain Babu Baranasi Das University, UttarPradesh, Lucknow, India

Umesh Chandra Jaiswal Madan Mohan Malviya University of Technology, Gorakhpur, India

S. Jeyalaksshmi Department of Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

Annapurani Panaiyappan K. Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur, India

Deepti Kakkar Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Thiyagarajan Chenga Kalvinathan LS Discovery, Cognizant, Chennai, India

Gayathri Karthick York St John University, London, UK

K. Kavitha Guru Nanak College, Chennai, India

Amber Khan Department of Electronics and Communications Engineering, Faculty of Engineering and Technology, New Delhi, India

Kaushal Kishor ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Anum Kiyani York St John University, London, UK

Ketan Kotecha Symbiosis Centre for Applied Artificial Intelligence, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, India

G. Kothai Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), KPR Institute of Engineering and Technology, Uthupalayam, Coimbatore, Tamil Nadu, India

S. Suresh Krishna Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India

Abhishek Kumar Department of ECE, NIT Jamshedpur, Jharkhand, India

Akash Kumar ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Ankit Kumar Department of Information Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, Chhattisgarh, India

Gaurav Kumar Department of Computer Engineering & Application's, GLA University, Mathura, India

Sandeep Kumar Department of CSE (AI), ABES Institute of Technology Ghaziabad, Ghaziabad, India

Santosh Kumar Department of Computer Science, ERA University, Lucknow, Uttar Pradesh, India

P. Ila Chandana Kumari CSE Department, Hyderabad Institute of Technology and Management, Hyderabad, India

Priya Kumari Babu Baranasi Das University, UttarPradesh, Lucknow, India

Hussein Attia Lafta College of Information Technology, University of Babylon, Babylon, Iraq

S. Lakshmi Thirumalai Engineering College, Kanchipuram, Tamil Nadu, India

S. Swarajya Lakshmi Department of CSE, KMIT, Hyderabad, India

Kamini Lamba Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Soonleh Ling York St John University, London, UK

K. Madhumathi Department of Computer Application, Anna Adarsh College for Women, Chennai, India

Jyoti Maini Chitkara Institute of Engineering and Technology, Chitkara University, Punjab, India

Mainuddin Department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi, India

Saikat Maity Department of Computer Science and Engineering, Sister Nivedita University (Techno India Group), Kolkata, India

V. Malathi Department of EEE, Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Kanchipuram, Tamil Nadu, India

K. Martin Sagayam Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

B. Mary Havilah Haque Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

Kousik Midya Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh, India

Vishwas Mishra Department of EEE, Swami Vivekananda Subharti University, Meerut, India

Neha Mittal Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, India

Vipin Mittal Department of Electronics and Communication Engineering, IIMT University, Meeut, India

Abulkareem Z. Mohammed Informatics Institute for Postgraduate Studies, Baghdad, Iraq

Jnyana Ranjan Mohanty KIIT University, Bhubaneshwar, India

D. Muthukumaran Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

S. Muzammil Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar, India

R. Mythili Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India

B. Naga Raju Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar, India

Shubham Nain Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, India

R. Narayananamoorthi Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

Ahmed J. Obaid Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

S. Omkumar SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

Ameer N. Onaizah School of Automation, Beijing Institute of Technology, Beijing, China;
Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

Souvik Pal Department of Management Information Systems, Saveetha College of Liberal Arts And Sciences, Saveetha Institute of Medical And Technical Sciences, Chennai, India;
Department of Computer Science and Engineering, Sister Nivedita University (Techno India Group), Kolkata, India

Sayantan Panda Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

Ashutosh Pandey Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh, India

Richa Pandey Sharda University, Greater Noida, Uttar Pradesh, India

Snigdha Parashar School of Computer Science (SoCS), University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

Raj Kumar Parida School of Computing Science and Engineering, Galgotias University, Greater Noida, India

C. Parthasarathy Vel Tech Multi Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, Chennai, Tamil Nadu, India

Pritesh Pathak Liverpool Business School, Liverpool John Moores University, Liverpool, UK

Prasant Kumar Patnaik KIIT University, Bhubaneshwar, India

Prasant Kumar Pattnaik School of Computer Engineering, KIIT Deemed to Be University, Bhubaneswar, India

G. Poornima SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

S. P. Porkodi Department of Electronics and Communication Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Tamil Nadu, India

S. Prabakaran P.T. Lee Chengalvaraya Naicker College of Engineering and Technology, Kanchipuram, India

Jay Prakash Madan Mohan Malviya University of Technology, Gorakhpur, India

Dinesh Prasad Department of Electronics and Communications Engineering, Faculty of Engineering and Technology, New Delhi, India

Balwinder Raj Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Abha Kiran Rajpoot Department of Computer Science Engineering, Sharda University, Greater Noida, India

Manoj Ramaiya Institute of Advanced Computing, Sage University, Indore, MP, India

Anup Rana Liverpool Business School, Liverpool John Moores University, Liverpool, UK

Rella Usha Rani Department of CSE (AI&ML), CVR College of Engineering, Hyderabad, India

Shalli Rani Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

K. Eswara Rao Department of CSE, Aditya Institute of Technology and Management, Tekkali, India

Dhaarna Singh Rathore School of Business, Auro University, Surat, Gujarat, India

B. Rebecca Jeyavadhanams York St John University, London, UK

Kaleemur Rehman ZHCET, AMU, Aligarh, UP, India

Danish Raza Rizvi Department of Computer Engineering, Faculty of Engineering and Technology, New Delhi, India

D. N. Sahai INMARSAT, ALTTC, Ghaziabad, India

Ayan Sar Informatics Cluster, School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

V. Sarada Department of Electronics and Communication Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Tamil Nadu, India

Amitav Saran LS Discovery, Cognizant, Chennai, India

K. Saraswathi SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India

Bikramjit Sarkar Department of Computer Science and Engineering, JIS College of Engineering, Kalyani, India

Ananta Sekia DRL, DRDO, Tezpur, Assam, India

H. Selvi Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed University, Chennai, India

D. Senthil Tagore College of Arts and Science, Chennai, India

Mays Ali Shaker Computer Science Department, College of Science for Women, University of Babylon, Babylon, Iraq

Devendra Kumar Sharma Department of Electronics and Communication Engineering, ABES Engineering College, Ghaziabad, Uttar Pradesh, India

Divya Sharma Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh, India

Priyank Sharma Department of ECE, MIET Meerut, Meerut, India

Shivam Sharma Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Ankit Shukla ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

Bipasha Shukla Deptartment of ECED, MIET, Meerut, India

Amit Prakash Singh University School of Information, Communication and Technology, GGSIPU, New Delhi, India

Anshika Singh Department of Computer Science and Engineering, Subharti Institute of Technology and Engineering, Swami Vivekanand Subharti University, Meerut, India

Anuj Kumar Singh School of Computing Science and Engineering, Galgotias University, Greater Noida, India

Jaideep Singh Liverpool Business School, Liverpool John Moores University, Liverpool, UK

Kalpana Singh Deptartment of UCRD, Chandigarh University, Punjab, India

Kamred Udham Singh School of Computing, Graphic Era Hill University, Dehradun, India

Mandeep Singh Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Mohini Preetam Singh Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Neha Singh Madan Mohan Malviya University of Technology, Gorakhpur, India

Rahul Singh Product Technology and Development Centre, Larsen & Toubro Ltd, Mumbai, India

Sourav Kumar Singh Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India

Teekam Singh Department of Computer Science and Engineering, Graphic Era Deemed to Be University, Dehradun, India

Vineet Kumar Singh Department of CSE (AI), ABES Institute of Technology Ghaziabad, Ghaziabad, India

Lakshay Singla Department of Computer Science Engineering, Thapar Institute of Engineering and Technology, Patialay, India

Nalinda Somasiri York St John University, London, UK

Jyoti Srivastava Madan Mohan Malviya University of Technology, Gorakhpur, India

Ramesh Kumar Sunkaria Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, India

Srutipragyan Swain School of Computer Engineering, KIIT Deemed to Be University, Bhubaneswar, India

Nakkina Sai Teja Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Anubhav Thakur ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India

M. Thyagaraj Nazareth College of Arts and Science, Chennai, India

Anmol Tyagi Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Dev Tyagi Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

Sanjive Tyagi Department of CSE, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India

K. Umapathy Department of ECE, Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya Deemed University, Kanchipuram, Tamil Nadu, India

Jyoti Upadhyay Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, India

D. Vanitha Department of EEE, Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Kanchipuram, Tamil Nadu, India

Manisha Vashisht Lingayas Vidyapeeth, Faridabad, Haryana, India

Vipul Vashisht Lagozon Technologies, Delhi, India

L. Venkata Subhash Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar, India

Priti Verma Sharda University, Greater Noida, Uttar Pradesh, India

Abhinav Vidwans Institute of Advanced Computing, Sage University, Indore, MP, India

Mohammed Joudah Zaiter Electrical Engineering Technical College Middle Technical University, Baghdad, Iraq

Transportation in IoT-SDN Using Vertical Handoff Scheme



Jyoti Maini and Shalli Rani

Abstract The Internet of Things (IoT) refers to the network that connects objects with various processing and sensing capabilities and communicates through the Internet. IoT facilitates connectivity between numerous technological gadgets and sensors to improve our quality of life. It has revolutionized the world around us. IoT has a great impact on transportation as IoT integration allows in growing a centrally managed network that may optimize the space protected via way of means of the vehicle, locate higher and secure routes in case of an important situation, effectively manipulate and keep the goods, fabric, and buy orders and universal offers a fine effect at the sales generated via way of means of the transportation sector. This paper briefly discusses the basics of IoT, the various technologies that comprise IoT, its integration with transportation, and its applications. A vertical handoff scheme is proposed for transportation and it is compared with various states of art techniques. It is observed that it outperforms over the other techniques in terms of throughput and number of end-users served.

Keywords Internet of Things · Technologies · Intelligent transportation · Software-Defined Network · Vertical handoff

1 Introduction

The Internet of Things (IoT) is becoming increasingly popular in our daily lives and can be felt around us [1]. The Internet is a network that connects people to information through which people can communicate. IoT, on the other hand, connects physical objects with separate processing, detection, and activation abilities that share interaction and communication abilities using the Internet [2]. Therefore, the main

J. Maini · S. Rani

Chitkara Institute of Engineering and Technology, Chitkara University, Punjab, India
e-mail: jyoti.maini@ggdsd.ac.in

S. Rani

e-mail: shalli.rani@chitkara.edu.in

purpose of IoT is to permit objects to connect to other objects anytime, anywhere via a network, path, or service. It is an emerging concept that makes our lives easy by introducing communication between sensors and electronic devices that are connected over the Internet. It provides new and effective solutions to various issues and challenges which affect different businesses, governments, and public-private industries around the world [3]. IoT combines several smart devices, intelligent systems, frameworks, and sensors over heterogeneous networks (Fig. 1) [4]. The Internet has developed into a network of countless devices rather than just a network of computers. On the other hand, the IoT is a network of networks because it consists of a variety of “connected” devices [5]. Today, gadgets like smartphones, cars, cameras, toys, buildings, appliances, and industrial systems may share information through the Internet. According to a report published by Lionel Sujay Vailshery, it is expected that the number of IoT devices all over the world will be three times more from 9 billion in 2020 and increase to approximately 25 billion IoT devices in 2030 (Fig. 2). By 2020, China has the most IoT devices with 3.17 billion devices and these devices are deployed in all industries and consumer markets, with approximately more than 55 percent of IoT-connected devices. This percentage is expected to remain as it is for the future next 10 years [6]. In addition, it has made possible the concept of quantum and nanotechnology that were previously impossible [1].

As the involvement of IoT devices and technologies increases, we see a significant transformation in our day-to-day lives. One major evolution of it is the concept of

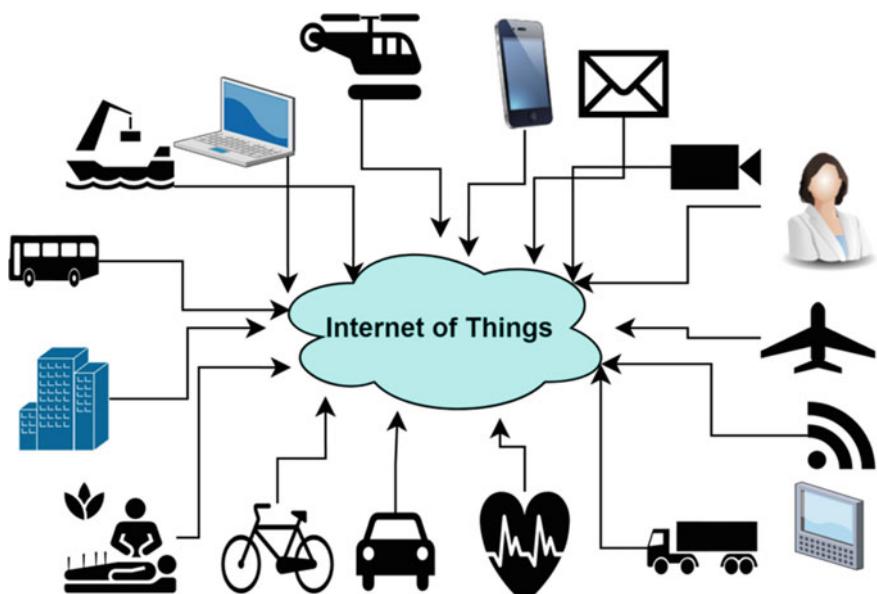


Fig. 1 Internet of Things

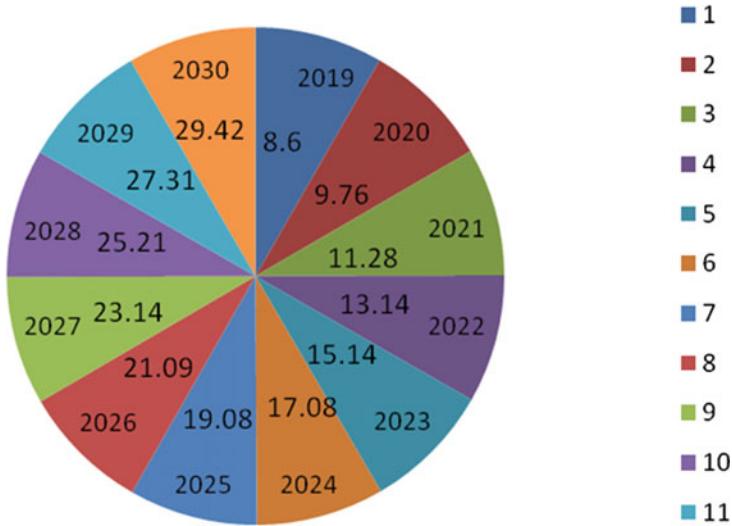


Fig. 2 No. of IoT-connected devices worldwide 2019–2030

(SHSs) smart home systems and devices, which become possible due to Internet-enabled IoT devices that include home automation systems, and energy management systems [3]. Smart health sensing system “SHSS” is another milestone of IoT with the involvement of intelligent human health support devices. These devices help to keep in monitor various health problems and fitness levels. It is also used in hospitals and trauma centers to monitor the critical health conditions of patients. The advanced technology with smart devices results in transforming the entire medical scenario. Similarly, in the transportation industry, IoT devices are used for various purposes, primarily in ticketing, security, and surveillance systems, that provide efficient and secure transportation in urban areas. Transportation in IoT integrates a large network of sensors embedded in systems, controllers, smart objects, and other intelligent devices [7–10].

The remaining paper has been organized into the following sections: Sect. 2 covers the background and related work. In Sect. 3, IoT evolution which has comprised the Internet of Things has been explained. Further in Sect. 4, IoT with transportation has been described. In Sect. 5, intelligent transforming with Software-Defined Network-based vertical handoff scheme has been presented. Section 6 consists of a brief analysis of various algorithms and schemes that have been proposed previously in which discussion and results have been included. Finally, the paper’s conclusion is provided in Sect. 7.

Contribution of the paper: I. In this paper, we tried to relate intelligent transportation with a vertical handoff scheme based on Software-Defined Networks for various different networks. II. Analysis of various handoff schemes (Table 1).

Table 1 No. of IoT-connected devices worldwide 2019–2030

Year	No. of connected IoT devices (in billions)
2019	8.6
2020	9.76
2021	11.28
2022	13.14
2023	15.14
2024	17.08
2025	19.08
2026	21.09
2027	23.14
2028	25.21
2029	27.31
2030	29.42

2 Related Work

Qiang et al. [11] proposed a scheme known as “Quality of Service (QoS) vertical handoff” with Software-Defined Network (SDN) for the different wireless networks. The authors ensured about this scheme that no. of vertical handoffs reduces by maximizing the overall QoS as compared to existing schemes.

In their study [12], Chandana Roy et al. introduced a scheme of edge node selection that has been proposed for an architecture called “Safe-aaS architecture” being used. In this paper, both static and mobile sensor nodes are considered. The count of edge nodes varies within the mobile sensor node’s range, as the geographical location of the vehicle changes. Further, they have talked about each type of edge node having different task executions as well as storage capacities. Therefore, edge nodes should be selected dynamically.

Kim et al. [13] discussed “a service history-based VHO algorithm” that contributes to quality by avoiding unstable VHO decisions in various networks. The authors ensure remarkable performance by minimizing the frequency of handoffs, probability of handoff aborts, and operational cost based on simulated results. Additionally, it is expected to become more productive as networks become larger and more complicated.

Roy et al. [14] proposed a scheme based on service handoff, called “Safe-Passé”, in a 5G environment to the end-users to provide Safety-as-a-Service. This can help the users to make personalized security-related choices. In practice, Safety Service Provider has limited coverage. But, the distance traveled by the end-user requesting service may span the service areas of multiple SSPs. To enable end-users to make safety-related decisions, the authors developed this handoff structure among the several service providers and explained how the service providers interact with one another.

For hybrid 5G environments, Qiang et al. [15] developed a user-centered handoff mechanism. In this mechanism, they have discussed about the solution of the handoff problem that simultaneously increases the receiving data rate and decreases the call block probability. Based on the limited local information available, the user will determine the receiving data rate and estimates the blocking probability of each accessible Base Station (BS) during handoff. Further, the maximization problem by using the throughput metric into account has also been considered. A user can choose the best new BS for a handoff by taking these two factors into account.

Chen et al. [16] introduced a management system in the Internet of Vehicles that can improve resource utilization efficiency and flexibility, improve QoS guarantees, and support multiple simultaneous requests. They have proposed an architecture that integrates advanced technologies including SDN, IoV, and NFV to obtain the future generation Intelligent Transportation System. They have also discussed GBGA and heuristic algorithms, which can give results that are close to optimal, depending on the type of traffic situation.

Liu et al. [17] proposed a service history algorithm that detects interference plus Noise Ratio. The information on service history and the interference plus Noise Ratio in the source network and the equivalent SINR in the target network has been collected according to the traffic features in an attribute matrix. Further, they have talked about the LS method, which is used to determine the weight correlation of various determinants. They have also concluded that this algorithm can achieve decent performance for the users and the networks.

Yang et al. [18] proposed a (MASVH) Multidimensional Adaptive SINR algorithm. This algorithm makes handoff decisions for multi-attribute QoS consideration based on the combined effects of interference plus Noise Ratio, the bandwidth required by the user, user traffic cost, and utilization from participating access networks. They ensured that this algorithm enhances the system performance with elevated throughput and lower outage probability and reduces the cost of user traffic for connecting to converged wireless networks.

“A two-step handoff method based on dynamic weight compensation” that Liu et al. [19] presented functions as a filter function for evaluating network performance. Users will continue to use their current networks if no other networks can satisfy the decision. Users will only hand off to the single network that has passed the pre-decision if there are just a few networks available. If multiple networks fulfill the pre-decision and a user lacks sufficient access, they would switch to a network randomly. A user will carry out the second stage if they have enough power and if multiple networks can pass the pre-decision.

From the above discussions, it is very clear that when using the Software-Defined Network (SDN) approach with a vertical handoff scheme, it performs better than existing schemes and maximizes the total QoS. It chooses and transfers to the best network at the right time, minimizing the number of vertical handoffs.

3 IoT Evolution

IoT is a worldwide information infrastructure that connects things (both real and virtual) using interoperable information and communication technologies to allow new services. In IoT, communications extend to everything that surrounds us via the Internet. It is not only a machine-to-machine communication, but lots more than this. It includes various technologies such as WI-FI, GPS, GSM, GPRS, sensor networks, WSN, 2G, 3G, 4G, and RFID, along with other similar technologies. These technologies enable IoT applications possible. IoT is not a single but collaborates both hardware and software technologies and provides extensive IT-based solutions. It refers to communication technologies, used to store, retrieve, and process data, and the electronic systems used to communicate with each other [20]. The core technologies for the Internet of Things are shown (Fig. 3).

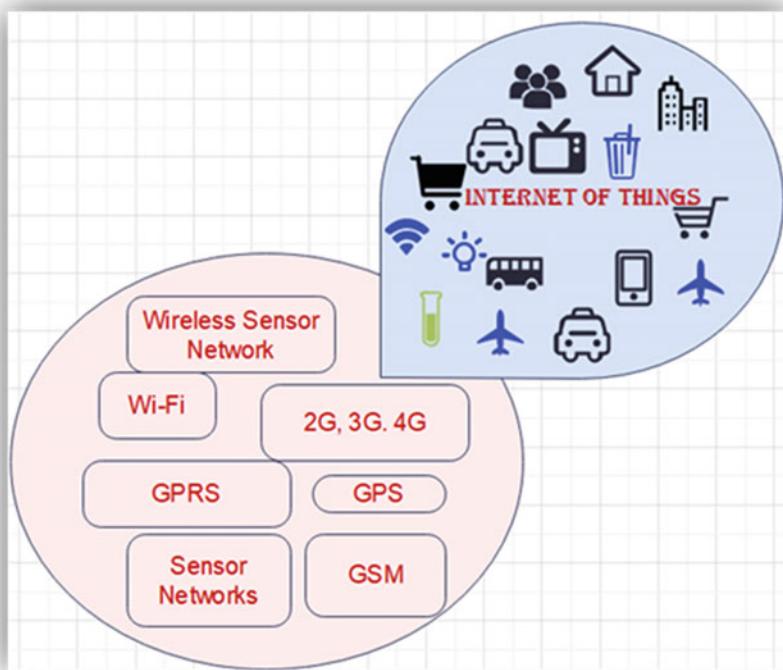


Fig. 3 Internet of things: IoT evaluation

4 IoT with Transportation

The transportation sector has a great role in the development of countries. Transportation is needed in many aspects like the supply of goods and services, logistics and movement of people, etc. Thus, transportation becomes an important part of the logistical connection between the customer and the supply chain [21]. Logistics refers to the process of managing how the material is brought in, managed, and delivered to its destination. Its management includes identifying future vendors and suppliers and understanding their strengths and accessibility [22]. IoT integration allows in growing a centrally managed network that may optimize the space protected via way of means of the vehicle, locate higher and secure routes in case of an important situation, effectively manipulate and keep the goods, fabric, and buy orders and universal offers a fine effect at the sales generated via way of means of the transportation sector. There are lots of problems like security, accountability, service reliability, convenience issues in navigation, an appropriate route or network among various routes, the cost of services, etc. in today's transportation that affects the development of the transportation sector. So the concept of the communication system in transportation is evolved which could be possible with IoT in the transportation industry.

4.1 IoT in Transportation: Applications

In the transportation business, IoT has many opportunities. IoT can be used to track vehicles' speed, position, whether they are moving or not, and other factors including whether they are in danger, etc. In the vast majority of situations, vehicles are deployed for transportation or to carry heavy weights. IoT also offers vehicle navigation management solutions for tracking and managing transportation (road, air, water, and transport) [23]. Let us have a brief look at the different applications of IoT in transportation. Traffic management: Where the deployment of IoT technologies is observed. IoT in transportation offers automated tolls as compared to traditional tolling and ticketing systems. The management of tolls and tickets has become simpler for traffic police officers due to RFID tags and other smart sensors. With the addition of IoT technology, the idea of self-driving automobiles has been transformed from a futuristic concept into a cutting-edge reality. With little to no human interaction, these vehicles can move securely by sensing their surroundings. Vehicle tracking or transportation monitoring technologies have become a need for businesses to efficiently manage supply chain operations. It can be made feasible using GPS Trackers. IoT in transportation provides the security of public transport. IoT-enabled devices help traffic police and municipalities to keep an eye on every transport and track traffic violations and take appropriate actions (Fig. 4). In this way, these are just a few examples of IoT's applications affect the transportation sector, but IoT will continue to make transportation smarter in the future [24].

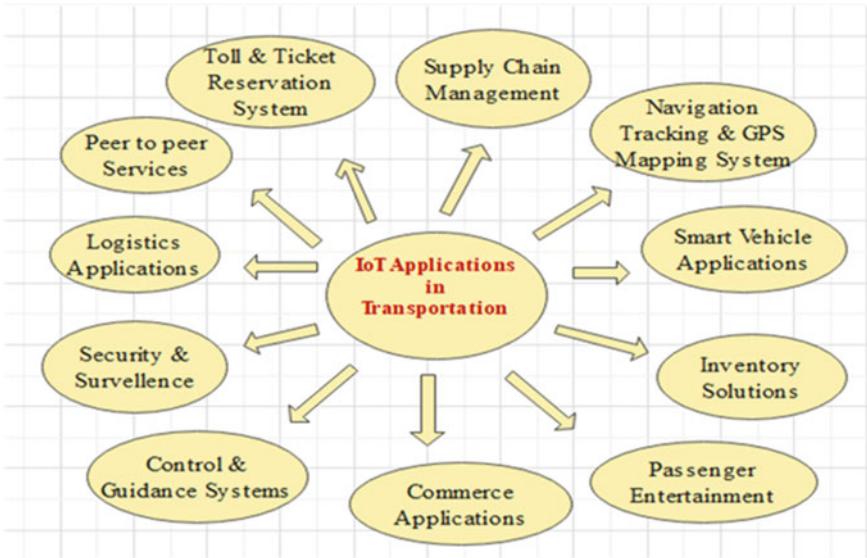


Fig. 4 Applications of IoT transportation

5 Intelligent Transportation Using a Vertical Handoff Method Based on Software-Defined Networks

In this research, we suggested an intelligent transportation scheme using Software-Defined Network (SDN)-based vertical handoff strategy for various routes. In networks with several routes, transportation requires switching inter-network connections from one route to another. In the vertical handoff, the process of shifting has been accomplished over many networks. Two significant concerns must be resolved before the vertical handoff. One is responsible for selecting a route for the transport vehicle. The second step is determining the handoff timing: or the moment at which the interconnected route transfers should be executed. Implementing SDN technology enables the resolution of these two vertical handoff concerns. SDN is a revolutionary networking paradigm that enables the worldwide centralization of network device management. In this research, we explore the vertical handoff issue in heterogeneous wireless networks using this SDN capability. This can be applicable to intelligent transportation. In this system, transportation may choose the optimal path to avoid congestion. SDN controller selects network routes for transport during the initiation, constructs the request matrix, and network selection stages of the vertical handoff network selection strategy [11].

- A. Strategy Phases: First of all, the current route will be evaluated for congestion as well as for timing to reach the destination through some GPS Tracker or IoT device and choose whether to start vertical handoffs or not. A vehicle will transmit request frames to the available network routes if it wants a vertical

handoff. The QoS values of the associated networks are represented by the values of request frames. In addition, even if the vertical handoff is not necessary, the vehicle transmits the request frames to its accessible networks. After receiving the information of different routes, The SDN controller prepares information in the form of matrix. The SDN controller chooses routes for the vehicle after creating the matrix.

- B. Handoff Timing: Because the vehicles are moving, they should not immediately switch to their chosen paths. The vehicle will have to determine how much distance to be covered and at what time to reach the destination. If the selected route will have a long distance and will take more time as compared to the current route, then the vehicle will not handoff. Vehicles only handoff to their chosen networks if the selected network routes are always more suitable than their existing networks. The rationale behind this is that, even if the chosen network is superior to the current one, if the latter can meet a vehicle's demand, no vertical handoff should be made by that vehicle.

6 Brief Analysis of Various Proposed Schemes and Results

We compare the various existing schemes VHO [13], Safe Passé [14], UCH [15], Multiplicative [25], Two-Step Scheme [19], and Multi-Objective Scheme [26] based on the ratio of end-users served and throughput. First, we analyzed the percentage of no. of users served over a certain time period in (Fig. 5). It has been noticed that with an increase in no. of end-users from 20 to 220, there is a declined trend, which can be due to handoffs, but in Safe-Passé Scheme, the rate of decrease in the ratio is less compared to other schemes. Then, we analyze the second concept of throughput in (Fig. 6). It depicts the increasing trend line of throughput of Safe-Passé's scheme, with an increase in the no. of end-users.

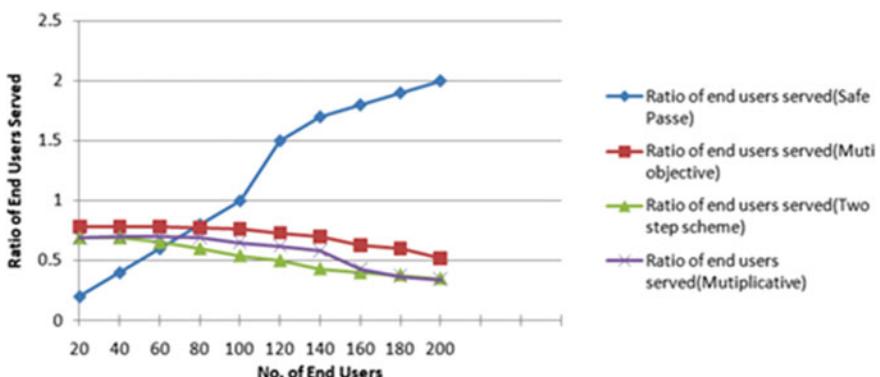


Fig. 5 Analysis of ratio of end-users served

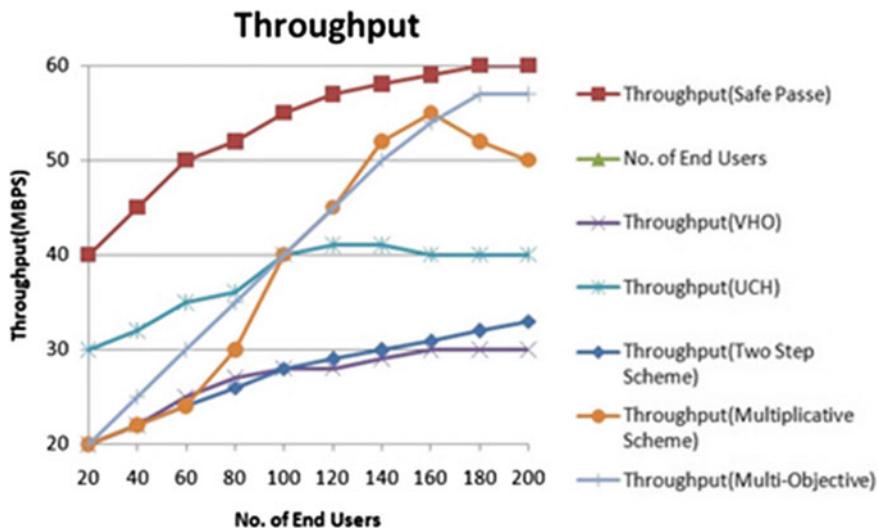


Fig. 6 Analysis of throughput

7 Conclusion

Internet of Things (IoT), an innovative concept, has a crucial role in the area of transportation. In this paper, the integration of IoT with transportation has been discussed. Further, we have proposed the scheme of selecting the appropriate route with the help of the SDN controller and if the chosen network routes are consistently better than their current networks in case of congestion or more traffic, a vehicle handoff scheme is applied to their selected networks. Further, we conducted a comparison of vertical handoff schemes with other schemes based on the ratio of users served and throughput. Then, we conclude that Safe-Passé's scheme is the best scheme till now, which gives better performance as compared to the other schemes despite handoffs. We would like to implement our proposed SDN-based vertical handoff scheme in the future for intelligent transportation in real life to make it smarter for smart cities.

References

- Gatsis K, Pappas GJ (2017) Wireless control for the IoT: power, spectrum, and security challenges. In: 2017 IEEE/ACM Second International conference on internet-of-things design and implementation (IoTDI). IEEE, pp. 341–342
- Jha S et al (2019) Deep learning approach for software maintainability metrics prediction. IEEE Access 7:61840–61855
- Ghanem S, Kanungo P, Panda G et al (2021) Lane detection under artificial colored light in tunnels and on highways: an IoT-based framework for smart city infrastructure. Complex Intell Syst. <https://doi.org/10.1007/s40747-021-00381-2>

4. Sharma S, Gupta K, Gupta D (2021) The amalgamation of internet of things and recommender systems. *J Phys: Conf Ser* 1969:012040
5. Sharma R, Kumar R, Sharma DK, Son LH, Priyadarshini I, Pham BT, Bui DT, Rai S (2019) Inferring air pollution from air quality index by different geographical areas: case study in India. *Air Qual Atmos Health* 12:1347–1357
6. Vailshery LS (2022) Number of internet of things (iot) connected devices worldwide from 2019 to 2021, with forecasts from 2022 to 2030. <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide>
7. Agnihotri S, Ramkumar K (2021) IoT and healthcare: a review
8. Kumar A, Sharma S, Goyal N, Singh A, Cheng X, Singh P (2021) Secure and energy-efficient smart building architecture with emerging technology IoT. *Comput Commun* 176:207–217
9. Nguyen PT, Ha DH, Avand M, Jaafari A, Nguyen HD, Al-Ansari N, Van Phong T, Sharma R, Kumar R, Le HV, Ho LS, Prakash I, Pham BT (2020) Soft computing ensemble models based on logistic regression for groundwater potential mapping. *Appl Sci* 10:2469
10. Sachan S, Sharma R, Sehgal A (2021) SINR based energy optimization schemes for 5G vehicular sensor networks. *Wireless Pers Commun.* <https://doi.org/10.1007/s11277-021-08561-6>
11. Malik PM, Sharma R, Singh R, Gehlot A, Satapathy SC, Alnumay WS, Pelusi D, Ghosh U, Nayak J (2021) Industrial internet of things and its applications in industry 4.0: state of the art. *Comput Commun* 166:125–139. ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2020.11.016>
12. Sachan S, Sharma R, Sehgal A (2021) Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. *Sustain Comput Inf Syst* 30:100504
13. Kim T, Han SW, Han Y (2010) A QOS-aware vertical handoff algorithm based on service history information. *IEEE Commun Lett* 14(6):527–529
14. Vo MT, Vo AH, Nguyen T, Sharma R, Le T (2021) Dealing with the class imbalance problem in the detection of fake job descriptions. *Comput Mater Continua* 68(1):521–535
15. Sharma R, Kumar R, Satapathy SC, Al-Ansari N, Singh KK, Mahapatra RP, Agarwal AK, Le HV, Pham BT (2020) Analysis of water pollution using different physicochemical parameters: a study of Yamuna river. *Front Environ Sci* 8:581591. <https://doi.org/10.3389/fenvs.2020.581591>
16. Chen J, Zhou H, Zhang N, Xu W, Yu Q, Gui L, Shen X (2017) Service-oriented dynamic connection management for software-defined internet of vehicles. *IEEE Trans Intell Transp Syst* 18(10):2826–2837
17. Sabeeha Begam S, Vimala J; Selvachandran G, Ngan TT, Sharma R (2020) Similarity measure of lattice ordered multi-fuzzy soft sets based on set theoretic approach and its application in decision making. *Mathematics* 8:1255
18. Yang K, Gondal I, Qiu B (2008) Multi-dimensional adaptive sinr based vertical handoff for heterogeneous wireless networks. *IEEE Commun Lett* 12(6):438–440
19. Liu C, Sun Y, Yang P, Liu Z, Zhang H, Wen X (2013) A two-step vertical handoff decision algorithm based on dynamic weight compensation. In: 2013 IEEE International conference on communications workshops (ICC). IEEE, pp. 1031–1035
20. Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I, Le D-N (2020) Global forecasting confirmed and fatal cases of COVID-19 outbreak using autoregressive integrated moving average model. *Front Public Health* 8:580327. <https://doi.org/10.3389/fpubh.2020.580327>
21. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Tien BD, Priyadarshini I, Sarkar M, Le T (2020) Crime rate detection using social media of different crime locations and Twitter part-of-speech tagger with brown clustering, pp 4287–4299
22. Kenton W (2022) Logistics: what it means and how businesses use it. <https://www.investopedia.com/terms/l/logistics.asp>
23. Dansana D, Kumar R, Parida A, Sharma R, Adhikari JD et al (2021) Using susceptible-exposed-infectious-recovered model to forecast coronavirus outbreak. *Comput Mater Continua* 67(2):1595–1612

24. 5 Major applications of IoT in transportation. <https://www.conurets.com/5-major-applications-of-iot-in-transportation>. 28 Feb 2022
25. Sharma R, Kumar R, Singh PK, Raboaca MS, Felseghi R-A (2020) A systematic study on the analysis of the emission of CO, CO₂ and HC for four-wheelers and its impact on the sustainable ecosystem. *Sustainability* 12:6707
26. Deb K (2014) Multi-objective optimization. *Search Methodol* 2014:403–449

MLP-Based Speech Emotion Recognition for Audio and Visual Features



G. Kothai, Prabhas Bhanu Boora, S. Muzammil, L. Venkata Subhash, and B. Naga Raju

Abstract Due to its potential applications in domains involving psychology, social smart machines, and human–computer interaction, speech emotion recognition has become an important and growing topic field in the study. That article offers the service comparative analysis of strategies in artificial intelligence classifiers for speech emotion recognition, including MLP, decision tree, SVM, and random forest. Using visual analysis methods like wave plot and spectrogram analysis as well as audio feature extraction, we evaluate these classifiers. According to our observations, MLP achieves better performance than the other classifiers, obtaining an accuracy of 85% when identifying different emotions. Moreover, we illustrate how audio feature extraction and visual analysis help to improve emotion recognition efficiency. Our research has applications for creating speech emotion recognition algorithms that may be applied in real-life scenarios.

Keywords Speech emotion recognition (SER) · MLP · MEL · MFCC · CHROMA · SVM · Decision tree · Random forest · Visual analysis

G. Kothai

Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning),
KPR Institute of Engineering and Technology, Uthupalayam, Coimbatore, Tamil Nadu, India

P. B. Boora (✉) · S. Muzammil · L. Venkata Subhash · B. Naga Raju

Department of Computer Science and Engineering, Kalasalingam Academy of Research and
Education, Krishnankoil, Virudhunagar, India

e-mail: 9920004615@klu.ac.in

S. Muzammil

e-mail: 9920004629@klu.ac.in

L. Venkata Subhash

e-mail: 9920004650@klu.ac.in

B. Naga Raju

e-mail: 9920004633@klu.ac.in

1 Introduction

An important area of study in the domains of signal processing and artificial intelligence is speech emotion recognition (SER). Our understanding of human communication and social interactions can be enhanced, in addition to the efficiency of speech-based human software applications, due to the capability to demonstrate emotions in speech. In recent years, digital computing approaches have been widely used to overcome the SER problem, with extensive research showing promising results. The Vanilla neural networks [1], decision tree [2], support vector networks [3], and random forest [4] are a few machine learning classifiers for speech emotion recognition that we compare and evaluate in this paper.

Our goals are to determine the best classifier for this paper and to provide a thorough analysis of the aspects that impact the performance of the classifier. To do this, we evaluated our techniques using a dataset of voice recordings that had been classified into emotional categories. Along with using wave plots and spectrogram plots to analyze the audio data, we also employed confusion matrices, accuracy graphs, and wave plots to evaluate classifier performance.

Our results indicate that the MLP classifier performed significantly better than the other classifiers, achieving approximately 85% accuracy. We also identified that specific audio features, such as timing and amplitude, were essential for accurately recognizing emotions in speech. Our results give valuable information into the SER problem and demonstrate the possibilities of machine learning approaches or classifiers for this important application. We also provide unique capabilities like audio-to-text conversion and visual analysis of wave plots and spectrogram plots. Our results demonstrate that the MLP classifier can properly recognize emotions in spoken language. This research will have far effects on the creation of speech-based technologies able to recognize and react to human emotions.

The following list summarizes this paper's main contributions:

- We examine the multi-layer perceptron against several machine learning classifications and offer an analysis.
- The RAVDESS dataset to acquire MFCC, CHROMA, and MEL from audio files along with emotions is mentioned.
- With the materials given and selected at random audio specimens and generated wave plots, spectrograms.
- Execute repetitions and start our model's MLP classifier using particular operations.
- Every class utilized in the training models incorporates the confusion matrices.
- Generating a categorization report, a framework of observed reactions (which combines Actual and Predicted), and we also assess the preciseness of all classes employed.
- Supply the creator's recording so that it may identify the speech and construct text.

2 Review of Literature Research

Khan et al. [5] Speech Clustering versus MI emotion acknowledgment in which a real-time database is used to discern emotions from speech signals. The findings show that the K-NN classifier has an average forward feature selection accuracy of 91.71%, whereas the SVM classifier has an accuracy of 76.57%. SVM classification for neutral and fear emotions is significantly superior to K-NN.

Tripathi et al. [6] presented numerous ConvNet structures that could be with voice characteristics and notations are proposed in this research. When paired with text, a two-dimensional neural network with voice features performs more accurately when compared to current best practices. Group frequency is 69.5% and a result of 75.1% is offered by the consolidated spectrum analyzer.

Huang et al. [7] proposed a technique for automatically extracting employing Bayesian belief nodes (DBNs), a form of most advanced neural networks, which can retrieve the affective attribute component into the emotional audio files. Fully connected channels (DBNs) and CV machines were merged into a classifier model that has been published with SVM (SVM). As a result of this method's ability to precisely extract emotion characteristic characteristics throughout the practical training phase, the model's final recognition rate—which is 7% higher than that of standard artificial extract—has increased significantly.

Kerkeni et al. [8] classified seven emotions by an autonomous mechanism for recognizing feelings in conversation employing three procedures for computer vision (multiple regressions, SV, and Vanilla Networks). Once the elements are exposed to featured picking and presenter standardization, all of the classifiers for the Berlin database attain a precision of 83%. This result demonstrates that given larger info, the hidden layer typically appears to be better, as it has features drawback involves enduring lengthy practice sessions.

Ingale et al. [9] explained the systems for recognizing speech emotions based on several classifiers. The signal processing unit, which extracts pertinent features from the available speech signal, and a classifier, which discerns emotions from the speech signal, are the two key components of a language processing tool for sentiments. Also, the accuracy of the speech emotion identification system can be improved by extracting additional useful speech elements.

Bertero et al. [10] introduced the emotion and sentiment recognition module for an interactive conversation system. Two components such as speech emotion and sentiment recognition were mainly concentrated. It is possible to achieve real-time performance on voice emotion recognition at 65.7% accuracy. Moreover, CNN sentiment analysis yields an 82.5 F-measure when trained on non-domain data.

Mirsamadi et al. [11] presented various RNN architectures for learning features for voice and expression mapping. This became demonstrated that users might discover descriptions at the required times, utilizing deep RNNs in addition to the aggregated spatial duration periods. Studies employing IEMOCAP data show that learned features have a higher classification accuracy than fixed-designed features used in typical SVM-based SER.

Khalil et al. [12] did the review utilizing pattern recognition to distinguish vocal sound techniques and additionally considered similarly recurrent neural networks, belief networks, and the convolutional layer. DC network has the highest accuracy by predicting Anger, Happy, and Sad with 98%.

Daniel et al. [13] took the databases to models and predicted the emotion which is given through speech with the help of some DL approach. CNN and LSTM algorithms are used for the models. For recognizing the emotion, they considered HMM, SVM, and ANN. Compared with other models, Support Vector Machine got more accuracy with 94% and worked as the most preferable algorithm.

Margaret et al. [14] prepared the architectures on SER by evaluating the DL techniques and proposed minimal speech processing by using deep learning. By using a multi-layered neural network, some speech spectrograms are filtered. Feedforward and backpropagation had been done to the neural networks. CNN has performed with discriminated architectures than the others.

Prasanna et al. [15] proposed for defining speech emotion; MFCC is used with two popular datasets known as EMO-DB and IEMOCAP. Constructed the SoftMax layer with fully connected layers and the bi-directional LSTM is defined. Confusion matrix for EMO-DB and IEMOCAP differs. By training the testing, the accuracy for EMO-DB has the highest. CNN and LSTM both are combined and predicted for the Mel-Frequency Cepstral Coefficients.

Maros et al. [16] with the attention mechanism and deep learning did the research and gave the review based on their performance. Human interactions are the basic part of emotions. The SER plays a vital role in developing the interactions between humans and computers. Attention-based neural networks are used for extracting the information which has been distributed in content. For data processing, LSTM and GRU networks are used. LSTM got the best accuracy.

Lim et al. [17] using recurrent and convolutional networks worked on the image, speech, and music recognition. Applied the proposed algorithms which are useful for developing the speech database with emotion and the obtained results are classified and compared with CNN methods. CNN purpose has the advantage of taking the properties of the signals. RNN is used for the computation in forward computation. CNN shows the results with distributed networks.

Gupta et al. [18] considered both speech signals and text information as the best features in threshold fusion. There are so many classifiers like Gaussian Mixture Models, Neural Networks, and Maximum likelihood which are generally used for classifying the models for speech emotion recognition. For obtaining the maximum accuracy, SVM is chosen. For normalizing and thresholding fusion, it gives the individual and actual results.

Kwon et al. [19] proposed the Hierarchical Convolutional Neural Architecture in deep learning. The input is given after obtaining the proposed system, it has been checked with Reyson's multi-media resource which includes authentic text and linguistic corpus, along with engaging expressive reciprocal gesture recognition. An 87% accuracy rate was obtained for this model.

3 Problem Statement

The objective is to develop a deep learning model MLP that can effectively diagnose a psychological situation behind the narrator that can be observed in its statement signal given a dataset of speech samples and appropriate emotion labels. A preprocessed group of acoustic data, including prosodic features, the best classifier for this issue still has to be developed, and performance-related problems should be considered carefully.

3.1 *Dataset Description*

There are nearly 1440 utterances in the speech recordings. The utterances vary in duration from 1.5 to 5 s, and the sample includes spoken word transcriptions. For academics focusing on speech and music emotion recognition, sentiment analysis, and relevant fields, the RAVDESS dataset provides an extensive range of resources.

3.2 *Dataset Details*

The dataset has been collected from Kaggle. The Ravdess dataset consists of many audio speech files, in which the duration of audio is nearly from 0.5 to 1.5 s. Emotions in the Ravdess dataset are Surprised, Disgusted, Happy, Calm, Angry, Neutral, Fearful, and Sad. Research works engaged in speech and musical emotion recognition, affective computing, and related topics can utilize the RAVDESS dataset as a rich and comprehensive resource.

4 Proposed System

See Fig. 1.

4.1 *Data Exploration*

To comprehend the features of the dataset and the correlations between the attributes and the target emotions, data exploration is a crucial stage in speech emotion recognition (SER). Here are a few common approaches for data exploration in SER:

Visualization: Exploring data through visual representation is rather effective. Patterns and connections between features and emotions can be detected using plots

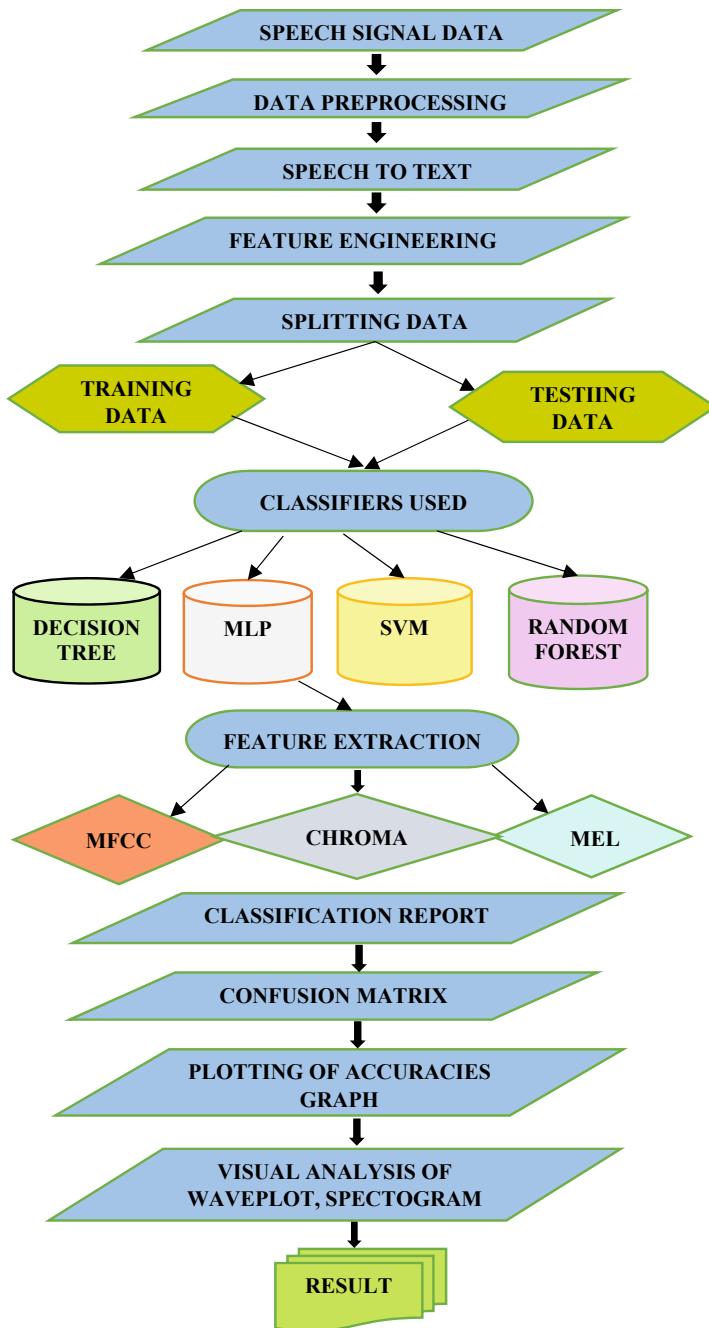


Fig. 1 Proposed system for SER

such as histograms, scatter plots, and box plots. For example, a box plot can demonstrate whether there are variations in feature distributions for emotional responses, while a scatter plot of two features can illustrate whether they are connected.

Correlation Analysis: To figure out the extent and axis with the straight connection that exists in the features, correlation analysis may be applied. To demonstrate that all feature pairs are correlated, a correlation matrix can be generated.

4.2 Feature Extraction

Describing the emotion recognized (SER) involves extracting important information from the speech signal. Feature extraction is highly essential to SER because it provides a way to recognise the unique features of the expressed emotional responses. There are an enormous number of often-used SER parameters, such as:

Mel-Frequency Cepstral Coefficients (MFCCs): These features combine the methods for determining the spectral properties of speech. To determine a set of parameters that are related to the convolution shell in the audio signal, they are established by determining the baseline of the speech signal's spectrum followed by applying the requirements of new approaches.

CHROMA: The process of extracting CHROMA features is applied widely to audio analysis. CHROMA features are always a useful method in speech emotion recognition because they can be utilized to record audio characteristics of speech and help differentiate between several emotions.

MEL: Although MEL frequency features are so capable of gathering the spectrum information of the audio stream, they are frequently employed in speech emotion recognition. Other speech features, including (duration, intensity) and spectral features (such as MFCCs), can be merged with MEL frequency which can boost voice-identifying application functionality.

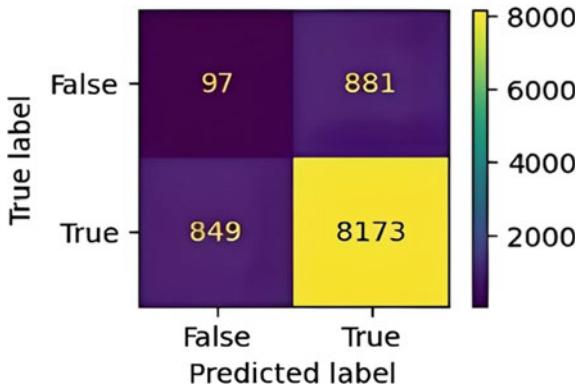
Pitch: These features capture information from the soundtrack, phrasing, and result of the pressure in speech. They are often used to draw boundaries between various feelings, such as anger and sadness.

Spectral Features: Spectral features have included the spectral midpoint, flatness, roll-off, and gradient. They record details regarding the energy distribution of the speech signal across several frequencies.

Features of the Time Domain: These include the zero-crossing rate, energy, and duration. They track information on the speech's spectral characteristics.

Formant Features: They collect information on the frequency ranges of the speech tract. They may be used to differentiate between various sounds as well as recognize emotions like surprise or fear. The speech signal is often divided into approximately

Fig. 2 MLP confusion matrix



30 ms-long frames to obtain these features, with overlapping filters also used to record information about phase movements.

5 Classifiers

5.1 Multi-layer Perceptron

A specific artificial neural network used within machine learning applications is the multi-layer perceptron (MLP). The word actually “MLP” is used to describe a specific type of artificial neural network consisting of numerous layers of interconnected nodes or neurons. For tasks involving supervised learning like classification and **regression**, the MLP model is frequently implemented. The true, false positive, and negative events are mentioned in Fig. 2.

5.2 Support Vector Machine

The ability of Support Vector Machines (SVMs) to integrate high-dimensional feature spaces and nonlinear relationships among input characteristics and output labels has resulted in their wide implementation in speech emotion recognition (SER), used as input features in SER. To classify speech signals into specified emotion classifications. SVMs are used in regression to accurately predict an emotion’s intensity on a continuous scale. The true positive, false negative values are shown in Fig. 3.

Fig. 3 SVM confusion matrix

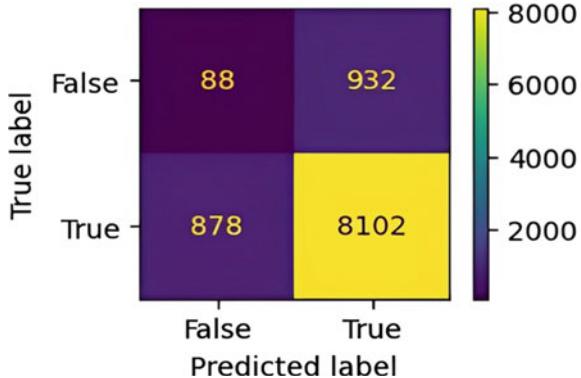
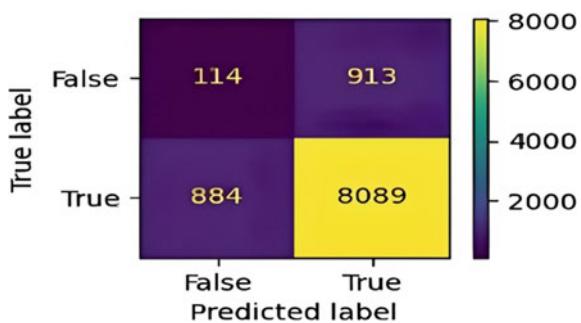


Fig. 4 Random forest matrix



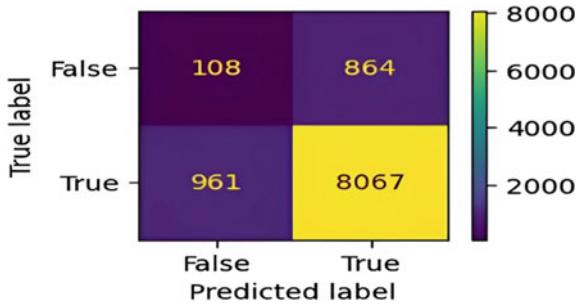
5.3 Random Forest Classifier

This classifier is popular in the machine learning approach that has been exploited in speech emotion recognition (SER). Due to its propensity for managing high-dimensional data and complicated connections among input features and output labels, the input parameters for SER are generally voice signals or extracted sound virtues such as prosodic features, and other spectrum characteristics. Depending on the emotion recognition challenge, the output may be segmented or continuous. Figure 4 consists of correlation matrix precision.

5.4 Decision Tree

Each leaf node in a decision tree provides a class label or a regression value, and each internal node in a decision tree provides a decision based on a specific feature. Enhance the information gain or the Gini impurity, the algorithm selects the most effective feature during training. The decision tree algorithm produces a prediction based on the class label or regression and provides the review with the leaf node

Fig. 5 Decision tree confusion matrix



after traversing the leaf node depending on values, and input attributes during testing (Fig. 5).

6 Experimental Results

By evaluating all the results, the deep learning classifier multi-layer perceptron acquired the highest accuracy of 85% compared to other machine learning classifiers. MLP has many advantages which are not present in other classifiers (Figs. 6, 7, 8, 9, 10, 11, 12, and 13).

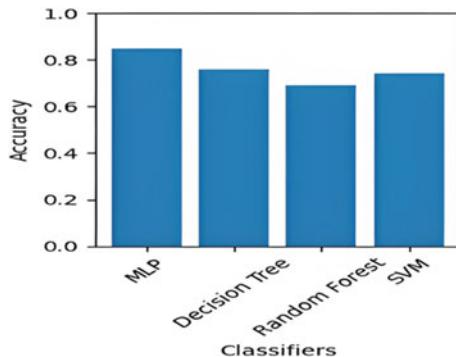


Fig. 6 The accuracy scores of classifier algorithms

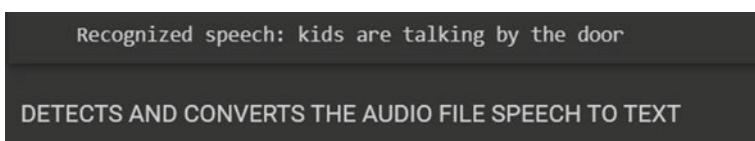
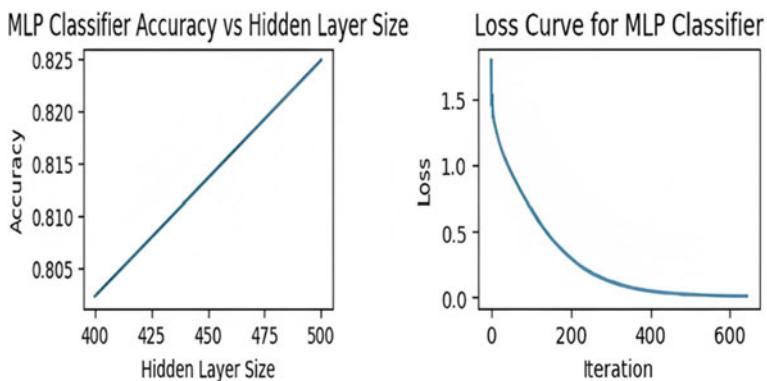


Fig. 7 Converts the audio speech file to text

	calm	disgust	fearful	happy	accuracy	macro avg	weighted avg
precision	0.812500	0.872340	0.844444	0.837838	0.841808	0.841781	0.843094
recall	0.928571	0.836735	0.826087	0.775000	0.841808	0.841598	0.841808
f1-score	0.866667	0.854167	0.835165	0.805195	0.841808	0.840298	0.841127
support	42.000000	49.000000	46.000000	40.000000	0.841808	177.000000	177.000000

Fig. 8 The classification report of the SER**Fig. 9** Graph between MLP hidden layer size and classifier accuracy**Fig. 10** How the SER predicts compared to actual

	Actual	Predicted
0	disgust	disgust
1	calm	calm
2	calm	calm
3	disgust	disgust
4	disgust	calm
5	disgust	disgust

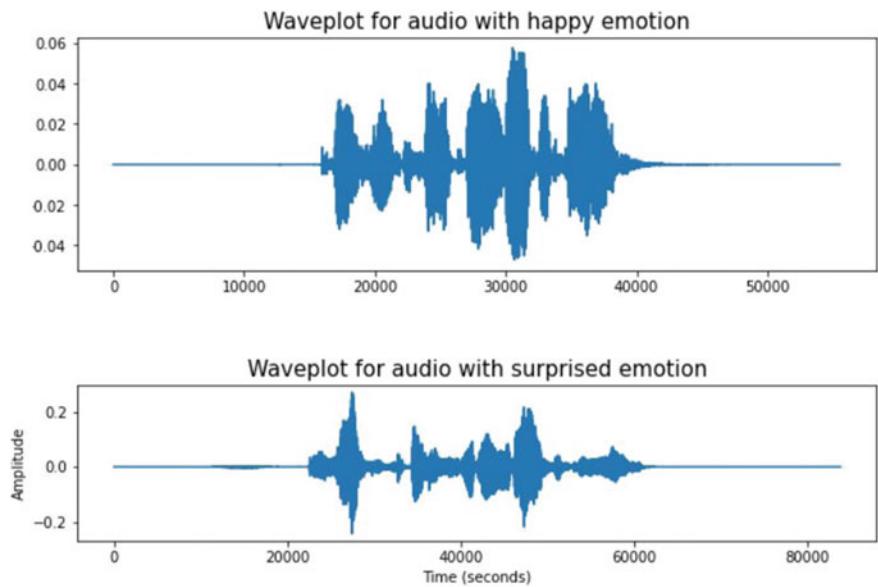


Fig. 11 Wave plots for audio with happy, surprised emotion

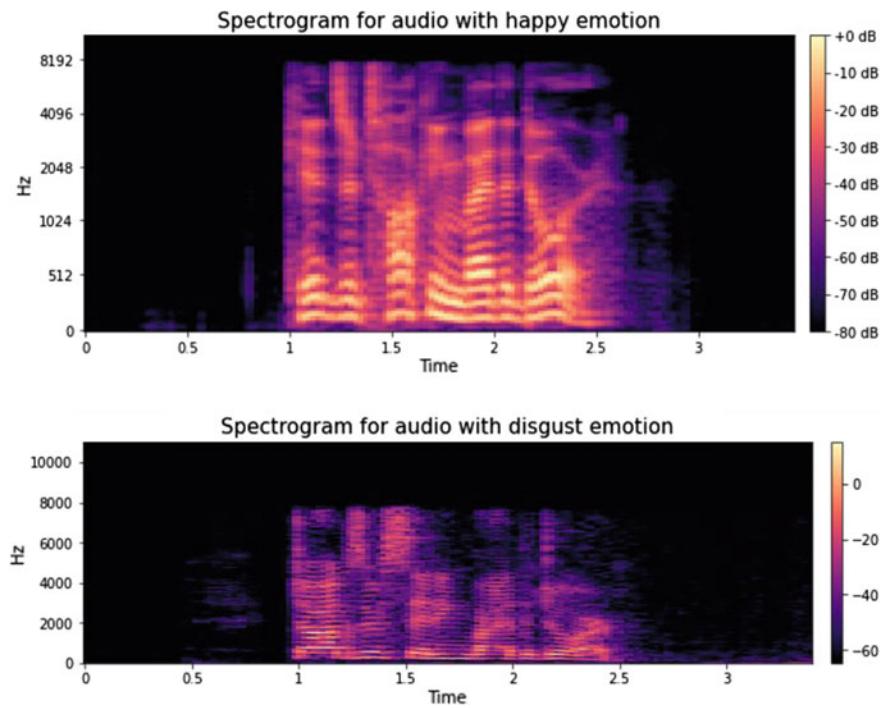


Fig. 12 Spectrogram for happy and disgusted emotions

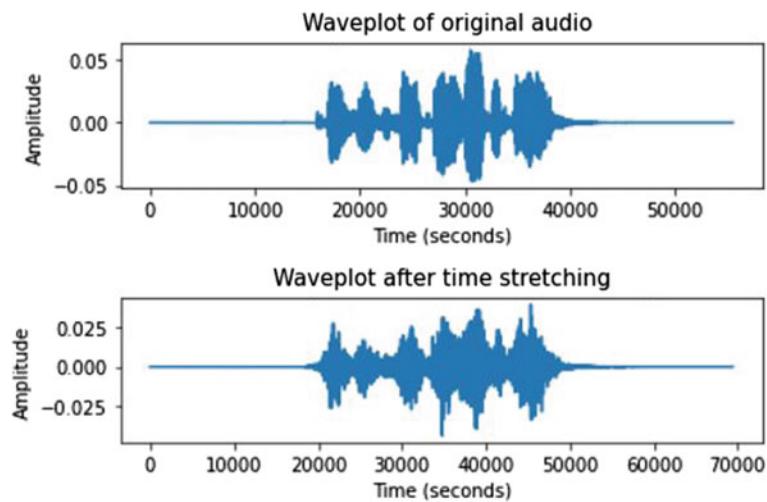


Fig. 12 (continued)

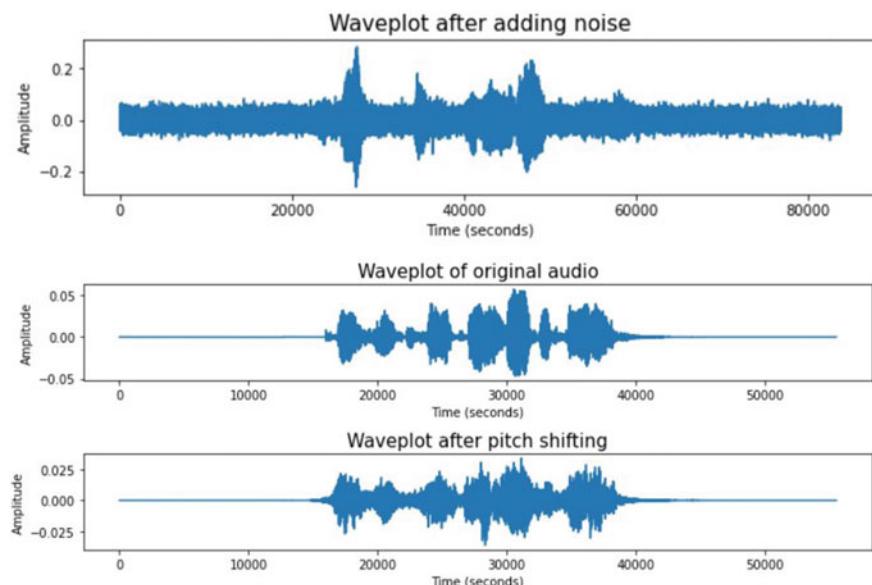


Fig. 13 Representation of noise and pitch

7 Conclusion

In conclusion, recognizing the emotions in speech is a crucial activity with a wide range of applications in domains like human–computer interaction, health care, and learning. Several other models, including SVM, random forest, decision tree, and MLP, are developed in recent years, and techniques based on machine learning have become commonly employed for SER. According to research and testing, MLP extensively exceeds SVM, decision trees, and random forests in terms of accuracy. This is an outcome of MLP’s ability to recreate complex connections between characteristics and feelings. It can generalize well to previously unexplored data and handle massive data with high dimensions.

References

- Chattopadhyay S, Dey A, Basak H (2020) Optimizing speech emotion recognition using manta-ray-based feature selection. arXiv preprint [arXiv:2009.08909](https://arxiv.org/abs/2009.08909)
- Wani TM, Gunawan TS, Qadri SAA, Kartiwi M, Ambikairajah E (2021) A comprehensive review of speech emotion recognition systems. IEEE Access 9:47795–47814
- Al Dujaili MJ, Ebrahimi-Moghadam A, Fatlawi A (2021) Speech emotion recognition based on SVM and KNN classifications fusion. Int J Electr Comput Eng 11(2):1259
- Fang Y, Yang H, Zhang X, Liu H, Tao B (2021) Multi-feature input deep forest for EEG-based emotion recognition. Front Neurorobot 14:617531
- Khan M et al (2011) Comparison between KNN and SVM method for speech emotion recognition. Int J Comput Sci Eng 3(2):607–611
- Tripathi S et al (2019) Deep learning based emotion recognition system using speech features and transcriptions. arXiv preprint [arXiv:1906.05681](https://arxiv.org/abs/1906.05681)
- Huang C et al (2014) A research of speech emotion recognition based on deep belief network and SVM. Math Probl Eng
- Kerkeni L et al (2019) Automatic speech emotion recognition using machine learning. In: Social media and machine learning. IntechOpen
- Ingale AB, Chaudhari DS (2012) Speech emotion recognition. Int J Soft Comput Eng (IJSCE) 2(1):235–238
- Bertero D et al (2016) Real-time speech emotion and sentiment recognition for interactive dialogue systems. In: Proceedings of the 2016 conference on empirical methods in natural language processing
- Mirsamadi S, Barsoum E, Zhang C (2017) Automatic speech emotion recognition using recurrent neural networks with local attention. In: 2017 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2227–2233
- Khalil RA, Jones E, Babar MI, Jan T, Zafar MH, Alhussain T (2019) Speech emotion recognition using deep learning techniques: a review. IEEE Access 7:117327–117345
- Abbaschian BJ, Sierra-Sosa D, Elmaghriby A (2021) Deep learning techniques for speech emotion recognition, from databases to models. Sensors 21(4):1249
- Fayek HM, Lech M, Cavedon L (2017) Evaluating deep learning architectures for speech emotion recognition. Neural Netw 92:60–68
- Pandey SK, Shekhawat HS, Prasanna SM (2019) Deep learning techniques for speech emotion recognition: a review. In: 2019 29th International conference radioelektronika (RADIOELEKTRONIKA). IEEE, pp 1–6
- Lieskovská E, Jakubec M, Jarina R, Chmuřík M (2021) A review on speech emotion recognition using deep learning and attention mechanism. Electronics 10(10):1163

17. Lim W, Jang D, Lee T (2016) Speech emotion recognition using convolutional and recurrent neural networks. In: 2016 Asia-Pacific signal and information processing association annual summit and Conference (APSIPA). IEEE, pp 1–4
18. Gupta S, Mehra A (2015) Speech emotion recognition using SVM with thresholding fusion. In: 2015 2nd International conference on signal processing and integrated networks (SPIN). IEEE, pp 570–574
19. Kwon S (2020) CLSTM: Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network. Mathematics 8(12):21

Drain Current and Transconductance Analysis of Double-Gate Vertical Doped Layer TFET



Mandeep Singh, Nakkina Sai Teja, Tarun Chaudhary, Balwinder Raj, and Deepti Kakkar

Abstract Due to its sharp subthreshold swing and low leakage current, VDL-TFET has become a potential option for low-power electronic devices. In this study, we examine the influence of various device settings on the electrical properties of VDL-TFETs to assess their performance in low-power applications. We analysis how the doping level, layer thickness, gate length, and temperatures affect the functionality of VDL-TFETs. Simulations conducted by us demonstrate that by carefully tuning those parameters, the device's ON-current, subthreshold swing, and delay properties may be greatly enhanced, rendering it appropriate for low-power digital and analogue electronics. We additionally contrast the efficiency of VDL-TFETs against that of various other low-power transistors, including FinFETs, and show that the subthreshold swing and delay features of VDL-TFETs are comparable to those of FinFETs. According to our research, VDL-TFETs are an intriguing technology for low-power semiconductor purposes.

Keywords VDL · TFET · NDR · EF · VDL-TFET

M. Singh (✉) · N. S. Teja · T. Chaudhary · B. Raj · D. Kakkar
Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India
e-mail: Mandeep.moni.singh@gmail.com

N. S. Teja
e-mail: nakkinast.vl.21@nitj.ac.in

T. Chaudhary
e-mail: chaudharyt@nitj.ac.in

D. Kakkar
e-mail: kakkard@nitj.ac.in

1 Introduction

A particular kind of transistor that uses the quantum tunneling effect is called a vertical tunnel field-effect transistor (TFET). It has been suggested as a potential replacement for traditional Metal–Oxide–Semiconductor Field-Effect Transistors (MOSFETs) in forthcoming efficient and low-power digital systems. A p-i-n junction makes up the 3D structure of a vertical TFET. The source is represented by the “*p*” phase, the drain by the “*n*” phase, and the tunneling zone by the “*i*” phase. A thin layer of a semiconductor material such silicon or germanium that has been doped with contaminants to produce a small bandgap makes up the tunneling zone. Through the barrier layer, this tiny bandgap enables effective electron tunneling from start to finish. A positive voltage supplied to the gate causes an electric field that pulls electrons from the source area toward the tunneling region, which is how a vertical TFET operates [1]. The electrons that have sufficient energy to pass through the tunneling region’s potential barrier do so by tunneling through it and moving in the direction of the drain. This causes a current to pass through the source to the drain, as the gate voltage may regulate. Vertical TFETs are superior than ordinary MOSFETs in a number of ways. Their steeper subthreshold slope allows them to operate more quickly by enabling faster on and off ratio. Second, they use less power because they operate at lower voltages and have lower leakage current. Thirdly, because they are less susceptible to temperature changes, they can function more dependably throughout a broader range of temperatures [2]. Getting high on currents is one of the difficulties in creating vertical TFETs. Because of its narrow bandgap, effectively a tiny percentage of electrons can tunnel over the tunnelling region. A number of strategies have been put out to deal with this problem, including employing hetero junctions to boost the likelihood of tunneling, strained materials to narrow the bandgap, and gate engineering to improve tunneling. Getting excellent efficiency at ambient temperature is another difficulty. This is due to the potential barrier of the tunneling zone which may be bypassed by thermal energy at ambient temperature, which lowers the tunneling current. The use of low-dimensional materials with a higher chance of tunneling at room temperature, including carbon nanotubes or graphene, and hybrid TFETs that combine tunneling with conventional injection have all been suggested as solutions to this problem [3]. A Vertical Doped Layer Tunnel Field-Effect Transistor (VDL-TFET) is utilized for overcoming the limitations of VT-FET. An example of a transistor that makes use of quantum tunneling characteristics to accomplish excellent performance and low-power use is the VDL-TFET. The source and drain connections are formed by a vertical channel of substantially doped semiconductor material sandwiched within two mildly doped areas. The gate electrode, which surrounds the channel, regulates how much current flows across the device. Upcoming technologies look hopeful thanks to the VDL-TFET’s numerous benefits over traditional transistors [4, 5]. These benefits include reduced power usage, increased operating speed, and enhanced scalability. We will explain concerning the VDL-TFET’s composition and working ideas in this article as well as some of the prospective industries it may find use in.

2 Schematics of VDL-TFET

A transistor type known as a VDL-TFET uses the idea of quantum tunneling to operate. The device has three terminals: a source, a drain, as well as a gate. Current flow through the source to the drain is controlled by the gate. In contrast to various other kinds of transistors, the VDL-TFET has a vertically doped film in the channel region. In order to enhance on-state current and decrease off-state current, the Vertical Doped Layer should act as a tunneling barrier [6, 7] (Fig. 1).

Based on the quantum tunneling theory, the VDL-TFET permits charge carriers to go across the channel even though they lack the necessary energy to penetrate above the energy barrier that the lightly doped areas have built up. This happens as a result of the channel's intense doping, which narrows the energy bandgap and enables electrons to tunnel past the barrier [8]. Whenever the gate voltage is positive, an electric field is produced that drains the channel of electrons, resulting in the formation of a p-type phase. The p-type area grows as the voltage level rises, narrowing the depletion zone and increasing the number of electrons accessible for tunneling. As a consequence, the equipment experiences an increase in current flow n-type section which is produced whenever the gate voltage is negative because it produces an electric field that draws electrons to the channel. The n-type zone enlarges as the gate voltage drops, widening the depletion area and lowering the total amount of electrons accessible for tunneling. The device's flow of current is reduced as a result of this [9]. There are several possible uses for the VDL-TFET in numerous industries. Low-power electronics is one of the finest viable approaches. The VDL-TFET uses a lot less energy than traditional transistors since it relies on the quantum tunneling principle [10]. Because of this, it is perfect for use in portable electronics and other battery-powered products where energy economy is important.

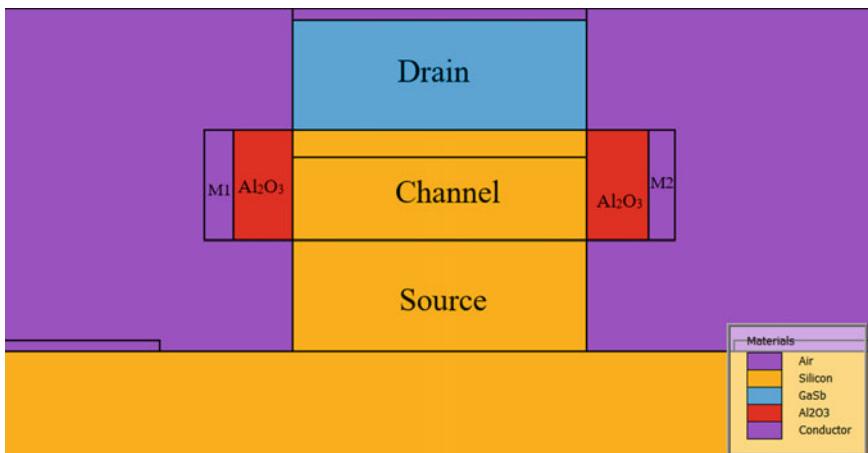


Fig. 1 Schematic of double-gate VDL-TFET

Table 1 Parameters of DG VDL-TFET

Parameters	Values
Gate length	20 nm
Drain current	6 nm
Source length	6 nm
Drain doping	5×10^{19}
Source doping	5×10^{18}
Gate doping	5×10^{18}
Work functions	4.3 eV

High-speed electronics is a further field in which the VDL-TFET may find use. The VDL-TFET can be utilized in fast speeds' communication networks as well as other purposes that call for quick switching due to its ability to flip on and off very rapidly. The field of quantum computing may benefit from the use of the VDL-TFET. The VDL-TFET is especially suited for usage in quantum computers, that utilize the laws of quantum mechanics to conduct calculations, as it functions on the principles of quantum tunneling [11] (Table 1).

3 Simulations and Result

The drain current generally declines with rising gate voltage in the negative differential resistance (NDR) area of the I_d - V_g graph of a VDL-TFET. The tunneling barrier's size is reduced by the gate voltage, that in turn lowers the likelihood that a tunnel will form and, as a result, the current. A crucial component of VDL-TFETs, the NDR region renders them desirable for low power and applications requiring high speeds. The NDR area, the peak current area, as well as the saturation region may be distinguished on the I_d - V_g chart of a VDL-TFET. The portion of the illustration known as the peak current region corresponds to where the drain current achieves its greatest value, and the region known as the saturation region is wherever the drain current keeps constant as the gate voltage increases. As the gate voltage increases in the NDR area, the drain current drops in order to reduce the likelihood of tunneling, for the reasons indicated above. The NDR zone can vary in breadth and depth based on the device specifications and is commonly seen at negative gate voltages [12]. The area of the illustration wherein the drain current obtains its greatest value is known as the peak current zone. According to the device's specifications, this area may vary in height as well as width and is often found at or close to zero gate voltage. In addition to rising gate voltage, the drain current remains constant in the normal state. The accessibility of the charge carriers in the channel places limits on this area, which is normally seen at positive gate voltages. In conclusion, since a VDL-TFET relies on quantum tunneling, its I_d - V_g graphs display an NDR zone. VDL-TFETs are desirable for low-power and high-speed uses due to their distinctive feature. The three parts

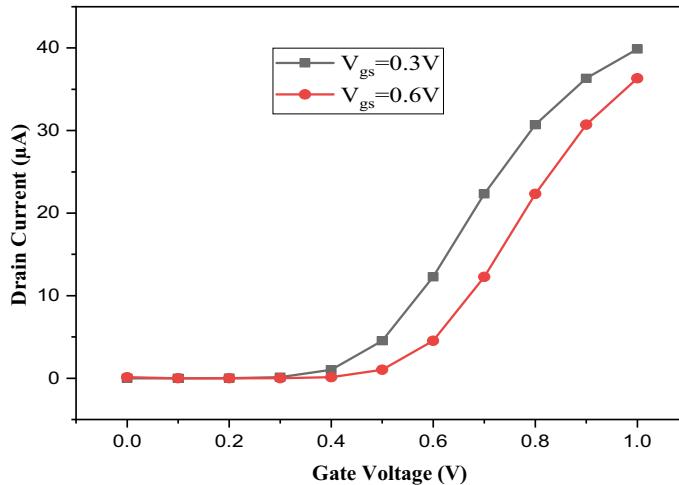


Fig. 2 Graph drain current versus gate voltage at different $V_{gs} = 0.3, 0.6\text{ V}$

of the graph—the NDR area, the peak current area, and the saturation zone—can be separated, and each is influenced by the device settings [13, 14] (Fig. 2).

A gate stack Vertical Tunnel Field-Effect Transistor (VT-FET) often has a multi-layered energy band diagrams. A gate electrode, a gate dielectric layer, and a tunneling barrier layer make up the gate stack. To potentially divide the two zones, the source and drain phases are doped. As depicted in Fig. 3, the energy band chart displays the device's energy levels, with the y-axis standing for energy and the x-axis for distance. During equilibrium, the gate electrode's Fermi level (EF) coincides with the valence band edge, whereas the source and drain areas' EF correlates with the conduction band margin [15]. The gate electrode gets increasingly positively charged when a positive gate bias is applied, forming an energy barrier that lowers the tunneling barrier and permits electrons to pass from the source area to the drain area. As a consequence, the I-V curve exhibits a NDR area that may be utilized for low-power switching purposes. By altering the doping level, layer thickness, and gate bias, a VDL-TFET's energy band shape may be changed. Greater subthreshold swing, on/off current ratio, and leakage current may be achieved by tuning these characteristics, which will also increase the efficiency of the device. Generally speaking, the energy band diagrams are a helpful tool for comprehending the functioning and operations of gate stack VT FETs.

Figure 4 illustrates how the electric field in the channel area may be modified when the gate bias voltage (V_{gs}) is raised, impacting the tunneling probability. The electric field might be strong and reasonably continuous across the channel area at low V_{gs} . A larger likelihood of tunneling might result from the electric field becoming more concentrated close to the tunneling location when V_{gs} is raised. As a result, as V_{gs} is raised, the electric field graph of a vertically doped layer TFET may exhibit a shift in the electric field peak in the direction of tunneling region. At larger V_{gs} levels,

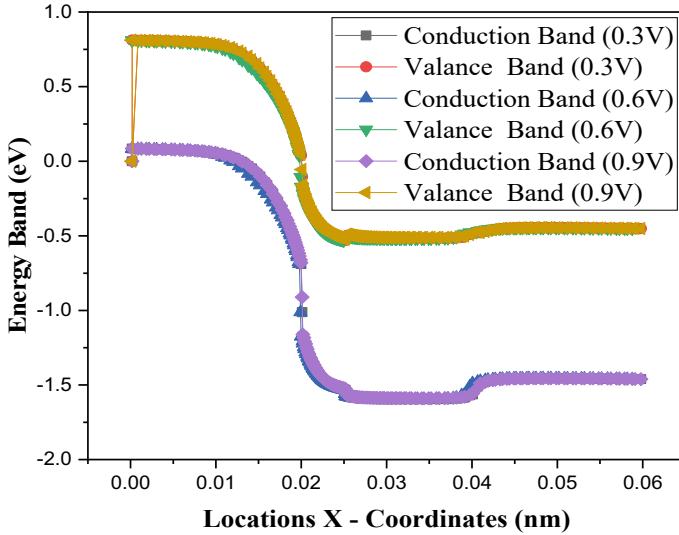


Fig. 3 Energy band diagram of VDL-TFET

the graph's shape will also alter, taking on a steeper slope close to the tunneling area [16].

A transconductance (g_m) plot for a VDL-TFET demonstrates the connection between the device's transconductance and V_{gs} . As seen in Fig. 5, the g_m graph of a VDL-TFET at $V_{gs} = 0.3$ V exhibits a lower transconductance value than at V_{gs}

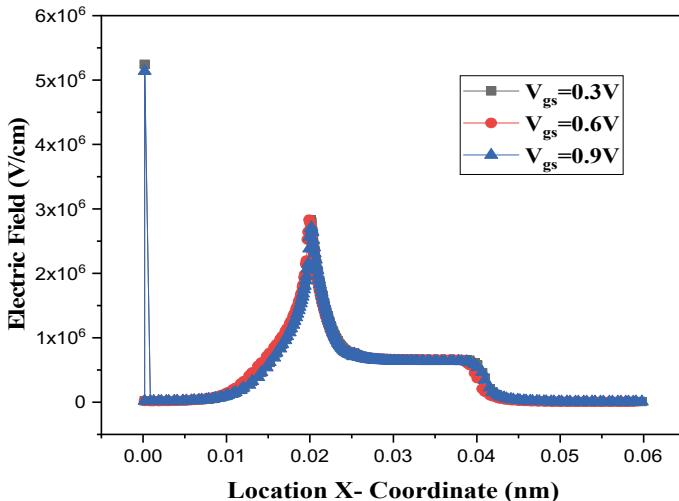


Fig. 4 Electric field graph of VDL-TFET at different $V_{gs} = 0.3, 0.6, 0.9$ V

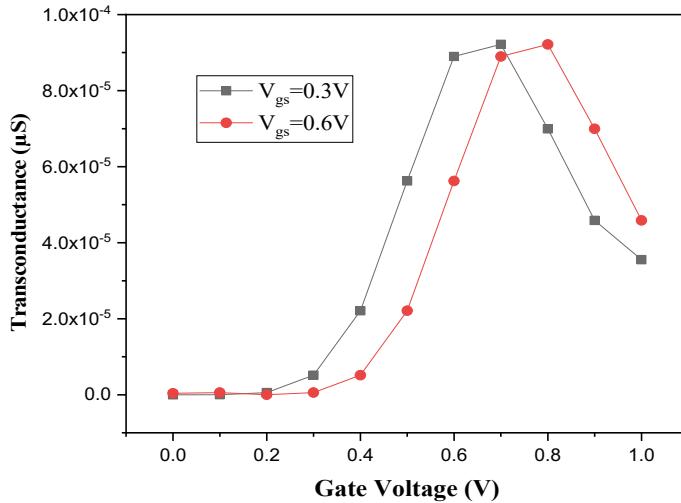


Fig. 5 Transconductance graph of VDL-TFET at different $V_{\text{gs}} = 0.3, 0.6\text{ V}$

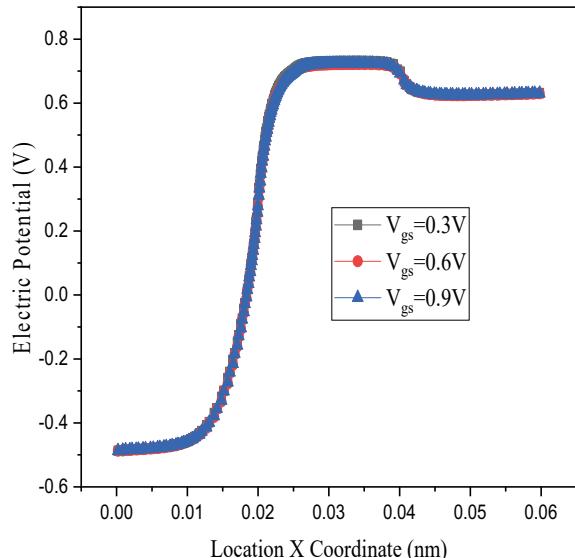
$= 0.6\text{ V}$. This is due to the fact that for $V_{\text{gs}} = 0.3\text{ V}$, the electric field on the channel's side is insufficient to remove obstacles at the source–channel interface. So, with a lower gm value, the gadget runs in the depletion mode [17]. On the contrary side, tunneling begins $V_{\text{gs}} = 0.6\text{ V}$ because the electric field is strong sufficient to pass through the potential barrier. As a consequence, the device runs in tunneling mode, which causes the gm value to drastically rise. The steeper slope of the gm plot indicates that the reason for this rise in gm is the greater tunneling current brought on by the higher gate voltage. In general, the transconductance plot of a VDL-TFET at various V_{gs} values demonstrates the device's capacity to control the flow of current in the active zone by adjusting the gate voltage. The greater the gate voltage and transconductance, the more efficiently the device can amplify an input signal [18, 19].

As seen in Fig. 6, the electric potential chart of a VDL-TFET at various V_{gs} demonstrates that the potential energy of the electrons varies as they pass across the transistor's channel. The acceptor concentration in the channel and the gate voltage both influence the electric potential pattern of a VDL-TFET, which in turn influences the channel's carrier concentration and tunneling likelihood [20]. The electric potential plot exhibits a comparatively flat pattern with a modest potential barrier close to the source and drain connections at a $V_{\text{gs}} = 0.3\text{V}$. This is due to the low likelihood of tunneling and insufficient gate voltage to create a large current flow across the channel. As consequently, the source and drain contacts' potential barriers have the greatest impact on how the current moves through the transistor. The electric potential plot depicts a stronger potential barrier near the source and a lower barrier at the drain as the V_{gs} rises to 0.6 V . This is due to a stronger gate voltage that causes a significant current to flow across the channel and a higher

possibility of tunneling [21]. As an outcome, any barrier that may be present close to the source gets more obvious, increasing the channel's carrier density and current flow. The electric potential curve at a V_{gs} of 0.9 V reveals a significant potential barrier close to the source and a significantly smaller barrier close to the drain. This is due to the fact that the gate voltage is currently in the area of saturation, where the greatest amount of current is flowing through the channel. As a result, the transistor operates like a conventional MOSFET with a high carrier density in the channel in this area, when the potential barrier close to the source becomes dominant. Additionally, the acceptor concentration and gate voltage have a significant impact on the electric potential graph of a VDL-TFET, which is a key component in defining its performance aspects. For the purpose of building and maximizing VDL-TFET devices for diverse applications, it is crucial to comprehend these interactions [22].

Figure 7 will display conventional behavior for an n-channel VDL-TFET while the work function is 4 eV. The energy barrier across the source as well as channel gets lower when the gate's voltage rises. Current will flow across the drain to the source when the barrier is sufficiently low for electron to tunnel over (V_{th}) at a specific voltage. As a result, the drain current quickly rises, resulting in an illustration with a high slope over the threshold voltage. The value of the threshold voltage will decline when the work function is significantly decreased to 4.3 eV [23]. Since a consequence, relative to the prior scenario, the drain current begins to rise at an inferior gate voltage. The chart general shape does not change, but the initial commencement of the current flow is caused by an alteration in the threshold voltage. In contrast, the threshold voltage rises whenever the work function is raised to 4.6 eV. Because of the larger energy barrier across the source–channel junction, a greater gate voltage is necessary [24]. The chart shifts to the opposite direction as an outcome of the drain

Fig. 6 Electric potential graph of VDL-TFET at different $V_{gs} = 0.3, 0.6, 0.9$ V



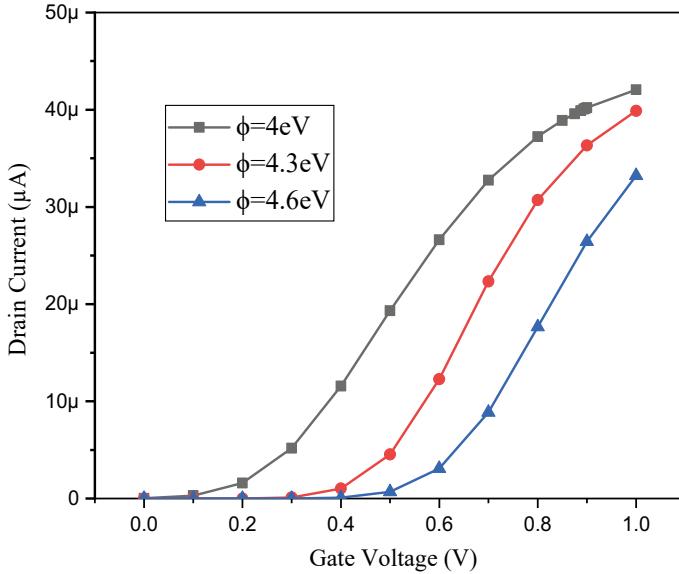


Fig. 7 Drain current versus gate voltage curve for different work function variations

current beginning to rise above a greater gate voltage. Although the entire behavior will appear prolonged in comparison to the 4 eV work function scenario, the slope above the threshold voltage will continue to be steep [25, 26].

The transconductance curve in Fig. 8 will operate like an n-channel VDL-TFET in a usual manner whenever the work function is 4 eV. The energy barrier across the source–channel junctions gradually gets lower when the gate-source voltage rises. This decrease in barrier height makes it possible for electrons to tunnel across, increasing the drain current [27]. As a result, the derivative of the drain current compared to gate-source voltage—the transconductance—increases. The transconductance begins to climb sharply on the chart, demonstrating the highly sensitive nature of the drain current to variations in the gate-source voltage. The threshold voltage falls when the work function is significantly decreased to 4.3 eV. Since a consequence, the transconductance chart, although having a smaller threshold voltage, will show an analogous pattern to the 4 eV instance. Contrasting to the 4 eV work functioning, the value of transconductance will grow as smaller gate-source voltages, causing the chart to climb more quickly and abruptly. In contrast, the threshold voltage rises since the work function is raised to 4.6 eV. Because of the larger energy barrier within the source–channel junction, a greater gate-source voltage is necessary. As a result, a gradual rise for transconductance will be shown within the transconductance chart. In contrast to the 4 eV work function scenario, the behavior will generally swing toward greater gate-source voltages and although the slope of the chart above the threshold voltage continues to be steep [28].

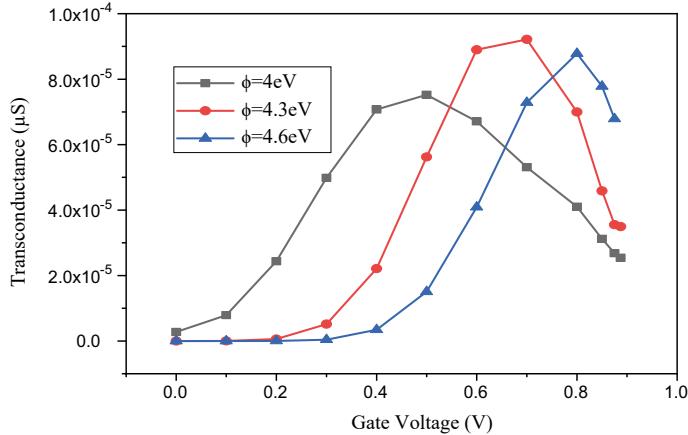


Fig. 8 Transconductance curve for different work function variations

The conduction current concentration plot (Fig. 9) will display conventional behavior for an n-channel VT-FET whenever the work function is 4 eV. The energy barrier at the source–channel junction inhibits considerable electron flow at lower electric fields; hence, the conduction current density is insignificant. The energy barrier gets lower as the electric field is stronger, making it possible for electrons to tunnel across it and contributing to the conduction current. As a result, since the electric field increases, the conduction current concentration will show a significantly increase. The conduction current density will initially display a steep slope on the chart, suggesting a strong sensitivity to variations in the electric field. Lowering the work function to 4.3 eV results in a smaller threshold voltage and a smaller energy barrier. Due to the smaller threshold electric field, the conduction current density chart will display an identical pattern to the 4 eV instance. Comparing to the 4 eV work function, the conduction current density will begin to grow at a lower electric field, causing the graph to rise more quickly and abruptly [29]. In contrast, the threshold voltage rises and a stronger electric field is needed to pass a higher energy barrier whenever the work function rises to 4.6 eV. As a result, a gradual rise within the conduction current density will be shown on the conduction current density chart. Above the threshold electric field, the slope of the chart remains sharp, but contrasted to the 4 eV work function instance; the general behavior will be changed to higher electric fields [30]. Figure 10 shifts to the right because larger work functions needed a stronger electric field to reach a substantial conduction current density. By the other hand, smaller work functions result in a smaller threshold electric field, which prompts a quicker and sharper increase in the plot of conduction current density.

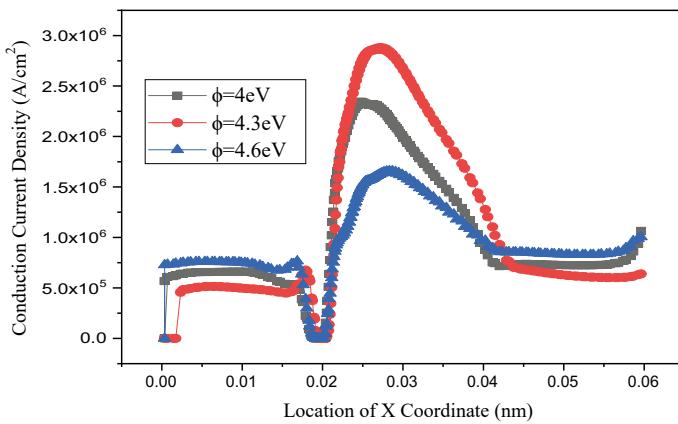


Fig. 9 Conduction current density graph for different work function variations

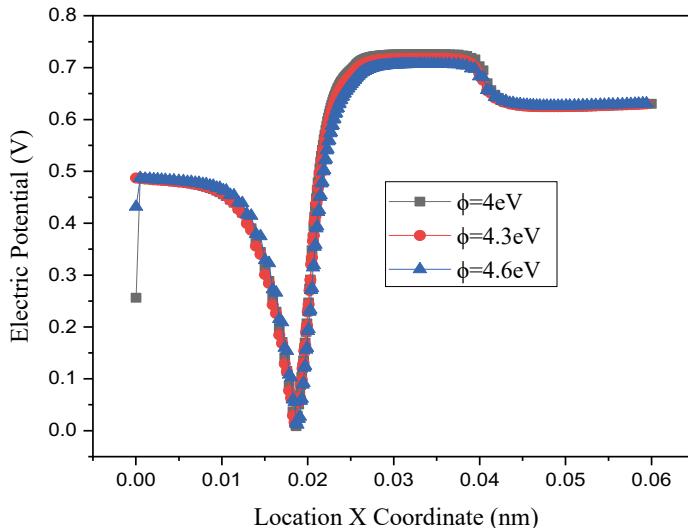


Fig. 10 Electric potential graph for different work function variations

Parameters	Units	Values
I_{ON}	μA	3.9×10^{-5}
I_{OFF}	μA	2.9×10^{-17}
I_{ON}/I_{OFF}	-	1.3
Transconductance	μS	9.8
Subthreshold	mV/dec	20.8

4 Conclusion

This study compares the performance of VDL-TFETs to other low-power transistors and finds that they can deliver competitive results in terms of transconductance and subthreshold swing. In low-power applications, this suggests that VDL-TFETs may eventually take the place of conventional transistors. Further study is necessary to improve the performance of VDL-TFETs for real-world applications because scaling them up remains a difficult problem. Furthermore, our research offers important information about the performance and enhancement of VDL-TFETs for low-power applications, which can direct future studies in this area. According to our analysis, VDL-TFETs might soon make it possible to create high-performance, low-power electronic devices and circuits. According to our research, VDL-TFETs are an innovative technology that has promise for low-power electronic circuits.

References

1. Gupta SK (2019) Analytical modeling of a triple material double gate TFET with heterodielectric gate stack. *SILICON* 11(3):1355–1369
2. Ko E, Lee H, Park JD, Shin C (2016) Vertical tunnel FET: design optimization with triple metal-gate layers. *IEEE Trans Electron Devices* 63(12):5030–5035
3. Goswami RN, Poorvasha S, Lakshmi B (2018) Tunable work function in junctionless tunnel FETs for performance enhancement. *Aust J Electr Electron Eng* 15(3):80–85. <https://doi.org/10.1080/1448837X.2018.1525173>
4. Gautam R, Saxena M, Gupta RS, Gupta M (2013) Gate-all-around nanowire MOSFET with catalytic metal gate for gas sensing applications. *IEEE Trans Nanotechnol* 12(6):939–944. <https://doi.org/10.1109/TNANO.2013.2276394>
5. Lee B, Oh J, Tseng H, Jammy R, Huf H (2006) Gate stack technology for nanoscale devices. *Mater Today* 9:32–40. [https://doi.org/10.1016/S1369-7021\(06\)71541-3](https://doi.org/10.1016/S1369-7021(06)71541-3)
6. Verhulst AS, Vandenberghe WG, Maex K, De Gendt S, Heyns MM, Groeseneken G (2008) Complementary silicon-based heterostructure tunnel-FETs with high tunnel rates. *IEEE Electron Devices Lett* 29(12):1398–1401
7. Avci UE, Morris DH, Young IA (2015) Tunnel field-effect transistors: Prospects and challenges. *IEEE J Electron Devices Soc* 3(3):88–95
8. Raad B, Nigam K, Sharma D, Kondekar P (2016) Dielectric and work function engineered TFET for ambipolar suppression and RF performance enhancement. *Electron Lett* 52(9):770–772
9. Singh S, Raj B (2018) Vertical tunnel-fet analysis for excessive low power digital applications. In: 2018 First International conference on secure cyber computing and communication (ICSCCC). IEEE, pp 192–197
10. Begam SS, Selvachandran G, Ngan TT, Sharma R (2020) Similarity measure of lattice ordered multi-fuzzy soft sets based on set theoretic approach and its application in decision making. *Mathematics* 8:1255
11. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Tien BD, Priyadarshini I, Sarkar M, Le T (2020) Crime rate detection using social media of different crime locations and Twitter part-of-speech tagger with brown clustering. *J Intell Fuzzy Syst* 4287–4299
12. Nguyen PT, Ha DH, Avand M, Jaafari A, Nguyen HD, Al-Ansari N, Van Phong T, Sharma R, Kumar R, Le HV, Ho LS, Prakash I, Pham BT (2020) Soft computing ensemble models based on logistic regression for groundwater potential mapping. *Appl Sci* 10:2469

13. Jha S et al (2019) Deep learning approach for software maintainability metrics prediction. *IEEE Access* 7:61840–61855
14. Sharma R, Kumar R, Sharma DK, Son LH, Priyadarshini I, Pham BT, Bui DT, Rai S (2019) Inferring air pollution from air quality index by different geographical areas: case study in India. *Air Qual Atmos Health* 12:1347–1357
15. Sharma R, Kumar R, Singh PK, Raboaca MS, Felseghi R-A (2020) A systematic study on the analysis of the emission of CO, CO₂ and HC for four-wheelers and its impact on the sustainable ecosystem. *Sustainability* 12:6707
16. Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I, Le D-N (2020) Global forecasting confirmed and fatal cases of COVID-19 outbreak using autoregressive integrated moving average model. *Front Public Health* 8:580327. <https://doi.org/10.3389/fpubh.2020.580327>
17. Malik PK, Sharma R, Singh R, Gehlot A, Satapathy SC, Alnumay WS, Pelusi D, Ghosh U, Nayak J (2021) Industrial internet of things and its applications in industry 4.0: state of the art. *Comput Commun* 166:125–139. ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2020.11.016>
18. Sharma R, Kumar R, Satapathy SC, Al-Ansari N, Singh KK, Mahapatra RP, Agarwal AK, Le HV, Pham BT (2020) Analysis of water pollution using different physicochemical parameters: a study of Yamuna River. *Front Environ Sci* 8:581591. <https://doi.org/10.3389/fenvs.2020.581591>
19. Dansana D, Kumar R, Parida A, Sharma R, Adhikari JD et al (2021) Using susceptible-exposed-infectious-recovered model to forecast coronavirus outbreak. *Comput Mater Continua* 67(2):1595–1612
20. Vo MT, Vo AH, Nguyen T, Sharma R, Le T (2021) Dealing with the class imbalance problem in the detection of fake job descriptions. *Comput Mater Continua* 68(1):521–535
21. Sachan S, Sharma R, Sehgal A (2021) Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. *Sustain Comput: Inf Syst* 30:100504
22. Ghanem S, Kanungo P, Panda G et al (2021) Lane detection under artificial colored light in tunnels and on highways: an IoT-based framework for smart city infrastructure. *Complex Intell Syst.* <https://doi.org/10.1007/s40747-021-00381-2>
23. Sachan S, Sharma R, Sehgal A (2021) SINR based energy optimization schemes for 5G vehicular sensor networks. *Wireless Pers Commun.* <https://doi.org/10.1007/s11277-021-08561-6>
24. Memisevic E, Svensson J, Lind E, Wernersson LE (2017) InAs/InGaAsSb/GaSb nanowire tunnel field-effect transistors. *IEEE Trans Electron Devices* 64:4746–4751
25. Moselund KE, Schmid H, Bessire C, Bjork MT, Ghoneim H, Riel H (2012) InAs–Si nanowire heterojunction tunnel FETs. *IEEE Electron Device Lett* 33:1453–1455
26. Nigam K, Kondekar P, Sharma D (2016) High frequency performance of dual metal gate vertical tunnel field effect transistor based on work function engineering. *Micro Nano Lett* 11(6):319–322
27. Sterkel M, Wang PF, Nirschl T, Fabel B, Bhuwalka KK, Schulze J, Eisele I, Schmitt-Landsiedel D, Hansch W (2005) Characteristics and optimisation of vertical and planar tunnelling-FETs. *J Phys: Conf Ser* 10(1):15
28. Saurabh S, Kumar MJ (2009) Impact of strain on drain current and threshold voltage of nanoscale double gate tunnel field effect transistor (TFET): theoretical investigation and analysis. *Jpn J Appl Phys* 48:064503
29. Verhulst AS, Vandenberghhe WG, Maex K, De Gendt S, Heyns MM, Groeseneken G (2008) Complementary silicon-based hetero-structure tunnel-FETs with high tunnel rates. *IEEE Electron Device Lett* 29(12):1398–1401
30. Wang PF, Hilsenbeck K, Nirschl T, Oswald M, Stepper C, Weis M, Schmitt-Landsiedel D, Hansch W (2004) Complementary tunneling transistor for low power application. *Solid State Electron* 48(12):2281–2286

OpenFace Tracker and GoogleNet: To Track and Detect Emotional States for People with Asperger Syndrome



Mays Ali Shaker and Amina Atiya Dawood

Abstract The goal of this work is to create an emotional model that can categorize emotions in real-time for persons with Asperger's syndrome (AS). The model is based on facial expressions, head movement, and eye gaze as significant features for emotions. Developing like this an emotional model can aid other people to communicate socially with people with autism. This model overcomes the limitations of previous research that used sensitive and invasive methods to capture physiological data in order to predict emotional states. These instruments are very costly and need a controlled environment. The proposed model succeeded in classifying the emotional states of the AS using a natural spontaneous dataset without invasive tools and in an uncontrolled environment. The dataset used in this study is an available dataset and contains videos recorder in an uncontrolled environment with different facial occlusion and illumination changes. The emotions were classified as fear, disgust, joy, anticipation, and sadness. The proposed model implements a modified version of the well-known and wide-used GoogleNet machine learning model to classify emotions. The key feature of GoogleNet is the inception module, which is designed to capture different levels of spatial features and perform dimensionality reduction using parallel convolutional layers with different filter sizes. There are several metrics that were used to measure the accuracy and generality of a DL model, including accuracy, precision, recall, and F1 score, depending on the nature of the task. Overall, measuring the generality of a DL model is an important step in evaluating its performance and ensuring that it will perform well on new, unseen data. The model achieved significant performance on unseen data with accuracy (98%).

Keywords Asperger syndrome · Deep learning · Emotional model · GoogleNet · Facial expressions · Head motion · Eye gaze

M. A. Shaker (✉) · A. A. Dawood

Computer Science Department, College of Science for Women, University of Babylon, Babylon, Iraq

e-mail: mays.hussein.gsci113@student.uobabylon.edu.iq

1 Introduction

Autism spectrum disorder (ASD) is a group of neurodevelopmental disorders [1] characterized by difficulties with communication, social interaction, and repetitive behaviors [2]. Asperger syndrome, autism, and non-specified pervasive developmental disorder (PDD-NOS) are the three different categories of autism spectrum disorders. Nevertheless, these categories are presently referred to as (ASD) [3]. Individuals with AS have a normal or above-normal IQ and normal language development [4], as we will be focused on this group in this study. There are currently no known interventions or treatments that can cure or fully alleviate (ASD) [5]. However, comprehensive behavioral and educational intervention programs play a major role in helping persons with ASD to overcome impairments in communication and social interaction [6].

To communicate socially, we need to understand theory of mind (ToM) [7] that means, we need ability to interpret mental states. The recognition of facial emotions is considered a fundamental component of social communication because it allows people to interpret and respond appropriately to the emotional cues of others [8]. This can help facilitate social interactions and improve social relationships [9]. Facial expressions play a crucial role in conveying significant non-verbal information, eye gaze and head movement also play an important role [10, 11].

There are many approaches of methods that have been developed for emotional state recognition, by capturing facial expressions and other non-verbal cues [12, 13].

However, despite the growing interest in enhancing the capacities of individuals with autism via study in interactive intervention, there is a shortage of literature on identifying the emotions shown by people with autism in the real world [14].

This research aims to explore the integration of OpenFace Tracker and GoogleNet, two powerful computer vision techniques, to track and detect emotional states in people with Asperger syndrome. OpenFace Tracker is a state-of-the-art facial tracking system that analyzes facial expressions, eye gaze, and head movements [15], while GoogleNet is a deep convolutional neural network, which proved the quality of its performance in classification tasks and possesses the ability to learn from imagery (features-map) [16].

By combining these technologies, we can create a robust framework capable of detecting and tracking emotional states with higher accuracy in the real world. The model will be used as assistance tool in social communication and help people to understand emotions which are expressed by people with autism. The proposed architecture is to monitor and predict emotions of people with autism.

Next section in this study will present the related works in the field of ER. Also, the methodology which is used in this study will be displayed. The final section will present the important results and conclusion.

2 Related Works

A lot of studies and tools have been developed to recognize emotions for people with autism. For instance a study [3], the researchers developed an effective model using CNN, LSTM algorithms to infer cognitive-affective states of students with Asperger's syndrome in real-time. A natural spontaneous affective dataset collected from head movement, facial expressions, and eye gaze was used while students interacted with a computer in an uncontrolled environment. The model successfully achieved an accuracy rate of 90.06% in detecting five emotional affective states: confidence, uncertainty, engagement, anxiety, and boredom.

The eye movements of 21 autistic individuals and 23 normally developing people with comparable ages, sexes, and IQs were recorded in [17] to investigate the significance of visual information received from facial stimuli for emotion identification. Autistic people concentrated more on the lower facial areas and less on the eyes on the emotion detection test than the usually developing group. Thus, the initial step of face processing, namely the extraction of visual information via eye-fixations, may be ascribed to at least a part of the autistic deficiency in emotion recognition.

In [18], the researchers developed an Intelligent Tutoring System (ITS) aimed at helping people with Asperger's syndrome, also known as Asperger's, to capture non-verbal signals, like facial expressions. The researchers extracted images for facial expression and the features were analyzed.

In [13], the researchers suggested combining machine learning and physiological signals to determine a student's emotions during a test. They utilized wearable sensors to gather EEG, ECG, and EMG data from 27 individuals taking an English language test on a computer. Despite yielding fruitful results, using physiological signal sensors may not be practical for ITS users outside of a laboratory setting due to its invasiveness and inconvenience.

A classifier was developed in Hassouneh et al. [19] to recognize fundamental emotions using a combination of facial expression feature points and electroencephalograph (EEG) data. To complete the classification job on a set of (35 men and 25 girls), the researchers used CNN and LSTM algorithms. The Haar method was used to extract the feature points from the face. The best accuracy for recognizing emotions was 87.25%. However, the research was hampered by the scarcity of themes on which to collect data, as well as the demand for additional EEG signal qualities.

In [20], the social issues associated with ASD may be partially attributable to neurotypical-autistic distinctions (i.e., variations in facial expressions); thus, the study centered on these facial expression disparities. In general, autistic and neurotypical individuals express themselves differently, with autistic individuals demonstrating less frequent and lower-quality expressions, as evaluated by neurotypical participants. In the case of facial expression, an autistic individual experiencing sorrow may assume a different facial expression than the traditional downturned mouth. As a consequence, we may be unable to perceive the individual's sorrow.

Face recognition, feature classification, and face feature extraction were the three steps that went into building the real-time emotion detection system for autistic children that was created as part of this study [21]. Only seven unique facial emotions may be identified by the proposed system: anger, disgust, fear, joy, sorrow, ridicule, and amazement. When evaluated on an average sample from 6 to 14-year-old youngsters, the suggested application had good results.

In [12], the researchers employed deep CNN to categorize facial expressions observed by individuals with and without autism spectrum disorder, using electroencephalography data that was recorded at the same time.

In [22], 37 people diagnosed with autism spectrum disorders and 43 people without the disorder were assessed on their capacity to copy and identify facial expressions. Researchers used a novel computerized face analysis tool to assess how successfully participants mimicked emotional facial expressions they were directed to emulate, as well as how proficient they were at distinguishing emotions. Although the autistic participants could imitate the facial expressions as directed, their imitation was slower and less accurate than that of the neurotypical ones.

Researchers at [23] created a system capable of automatically understanding emotions based on facial expressions and combined it with a robotic platform to help youngsters with autism spectrum disorders connect socially. The experimental setup and approach for a real-time face emotion recognition system included an Intel® RealSense™ 3D sensor, facial feature extraction, and a multiclass support vector machine classifier. The findings showed that the suggested strategy is suitable for use in therapy sessions and may be used to improve emotion recognition and imitation abilities. This work's limitations are as follows: the low identification rates of 'Anger' and 'Fear' in the activity IMITATE might be attributed to the fact that children had to assess the facial expression shown by ZECA, which could indicate that they did not interpret the facial expression properly [28, 29].

3 Methodology

The model structure can be seen in Fig. 1. The model consists of several steps that start with data collection and preparing steps and end with classification and evaluation steps.

The dataset used in this work [24] structured in two folders one for typical development students (TD) and another for autistic students (ASD). Each folder contains videos taken to students during uncontrolled environment engaging in a learning process. The videos were taken using computer digital camera.

OpenFace API [15] was used to track and process the videos using a software tool (i.e.,) that generates files of features tracking (in.CSV file format) for the eye gaze, head-pose, and set of face action units (AUs). The tracking features contains values about eyes pupil position and gaze, head position relative to the screen and orientation, in addition to face action units extracted from learned models.

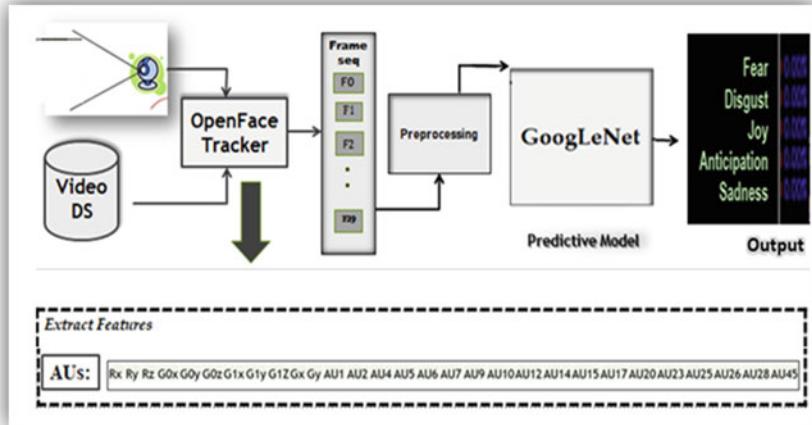


Fig. 1 User interface for the proposed model

3.1 Data Preprocessing

This step prepares and processes the data for using by the model. First step is data cleaning, as it is an important step in any mathematical model. It includes removing duplicates, eliminating duplicate entries guarantees the uniqueness of each data point and prevents skewed analysis. Evaluate missing data, outliers excluding them from analysis or include them based on their relevance, and ensure consistency. By accomplishing these steps ensure the dataset becomes more reliable and confident. The dataset which is used in this paper was cleaned previously by its authors.

3.2 Cues Generation

Building the cues generation is the next step in this model. Cues generation is the important contribution in this work; as based on my knowledge this is a new method used to build cues vector to be used in emotion prediction. The work in [3] accumulates cues with heterogeneous features (gaze, pose, and Au's) in one map. On the other hand, this paper categorizes homogenous features into separate maps, as shown in Fig. 2a and b. Figure 2a for researches in [3], and Fig. 2b for this work, where g_n , p_n , and AU_n is a gaze, pose, and action unit, respectively; $n = 0$ to number of features per frame – 1; and t is the time spent on each frame.

These cues are input channel for the proposed work in this paper. The features are stacked into three matrix or categories one for gaze, pose, and another for AU's. Each matrix is organized in a sliding window where consecutive rows are separated

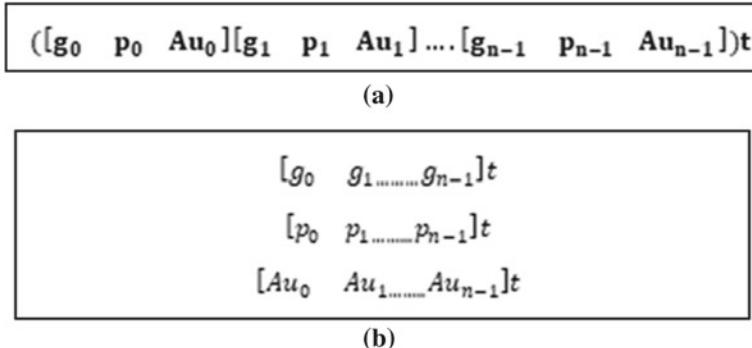


Fig. 2 **a** Cues generation for Dawood et al. [3]. **b** Cues generation for proposed work

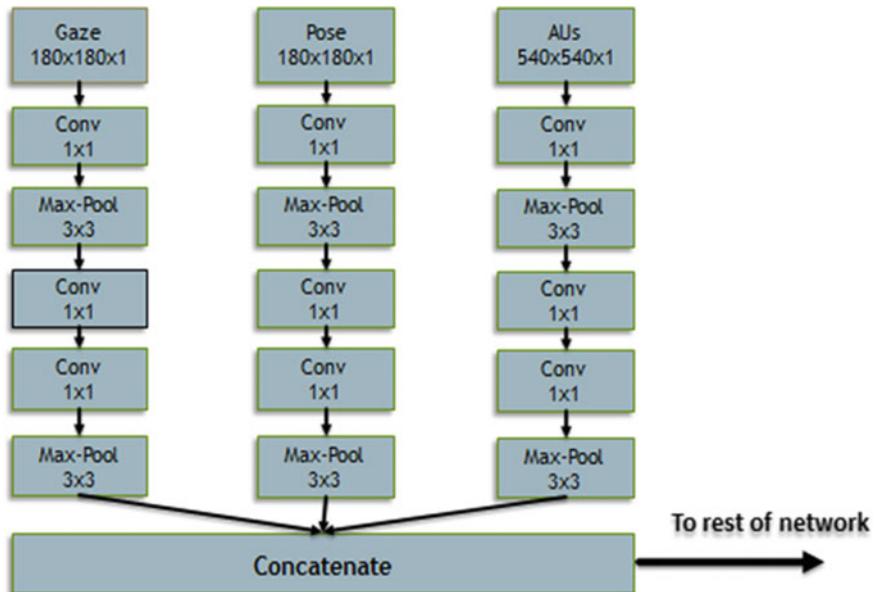
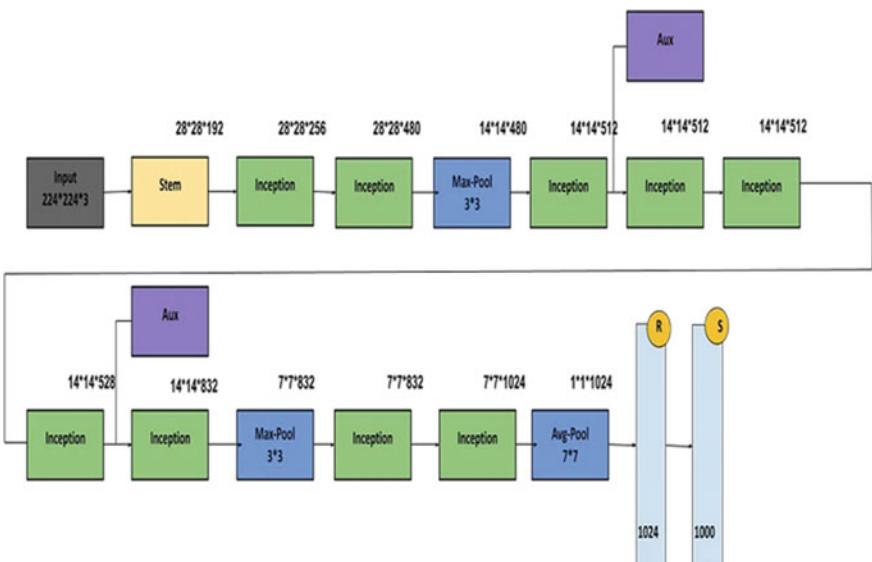
by a time difference of 1/60 of a second. As a result of this process, three files are generated, each representing the cue at the same specific point in time.

The model proposed in this study utilizes a customized adaptation of the popular and widely used GoogleNet machine learning model for emotion classification. GoogleNet, commonly known as Inception v1, is a deep CNN architecture developed initially by Google researchers for picture categorization. In the original model, the input layer except images of multiple channels is followed by an inception layer. The proposed modified model changes this single input model into three input layers. Each branch of the proposed architecture contains a stem, two inception layers, and one max-pooling layer before concatenating the three main branches; (Fig. 3) illustrates the flow of the modified GoogleNet. The purpose of these layers is to get initial knowledge from the extracted cues, when the network reaches the layer before concatenation the values now are insights gathered by the model rather than values that represent the original features (i.e., abstractions). After this step the structure continues as the original network.

GoogleNet works by using a deep CNN architecture to classify images into different categories. The network consists of multiple layers of convolutional filters, pooling layers, and fully connected layers that learn increasingly complex features of the input image.

One of the key features of GoogleNet is the inception module, which is designed to capture different levels of spatial features and perform dimensionality reduction using parallel convolutional layers with different filter sizes. The inception module allows the network to balance computational complexity and accuracy by avoiding the expensive operation of processing all the data through a single convolutional layer. Figure 4 states GoogleNet architecture.

The input image is first preprocessed with standard data augmentation techniques, such as random cropping and horizontal flipping. Then, the image is then sent through a series of convolutional layers, each followed by a rectified linear unit (ReLU) activation function and a max-pooling layer. The final pooling layer's output is fed

**Fig. 3** Modified GoogleNet**Fig. 4** GoogleNet

into many inception modules, which are coupled to a global average pooling layer and a softmax layer for classification.

During training, the network uses stochastic gradient descent (SGD) with momentum as the optimization algorithm and cross-entropy loss as the objective function. The network was trained using the ImageNet dataset, which contains 1.2 million labeled pictures from 1000 different categories.

In summary, GoogleNet works by using a deep CNN architecture with an inception module to capture different levels of spatial features and perform dimensionality reduction, and is trained on a large dataset of labeled images using SGD with momentum and cross-entropy loss.

The functions that used to generate the proposed model are illustrated in the algorithms below.

Algorithm 1: EmotioNet Function

1. Create input layers for gaze features, pose features, and AUs features.
2. Create a gaze Store the output of these layers as GX.
3. Create a pose Store the output of these layers as PX.
4. Create an AUs subnet Store the output of these layers as AX.
5. Concatenate the outputs of GX, PX, and AX, and store the result as X.
6. Apply two Inception blocks to X.
7. Apply a MaxPooling2D layer to X.
8. Apply additional five inception blocks.
9. Apply a 2D max-pooling layer with pool size of (3,3) and strides = 2.
10. Apply further two inception blocks to have a total of nine inception blocks.
11. Apply a global average pooling layer for the resulted model network.
12. Enforce a weights dropout layer with dropout percentage of 40%.

Generate the final classification dense layer with five classes (emotions classes) using SoftMax as an activation function.

3.3 Training Step

Once all the preceding steps and processes have been completed, the model enters the training phase. In machine learning, dataset splitting involves dividing a dataset into subsets to train and evaluate a model. The most common type is the train-test split, which creates a training set and a test set. The training set is used for model training, while the test set assesses its performance.

The dataset splitting ratio depends on the dataset size and task. Typically, an 80/20 or 70/30 ratio is used, with 80 or 70% for training and the remainder for testing. In this study, a three-way split is applied: 70% for training, 15% for validation (fine-tuning and preventing overfitting), and 15% for testing (evaluating model generality).

During training, data flows in batches through the model's network structure. The model's output is validated using the validation set to measure intermediate metrics.

These metrics are compared with previously collected ones to assess the model's progress and detect overfitting.

The training-validation process, from data-batch input to model validation, is called an epoch. After completing multiple epochs, the model is considered trained but not final.

4 Results and Discussions

Capturing features and producing likelihood probability of each emotional class along the time-series (t-secs) of the video-clip frames, was the model's task. Figure 1 states the user interface which display the emotional state for people behavior, as maximum probability was considered as the current emotion.

Number of videos which were used in this study is 862 video clips. The author generated into 684,364 cues as mentioned previously. Then the cues divided into training, validation, and test sets, as in Table 1.

Typically, the accuracy, error rate, generality, and detection speed are key indicators used to evaluate the performance and capabilities of a model. These metrics provide insights into the model's effectiveness, precision, versatility, and efficiency.

The model gained best results after 100 epochs of training steps. The training and validation accuracy and loss error graphs are shown in (Fig. 5a and b, respectively. From (Fig. 6) can be noticed that the model's performance improved with an increase in the number of epochs.

The accuracy metric for training was computed to evaluate the overall number of correct predictions made by the model, encompassing all five classes. Formula (1) states the process of calculating the model accuracy (ACC) [25], Table 2 states the value of ACC and error rate.

$$\text{ACC} = \frac{\text{TP}}{n} \quad (1)$$

where TP represents the count of accurate predictions made by the model (true predictions), while n represents the total number of testing cues. This measure is regarded as an optimistic metric for assessing the accuracy of categorical models.

The model exhibits a high level of performance across all classes, demonstrating accurate predictions for both true (correctly identifying the correct class) and false (correctly identifying when a class is not required) classifications. To gain a better

Table 1 Training, validation, and test

Split	Number of cues
Training	554,334
Validation	61,593
Test	68,437



Fig. 5 **a** Specified the accuracy of training and validation. **b** Described the loose training and validation mistake

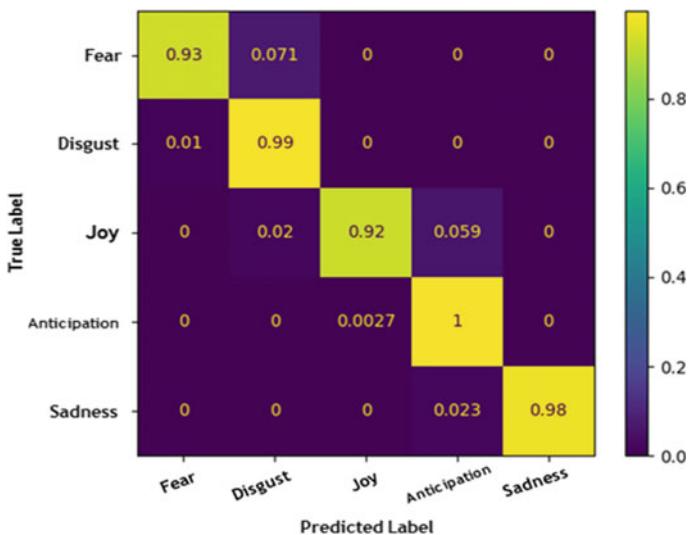


Fig. 6 Confusion matrix for predicted label and true label

Table 2 Accuracy and loss values for model

	Accuracy (%)	Loss
Train	99.88	0.00914
Test	98.54	0.016631
Validation	98.45/	0.018632

understanding of the model's performance for inter-class prediction, it was tested on an unseen dataset. This test dataset was distinct and separate from the training dataset, as it was excluded prior to the model's training phase and kept aside specifically for

evaluation purposes. The significant value of unseen data is to measure the model efficiency by predicting a correct class. The statistical results of the model's test experiments are summarized in a confusion matrix (Fig. 6), it states the correct and incorrect prediction rates for each emotional state.

The rows, in confusion matrices, describe the model's class prediction, while the columns are the true classes obtained from the labeled dataset. The highest classification rate was (100%) for anticipation class, while the lowest score was (0.93%) for fear class. These results state the model achieved solid performance and significant generality toward unseen data. Additional important metrics can be utilized to assess the model's generality. These metrics encompass calculating the true positive rate (TPR) for each class, also known as recall, which measures the number of true correct predictions that align with the ground truth label data. Additionally, the false positive rate (FPR) is employed to identify falsely classified classes. Another metric employed is precision or positive prediction value (PPV), which evaluates the accuracy of positive predictions made by the model.

The mathematical formula of these metrics is illustrated below, Eqs. (2) and (3) [26, 27].

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

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

where FN is the number of inaccurate class forecasts that differ from the ground truth class in reality. This is the quantity of erroneous forecasts that match the ground truth class for a given class. The term false positives (FP) describe how often wrong predictions for a given class match the ground truth class.

5 Conclusion

In this study, the captured normalized data of head movement, facial expression, and eye gaze were used as inputs to a prediction model to infer basic emotions for individuals with Asperger's syndrome by observing momentary dynamic changes in their facial expressions while interacting with a computer, where an emotional model was developed based on the model of deep learning (GoogleNet) to extract the temporal and spatial features; the signals are read from the captured video clips by a webcam installed on the computer without the need for wearable tools. It was challenging to infer naturally affective states in an uncontrolled environment, particularly when there were numerous variations and occlusions. The concepts employed in GoogleNet aided in producing active and automatic assistance tool to help others understand the emotions of people with autism. The model classified fear, disgust, joy, anticipation, and sadness emotions. As there is no another available dataset for

Table 3 Testing results of the proposed model

Precision	Recall
0.97	0.96

autism, 15% of data were excluded from all datasets before training process to ensure model generality toward unseen data. The model testing produced significant results with accuracy equal to 98.54% and error rate 0.0166. That means high generality for model was done toward unknown data.

The model results proved that using state-of-the-art techniques can participate in increasing the applications and models that have an ability to recognize emotional states of ASD even with their facial impairment.

In addition, it can be concluded that collaborating facial expressions, eye gaze, and head movements has potential role as intervention tools in helping ASD like other diagnoses methods. Generally, the model was compared with another previous work which mentioned in cues generating section. The comparison was based on cues generation methods only, because the previous work focused on extracting complex emotions with using convolution neural network and long short-term memory, in spite of the test of previous work that achieved accuracy 89% toward unseen data. That approves the efficiency methods which are used in this work (Table 3).

References

1. Hassan A, Pinkwart N, Shafi M (2021) Serious games to improve social and emotional intelligence in children with autism. *Entertainment Comput* 38:100417. <https://doi.org/10.1016/j.entcom.2021.100417>
2. Sharma SR, Gonda X, Tarazi FI (2018) Autism spectrum disorder: classification, diagnosis and therapy. *Pharmacol Ther* 190:91–104. <https://doi.org/10.1016/j.pharmthera.2018.05.007>
3. Dawood A, Turner S, Perepa P (2018) Affective computational model to extract natural affective states of students with Asperger syndrome (AS) in computer-based learning environment. *IEEE Access* 6:67026–67034. <https://doi.org/10.1109/ACCESS.2018.2879619>
4. Faras H, Al Ateeqi N, Tidmarsh L (2010) Autism spectrum disorders. *Ann Saudi Med* 30(4):295–300. <https://doi.org/10.4103/0256-4947.65261>
5. Lin E, Tsai SJ (2016) Genome-wide microarray analysis of gene expression profiling in major depression and antidepressant therapy. *Prog Neuropsychopharmacol Biol Psychiatry* 64:334–340. <https://doi.org/10.1016/j.pnpbp.2015.02.008>
6. Joseph L, Pramod S, Nair LS (2017) Emotion recognition in a social robot for robot-assisted therapy to autistic treatment using deep learning. In: 2017 International conference on technological advancements in power and energy (TAP Energy). IEEE, pp 1–6. <https://doi.org/10.1109/TAPENERGY.2017.8397220>
7. Perner J, Wimmer H (1985) “John thinks that Mary thinks that...” attribution of second-order beliefs by 5-to 10-year-old children. *J Exp Child Psychol* 39(3):437–471. [https://doi.org/10.1016/0022-0965\(85\)90051-7](https://doi.org/10.1016/0022-0965(85)90051-7)
8. McKenzie K, Russell A, Golm D, Fairchild G (2022) Empathic accuracy and cognitive and affective empathy in young adults with and without autism spectrum disorder. *J Autism Dev Disord* 52(5):2004–2018. <https://doi.org/10.1007/s10803-021-05093-7>

9. Berggren S, Fletcher-Watson S, Milenkovic N, Marschik PB, Bölte S, Jonsson U (2018) Emotion recognition training in autism spectrum disorder: a systematic review of challenges related to generalizability. *Dev Neurorehabil* 21(3):141–154. <https://doi.org/10.1080/17518423.2017.1305004>
10. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Alkwai LM, Kumar S (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electr Eng* 107:108617. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
11. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerging Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>
12. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electr Eng* 108:108715. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
13. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun.* ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
14. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM (2023) An efficient hybrid deep learning model for denial of service detection in cyber physical systems. *IEEE Trans Netw Sci Eng.* <https://doi.org/10.1109/TNSE.2023.3273301>
15. Gupta U, Sharma R (2023) Analysis of criminal spatial events in India using exploratory data analysis and regression. *Comput Electr Eng* 109(Part A):108761. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108761>
16. Goyal B et al (2023) Detection of fake accounts on social media using multimodal data with deep learning. *IEEE Trans Comput Soc Syst.* <https://doi.org/10.1109/TCSS.2023.3296837>
17. Sneha PM, Sharma R, Ghosh U, Alnumay WS (2023) Internet of things and long-range antenna's; challenges, solutions and comparison in next generation systems. *Microprocess Microsyst* 104934. ISSN 0141-9331. <https://doi.org/10.1016/j.micpro.2023.104934>
18. Vohnout R et al (2023) Living lab long-term sustainability in hybrid access positive energy districts—a prosumager smart fog computing perspective. *IEEE Internet Things J.* <https://doi.org/10.1109/IOT.2023.3280594>
19. Yu X, Li W, Zhou X et al (2023) Deep learning personalized recommendation-based construction method of hybrid blockchain model. *Sci Rep* 13:17915. <https://doi.org/10.1038/s41598-023-39564-x>
20. Yadav S et al (2023) Video object detection from compressed formats for modern lightweight consumer electronics. *IEEE Trans Consum Electron.* <https://doi.org/10.1109/TCE.2023.3325480>
21. Singh A, Dewan S (2020) AutisMitr: emotion recognition assistive tool for autistic children. *Open Comput Sci* 10(1):259–269. <https://doi.org/10.1515/comp-2020-0006>
22. Drimalla H, Baskow I, Behnia B, Roepke S, Dziobek I (2021) Imitation and recognition of facial emotions in autism: a computer vision approach. *Mol Autism* 12:1–15. <https://doi.org/10.1186/s13229-021-00430-0>
23. Silva V, Soares F, Esteves JS, Santos CP, Pereira AP (2021) Fostering emotion recognition in children with autism spectrum disorder. *Multimodal Technol Interact* 5(10):57. <https://doi.org/10.3390/mti5100057>
24. Dawood A, Turner S, Perepa P (2019) Natural-Spontaneous affective-cognitive dataset for adult students with and without Asperger syndrome. *IEEE Access* 7:77990–77999. <https://doi.org/10.1109/ACCESS.2019.2921914>
25. Zhou X (ed) (2015) Proceedings of the 23rd ACM international conference on multimedia. Association for Computing Machinery.
26. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Li L (2020) A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 6:1122–1129. <https://doi.org/10.1016/j.eng.2020.04.010>

27. Cook J, Ramadas V (2020) When to consult precision-recall curves. Stand Genomic Sci 20(1):131–148. <https://doi.org/10.1177/1536867X20909693>
28. Sasank TS et al (2021) J Phys: Conf Ser 1879:032124
29. Shahab S, Agarwal P, Mufti T, Obaid AJ (2022) SIoT (Social Internet of Things): a review. In Fong S, Dey N, Joshi A (eds) ICT analysis and applications. Lecture notes in networks and systems, vol 314. Springer, Singapore. https://doi.org/10.1007/978-981-16-5655-2_28

Vehicle Classification and License Number Plate Detection Using Deep Learning



Kaushal Kishor, Ankit Shukla, and Anubhav Thakur

Abstract The amount of automobiles on the road is rising dramatically as well as the commercial revolution and the economy increase. This work seeks to recognize and categorize significant objects in a video stream (surveillance video); moreover, it is desirable to detect and recognize information on four-wheeler license plates. Furthermore, we examine the dataset and discuss the findings. Visualizations and conclusions are also included in the discussion. The performance of several classifiers is then compared using object classification. The accuracy of the model is investigated as a function of data multiplication. Finally, we will go through the methods and strategies utilized to recognize license plates. ALPR makes use of image processing for recognizing vehicles by their license plates without the direct involvement of a person by extracting the information from the picture of the vehicle or from a series of images. In this study, CNN, a deep learning approach, is used to distinguish as well as identify car license plates. This technology can recognize and locate license plate numbers. Digital cameras are used to detect vehicles.

Keywords Automatic license plate recognition · Convolution neural network · Vehicle detection · Object classification · Optical character recognition

1 Introduction

The automatic licence plate recognition system scans and identifies the licence plate of a vehicle. A vehicle is detected using optical character recognition (OCR) on photos. Closed circuit or traffic regulation enforcement webcams, for example, can be used for this strategy. These kinds of cameras may be found in a variety of areas, including police stations and computerized tollbooths. The ALPR can be used to save pictures recorded as well as cameras as text from the license plate; certain cameras may be capable of storing the driver's picture. ALPR is broken down into three stages:

K. Kishor (✉) · A. Shukla · A. Thakur
ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India
e-mail: kaushal.rastogi07@gmail.com

number plate removal, text categorization, and character recognition. The Number Plate Recognition approach involves gathering, recognising, and storing data such as pictures, number plate numbers, and database locations in order to identify persons or do additional research [1, 2].

The projection algorithm, template-matching technique, clustering analysis algorithm, and connecting domain algorithm were some of the most commonly utilized license plate character segmentation techniques at the time. The proposed algorithm in the memorandum can start by removing specific noise and lateral border intervention on both edges of the license plate, yet it exhibits poor categorization effect on Chinese protagonists under illumination variations and hostile climate; to overcome the text identification challenge, the literature employs lateral chopping and vertical prediction; nevertheless, this technique is extremely sensitive to license plate inclination [3].

Automatic license plate recognition is also utilized in parking reservation systems, toll online payment systems, as well as other systems that require authorization. By automating the procedure, it saves security officials a significant amount of time. In recent decades, computer vision technology has made considerable breakthroughs in a variety of real-world settings. Historically, car license plates were detected by determining the width, height, contour area, and so on using template-matching algorithms. Live CCTV footage is used to generate frames. To detect the autos, the frames are sent via the YOLO algorithm. The identified vehicles are saved as separate photos in a folder. These photographs are evaluated for plate information. The recognized license plate will be clipped and placed in its own folder. OCR is used to recognize the letters and numbers on these license plates. The collected text, along with the moment, date, and vehicle's license number, is then saved in a spreadsheet [4]. This is a technological combo in whom the Smartphone app enables the system to recognize and instantly retrieve a vehicle's registration plate number from electronically taken photographs. The procedure of automatically gathering the number plate is the detection and translation of visual data from an electronic picture to text format textual information or the ASCII code of a registration number [5].

The device is deployed near the entrance to regulate surveillance within severely limited territories, for example, militarized areas or regions within key government buildings such as Parliament and the Supreme Court, among others. The proposed technology scans the vehicle and pictures it. The division of an image is used to extract a photograph's automobile number plate region. For character segmentation, an OCR approach is utilized [6].

2 Literature Review

K-Nearest Neighbour (KNN) Algorithm for Licence Plate Letter Recognition and Classification. Automating licence plate identification is critical for movement management, which includes digitally collecting road tolls and managing parking

lots. This automation may cut management costs while increasing implementation efficiency. As image processing, classifiers, and computing performance on computers improve, we use Sobel operators to recognize object boundaries in order to extract license plate areas [7–9].

An automated approach for recognizing vehicle license plates using a classification algorithm. Multiple governments have conducted extensive research on license plate recognition. Because of the various types of number plates in use, the criteria for an automatic number plate recognition system vary by nation [10, 11].

Deep learning-based automatic license plate identification Two CNN models are employed in this investigation. As a result, two datasets including automobile photos and photos of license plates are necessary. Stanford vehicles' dataset from the internet is utilized to train the automobile pictures. The technology employs the YOLO method to teach the model how to recognize cars. The letters on these license plates are recognized using OCR. The information is then saved to an excel file, along with the time, date, and car's serial number. This method is more accurate than prior methods, and it also has the benefit of real-time implementation [9, 12, 13]. The goal of the Automated Machine Learning-Based Vehicle Plate Recognition Software is to link all smart vehicle number plate detection systems while improving efficiency in cost and reducing inaccuracy. This system may be utilized for a variety of purposes and has a low setup cost for the necessary gear [14].

Toll Collection Using Vehicle and License Plate Recognition and a New Dataset They propose an automatic framework for toll collection in this paper, which consists of three steps: vehicle type recognition, license plate localization, and reading. Unfortunately, due to visual differences produced by a variety of circumstances, each of the three processes becomes non-trivial. The traditional front-end vehicle decorations cause differences between cars of the identical kind [15].

A system for identifying vehicles that takes use of automated number plate recognition. He created a system that identifies the car and then captures the vehicle picture using Optical Character Recognition. A picture's vehicle number plate area is retrieved using image segmentation. For character recognition, an OCR approach is utilized. The collected data is then juxtaposed with database records to determine specific details such as car ownership, registration location, location, and so on [11, 16].

Image processing techniques are used in an innovative approach to the detection and recognition of vehicle registration number plates belonging to Indians. This study presents a new picture analysis approach for Indian registration plate recognition and detection that can handle noisy, cheap, cross-angled, non-standard font plate numbers. This work makes use of numerous methods for image processing in the preprocessing stage, including morphological modification, Gaussian smoothing, and Gaussian thresholding [17, 18].

Automated Recognition of License Plates They presented a system that focuses on the security of parking at any site using image processing and neural network. As a result, because when vehicle approaches the premises, the system must log the time and license plate data. While every ALPR system contains an image capturing system and a letter projection network, this system activates the camera with a sensor

network and an user interface with graphics that allows the user to handle the whole setup [19, 20].

This study describes methods for vision-based vehicle recognition and classification in a monocular picture series of traffic scenes captured by a stationary camera. There are three layers of processing: raw photos, region level, and vehicle level. Vehicles are represented as rectangular patches with dynamic behavior. The proposed technique is based on the formation of correspondences between areas and vehicles as the vehicles traverse the picture sequence [21]. The method's efficiency is demonstrated by experimental data from roadway situations [22]. We also present briefly an interactive camera calibration tool that we created for retrieving camera settings using user-selected picture attributes [23, 24].

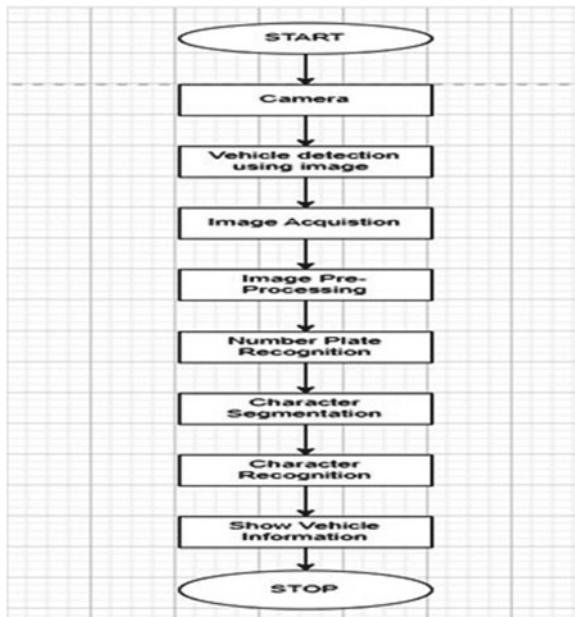
3 Proposed Model

The identification and recognition of license plates are primarily accomplished through three processes: license plate detection, text categorization, and pattern matching.

Figure 1 depicts the solution process for the suggested technique at its most basic level.

The architectural foundation for license plate identification and recognition is depicted in Fig. 2. It is divided into five stages:

Fig. 1 Block diagram



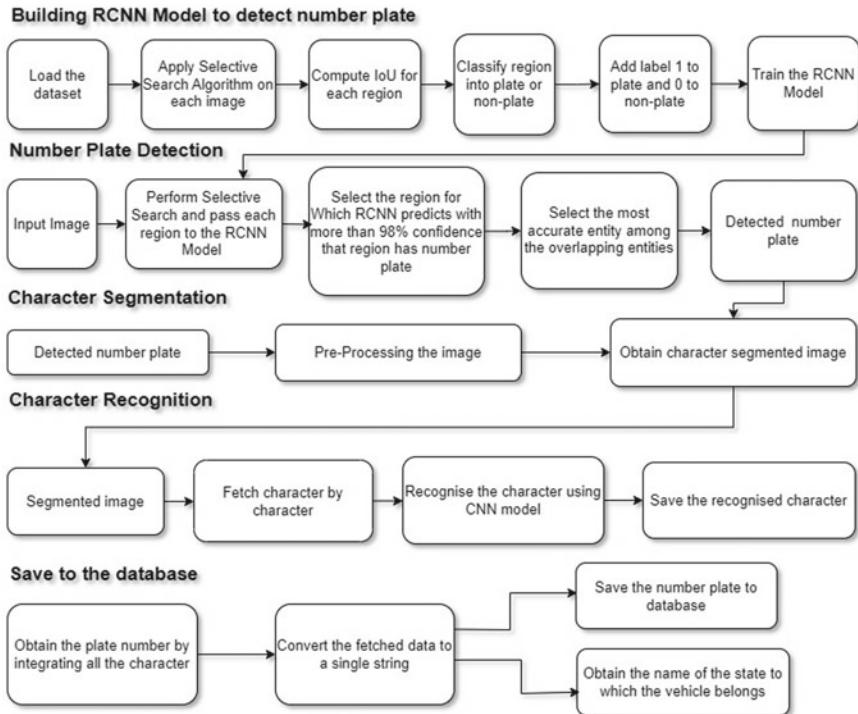


Fig. 2 Architectural framework

Step 1: Creating an RCNN model for detecting license plates

Step 2: Identification of number plates.

Step 3: Pattern extraction.

Step 4: Text recognition..

Step 5: Save to the database.

Developing the RCNN Model for License Plate Recognition—During the first phase, a regional convolution neural network (RCNN) model is constructed by applying a selective search algorithm to each image in the dataset. The algorithm divides the picture into 2000 areas and classifies each region based on the estimated IoU, identifying the regions as tiles or non-tiles, resulting in the creation of a new dataset known as use data. Train the RCNN model using the newly generated data.

Number Plate Identification—Phase 2: The number plate for the provided image is located. Running the Search Selection method and forwarding each region to the RCNN model accomplish this. The model offers confidence in each region, and the model predicts all areas with higher than 98% confidence. The maximum pressure function selects the highest accurate pressure among all chosen locations. And that area is classified as a fixed number region.

Character segmentation is a technique for dividing a photograph of a group of letters into sub-images that represent distinct characters. Phase 3 involves character segmentation from a specific number plate location.

Symbol Recognition—In phase 4, each symbol in the segmented picture is provided as input to the CNN model independently. The text of the numbers will be obtained after all of the characters have been identified.

In step 5, the numerical text is transformed to plain text and the first two characters recognized by the CNN model produced from the vehicle state are used. The wording on the plate and the vehicle's state are then saved in a database. Also, by entering the number plate text, you may get complete information on the car on the Vahan-Info website.

4 Result

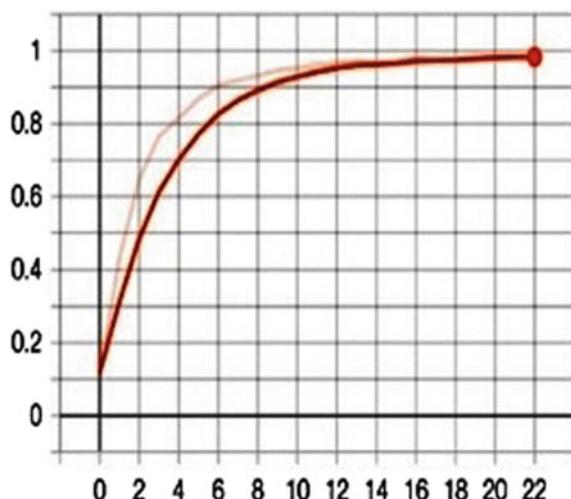
The results of the proposed methodology are discussed in this section.

RCNN Model—RCNN Model Results: This type is designed to detect license plates. The model produces the following outcomes.

- (1) During 23 periods, the model had an accuracy of 99.57%.
- (2) The RCNN model works well for the majority of the input photos.
- (3) It also works better if the image has numerous automobiles.

The system achieves an accuracy of 99.54% while using 23 periods. However, as the number of attributes rises, the accuracy decreases. Therefore, the accuracy is inversely related to the number of data attributes (Fig. 3).

Fig. 3 Attribute versus accuracy



After applying attributes to variable, we get a variable accuracy 99.07%, and if the attributes increases, the variable accuracy gets increased, so we can say by this project the variable accuracy is directly proportional to the data attributes (Fig. 4).

After applying attributes, we get a loss of 62%, and if we increase the attributes' number, there is no any loss of data, so loss is not proportional with the attributes. The loss is the fixed on any number of attributes (Fig. 5).

After applying attributes on variables, we get a variable loss of 49.9%, and if the attributes increases, then the loss is reduced, so we can say that variable loss is indirectly proportional to attributes (Fig. 6).

Fig. 4 Attribute versus accuracy

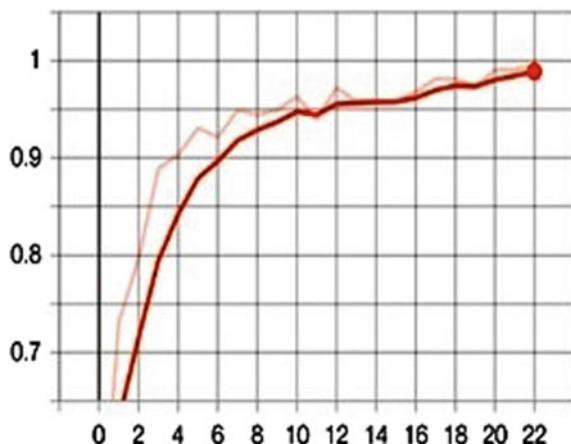


Fig. 5 Attribute versus accuracy

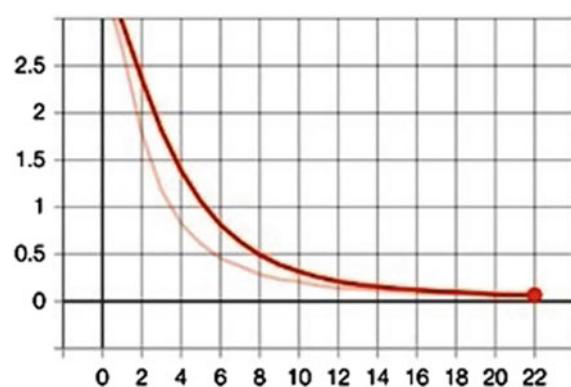
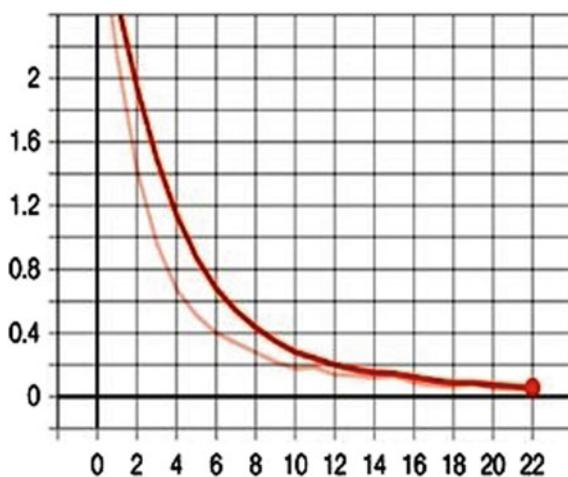


Fig. 6 Attribute versus accuracy



5 Conclusion

Using image processing methods, an innovative method for identifying and recognizing Indian vehicle license plates has been developed. In this study, the deep learning-based methodology outperforms image processing approaches in terms of reliability in real-world data. To fill this void, we propose a deep learning-based system for identifying and recognizing license plates. For digit recognition, we employ region-based convolution neural networks, and for character recognition, we use convolution neural networks. The location and details of the car are obtained after recognizing the text on the plate. The text of the vehicle's license plate and the state name are stored in a database to keep track of detected automobiles. The suggested deep learning approach for number plate identification and recognition has obtained 99.57% accuracy for the RCNN model. In the future, we may use the same methodology/approach to the dataset, which contains different number plate kinds in the context of font style, background color, and number plate text separated into two lines (usually seen in two-wheelers).

References

1. Khare P, Dudhe R, Chungade A, Naykinde A (2015) Advanced license number plate recognition system. *Int J Eng Res Technol (IJERT)* 3(6)
2. Wawage P, Oza S (2013) An approach for automatic detection of vehicle license plate and character recognition using classification algorithm
3. Yang Y (2020) License plate character segmentation algorithm based on improved regression model. *J Phys: Conf Ser* 1453:012030
4. Gnanaprakash V et al (2021) IOP Conf Ser: Mater Sci Eng 1084:012027. <https://doi.org/10.1088/1757-899X/1084/1/012027>

5. Kumar J, Birudu S, Narkedamilly L (2021) Automatic vehicle number plate recognition system using machine learning. IOP Conf Ser: Mater Sci Eng 1074:012012. <https://doi.org/10.1088/1757899X/1074/1/012012>
6. Usama M, Anwar H, Shahid M, Anwar A, Anwar S, Hlavacs H (2022) Vehicle and license plate recognition with novel dataset for toll collection. <https://doi.org/10.48550/arXiv.2202.05631>
7. Kishor K, Tyagi R, Bhati R, Rai BK (2023) Develop model for recognition of handwritten equation using machine learning. In: Mahapatra RP, Peddouji SK, Roy S, Parwekar P (eds) Proceedings of International conference on recent trends in computing. Lecture notes in networks and systems, vol 600. Springer, Singapore. https://doi.org/10.1007/978-981-19-8825-7_23
8. Gupta S, Tyagi S, Kishor K (2022) Study and development of self sanitizing smart elevator. In: Gupta D, Polkowski Z, Khanna A, Bhattacharyya S, Castillo O (eds) Proceedings of data analytics and management. Lecture notes on data engineering and communications technologies, vol 90. Springer, Singapore. https://doi.org/10.1007/978-981-16-6289-8_15
9. Qadri MT, Asif M (2009) Automatic number plate recognition system for vehicle identification using optical character recognition, pp 335–338. <https://doi.org/10.1109/ICETC.2009.54>
10. Kiran R, Varma P, Ganta S, Hari Krishna B, Svsrk P (2020) A novel method for Indian vehicle registration number plate detection and recognition using image processing techniques. Procedia Comput Sci 167:2623–2633. ISSN 1877-0509
11. Paneerselvam S, Gurudath P, Prithvi R, Ananth VG (2018) Automatic license plate recognition using image processing and neural network. ICTACT J Image Video Process 8:1786–1792. <https://doi.org/10.21917/ijivp.2018.0251>
12. Kishor K, Sharma R, Chhabra M (2022) Student performance prediction using technology of machine learning. In: Sharma DK, Peng SL, Sharma R, Zaitsev DA (eds) Micro-electronics and telecommunication engineering. Lecture notes in networks and systems, vol 373. Springer, Singapore. https://doi.org/10.1007/978-981-16-8721-1_53
13. Kishor K (2022) Communication-efficient federated learning. In: Yadav SP, Bhati BS, Mahato DP, Kumar S (eds) Federated learning for IoT applications. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-85559-8_9
14. Gupte S, Masoud O, Martin R, Papanikopoulos N (2002) Detection and classification of vehicles. IEEE Trans Intell Transp Syst 3:37–47. <https://doi.org/10.1109/6979.994794>
15. Kishor K, Singh P, Vashishta R (2023) Develop model for malicious traffic detection using deep learning. In: Sharma DK, Peng SL, Sharma R, Jeon G (eds) Micro-electronics and telecommunication engineering. Lecture notes in networks and systems, vol 617. Springer, Singapore. https://doi.org/10.1007/978-981-19-9512-5_8
16. Kishor K (2022) Personalized federated learning. In: Yadav SP, Bhati BS, Mahato DP, Kumar S (eds) Federated learning for IoT applications. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-85559-8_3
17. Selmi Z, Ben Halima M, Alimi AM (2017) Deep learning system for automatic license plate detection and recognition. In: 14th IAPR International conference on document analysis and recognition (ICDAR), Kyoto, Japan, pp 1132–1138. <https://doi.org/10.1109/ICDAR.2017.187>
18. Sharma A, Jha N, Kishor K (2022) Predict COVID-19 with chest X-ray. In: Gupta D, Polkowski Z, Khanna A, Bhattacharyya S, Castillo O (eds) Proceedings of data analytics and management. Lecture notes on data engineering and communications technologies, vol 90. Springer, Singapore. https://doi.org/10.1007/978-981-16-6289-8_16
19. Tas S, Sari O, Dalveren Y, Pazar S, Kara A, Derawi M (2022) Deep learning-based vehicle classification for low quality images. Sensors 22:4740. <https://doi.org/10.3390/s22134740>
20. Rai BK, Sharma S, Kumar G, Kishor K (2022) Recognition of different bird category using image processing. Int J Online Biomed Eng (iJOE) 18(07):101–114. <https://doi.org/10.3991/ijoe.v18i07.29639>
21. Kishor K, Nand P (2023) Wireless networks based in the cloud that support 5G. In: Cloud-based intelligent informative engineering for society 5.0, 1st edn. Chapman and Hall/CRC, New York, pp 23–40. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895-2>

22. Kishor K, Saxena N, Pandey D (2023) Cloud-based intelligent informative engineering for society 5.0, 1st edn. Chapman and Hall/CRC, New York, pp 1–234. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895>
23. Kishor K (2023) Cloud computing in blockchain. In: Cloud-based intelligent informative engineering for society 5.01 1st edn. Chapman and Hall/CRC, New York, pp 79–105. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895-5>
24. Ammar A, Koubaa A, Ahmed M, Saad A (2019) Aerial images processing for car detection using convolutional neural networks: comparison between faster R-CNN and YoloV3. <https://doi.org/10.20944/preprints201910.0195.v1>

Car Price Prediction Model Using ML



Kaushal Kishor , Akash Kumar, and Kabir Choudhary

Abstract Car manufacturing rates have climbed consistently over the previous decade, with million cars manufactured per year. This has given a significant boost to the vintage car market, whose is now entrenched as a thriving industry. The recent entrance of many internet portals and platforms has supplied buyers, customers, merchants, and vendors with the necessity to keep informed with the current situation and trends in order to know the true worth of a used automobile in the contemporary market. While there are several uses of machine learning in realistic life, one of their most apparent aspects is its usage in addressing prediction issues. Furthermore, there are an infinite variety of topics on which predictions can become made. This paper is about and revolves around a particular application. We will anticipate the market value of an obsolete vehicle using a machine learning method such as linear regression and construct a mathematical structure utilizing data that has been provided with a certain set of attributes.

Keywords Vehicle pricing prediction · Linear regression · Data comprehending · Data cleaning

1 Introduction

Predicting car prices is a crucial and significant endeavor, especially since the automobile is old and not brand-new. Consumers like leasing their vehicles, so it is a formal agreement among the consumer and dealer. Straight sellers, third parties, commercial entities, and insurance companies all fall under the seller group. Under a lease agreement, the purchasers make regular payments for the item over a specified time period. Due to the fact that these lease obligations depend on the projected value of the car, sellers are curious about the accurate projected value of their automobiles. Surveys indicate that it might be difficult and crucial to determine an aged car's

K. Kishor · A. Kumar · K. Choudhary
ABES Institute of Technology, Ghaziabad, Uttar Pradesh, India
e-mail: kaushal.rastogi07@gmail.com

fair projected price. Consequently, a trustworthy technique of estimating prices for used cars is needed. In this case, machine learning methods for prediction could be helpful. Machine learning employs both deductive and inductive thinkings. Deductive learning uses existing facts and knowledge to infer new facts and knowledge, as opposed to inductive machine learning, which develops new computer programs by finding patterns and trends in novel, undiscovered sets of data. We choose the multiple linear regression deductive techniques because it generates fresh values according to preexisting data. With this method, there is just one dependent variable, Y, and several independent variables, X. Direct or linear relationships exist between the variables. These are the objectives of this essay: Since there are several independent variables in this study, multiple linear regressions, one of the various varieties of linear regressions, are used. Due to the fact that there are thousands of used automobiles and the data for each one has values for several aspects, the data related with the inquiry was quite extensive. Data collection and analysis are both challenging. The declared purpose of this study is to explain why automotive costs in India are what they are. Linear regression was used to assist estimate predictions. Because quality is often determined by a variety of features and circumstances, a correct evaluation of automotive pricing needs specialized expertise. The costs of used vehicles fluctuate on the marketplace; therefore, both purchasers and vendors want a smart system that will facilitate accurate price forecast of market pricing as well as the proper price based on vehicle categorization. The gathering of data that contain the most crucial components, such as the following: (1) period of manufacture, (2) engine layout, (3) condition, (4) km covered, (5) horsepower, (6) number of doors, and (7) vehicle mass, is a significant constraint in such a system. It is obvious that the stated characteristics have an impact on the product's pricing; sadly, information regarding these qualities is not always accessible [1]. The benefit of machine learning (ML) is the capability of machines to gain knowledge via the use of mathematical frameworks and processed information without explicit guidance. A part of machine learning is computational intelligence. Predictive models are built when the data is examined using pattern detection techniques. Like a person, machine learning's findings allow for precise prediction using the most information and expertise [2].

Furthermore, the price of a car is heavily determined by the quantity of petrol it consumes and the amount of gasoline it consumes per mile, causing fuel demand to change often.

The following is the structure of this analysis:

- Data exploration and interpretation.,
- Data cleaning..
- The procedure for data preprocessing includes implementing feature engineering and expansion, feature selection by recursive feature elimination (RFE), model development, and evaluation and elimination of aberrations from linear regression assumptions.

2 Literature Review

Multiple research studies are under underway to provide forecasts on the expenses linked to the acquisition of pre-owned vehicles. Using historical data, researchers periodically forecast product prices. This research article forecasted automobile prices in Mauritius using multiple linear regression, k-nearest neighbor, Naive Bayes, and decision tree approaches. When the predictions from several tactics were evaluated, it turned out that the rates of these methods were relatively close. The decision tree technique and the Naive Bayes strategy, on the other hand, have been demonstrated to be incompetent categorizing and predicting numerical values. According to this research, a small sample size does not provide accurate prediction [3–5].

This paper developed random forest as an approach for forecasting used vehicle prices in 2019 [6]. The Kaggle dataset was utilized in the study to anticipate vintage vehicle rates, with a precision of 83.62% for test data and 95% for train data. By removing outliers and unnecessary variables from the dataset, the primary criteria for predicting this outcome were determined to be price, kilometers, brand, and kind of vehicle. In regards to precision, random forest beat earlier work utilizing these datasets [7–9]. This paper highlights the need for a model to anticipate the cost of used automobiles in Bosnia and Herzegovina [10].

As machine learning methodologies, this study used artificial neural networks, support vector machines, and random forests. However, the previously mentioned tactics were used in combination. The web scraped, implemented in the PHP computer language, was used to collect data from the internet address autopijaca.ba for the forecast. The relative capabilities of several algorithms were then assessed in order to determine which approach best fits the provided data [5, 11, 12].

The purpose of the system created by this article is to provide the user with a realistic estimate of how much the car will cost those [13].

The platform, which is a web application, can additionally present the individual with an inventory of alternatives various car types according to the specifications of the vehicle the user is searching for. It contributes to providing useful information to the prospective purchaser or seller upon which to base their decision. The method forecasts utilizing multiple linear regressions, and it was developed based on past information collected over extended amount duration [14, 15].

An associated effort that makes use of Support Vector Machines (SVM) to make predictions about the cost of leasing automobiles was given in this publication [16]. SVM beats multiple linear regressions in forecasting prices when a large dataset can be obtained, according to this study [17]. SVM is also effective at dealing with data with numerous dimensions while avoiding overpriced and under-fitting. SVM's main traits are discovered through a genetic process. However, the approach does not give any evidence to back up the assertion that support vector machines (SVM) outperform classical multiple regression analysis in terms of variance and mean standard deviation [18–20].

In order to estimate utilized automobile costs using a collection of information gathered from Kaggle that had 14 distinct attributes, it was suggested to apply a

supervised machine learning model utilizing K-nearest neighbors. After varying the value of K and changing the percent, the precision of this technique went up to 85%. As expected, increasing the ratio of training data to testing data leads to enhanced accuracy outcomes [21].

3 Proposed Model

Let us start by inspecting the dataset to determine its size, attribute names, and other specifics. The data types and column names of the dataset are introduced, and Python is used when linear regression is appropriate depending on the estimate.

Linear regression algorithms using training and test datasets yielded the following results: Linear regression is a type of supervised machine learning technique depending on the significance of an additional variable, which anticipates the outcome of a separate variable. In this case, the model searches for the best-fitting linear line between the independent and dependent variables.

3.1 Algorithm

- Step 1. Data Sourcing and Analysis
- Step 2. Data cleansing, manipulation, visualization, and outlier detection
- Step 3. Conduct EDA on the Prepared Dataset
- Step 4. Model Development
- Step 5. Data Splitting for Training and Testing
- Step 6. Model Construction
- Step 7. Train Data Residual Analysis
- Step 8. Making Predictions
- Step 9. Model Assessment
- Step 10. Final Interaction.

Least-Absolute the reduction and Picking Operator Lasso Regression, This is an L_1 regularization procedure, and the values $d1, d2, d3$, and so on indicate the spread between the real data lines and the hypothetical line in the picture above. The total of the squared of the distances across the points and the depicted curve is least squares. The optimal model for minimizing least squares is chosen in linear regression. A penalty component is introduced to the least squares in lasso regression. In other words, the model is chosen to minimize the loss function.

$D = \text{least squares} + \lambda * \sum (\text{absolute magnitudes of the coefficients})$.

The Lasso regression penalty is made up of all calculated parameters. Lambda may take any value from 0 to infinity. This number specifies the intensity with

which the regularization is carried out. Cross-validation is commonly used to pick it. The total of the coefficients' value in absolute terms is penalized by Lasso. The coefficients diminish and finally reach 0 as the lambda value grows. Lasso regression removes unimportant variables from our model in this way. Our regularized model may be somewhat more biased than linear regression, but it has less volatility in its subsequent predictions.

4 Result

This section discusses the suggested methodology's outcomes.

Figure 1 shows the result between actual price and predicted price, and we have an R-score error of 0.8880740, and accuracy 88.807%.

Whenever the r_2 metrics for all the regression techniques are compared, the decision tree methodology has an outstanding r_2 value of 0.9544, meaning simply said, the picking tree algorithm has produced the most precise forecasts when compared to the alternative method.

In Fig. 2, we can see that the red line indicates the initial parameters of the dataset and the blue line indicates the values projected using the decision tree regression. Both lines are extremely close, showing that the predictions are fairly accurate.

Figure 3 illustrates the results of a significant price reduction, which may be largely attributed to the increase in the number of miles driven and the age of the car, as shown by the data that was trained. As a result, the anticipated vehicle price is the response factor, and the number of km traveled each observation is the explanatory variable. If

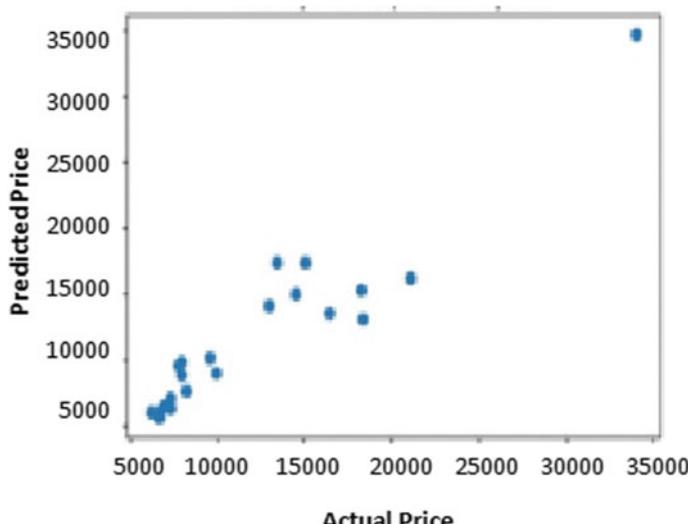


Fig. 1 Actual price versus predicted price

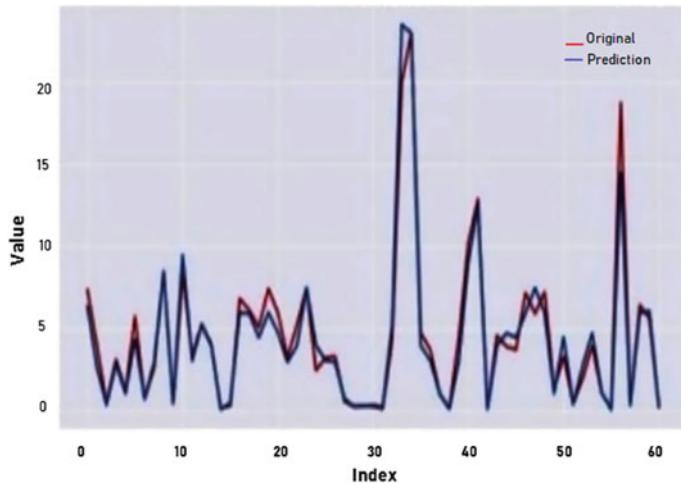


Fig. 2 Original versus prediction decision tree regression

factors like driving mode, vehicle circumstance, horsepower for transportation and number of doors, and car weight were also taken into account for secondhand cars, predictions would be more precise. The cost of automobiles will undoubtedly rise given the status market volatility, inflation, a scarcity of materials for the manufacture of new cars, a decline in the number of used cars available, and the volatility of petrol prices.

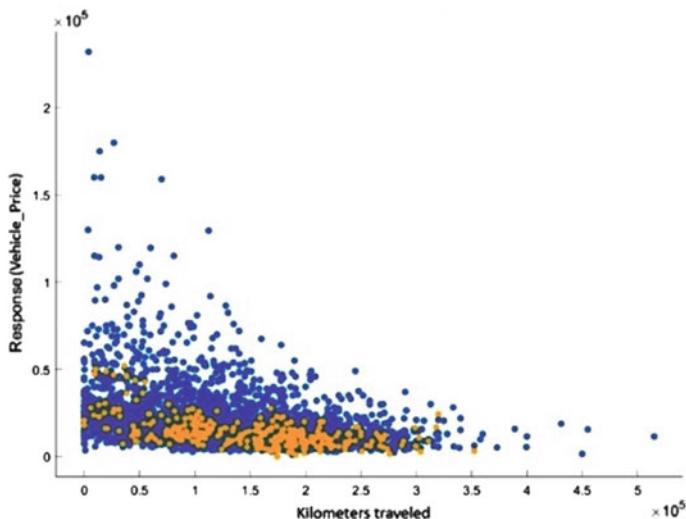


Fig. 3 Using regressors from a decision tree, predict the cost of vehicles

5 Conclusion

Analyzing used-car price is difficult owing to the enormous assortment of indicators and elements that must be considered in order to provide relevant results. The first and most important phases are data collected and handled. The representation was then built and implemented in order to perform calculations and provide results. During the integration of several regression algorithms to the model, it was determined that the decision tree technique performed the best, having the highest r^2 value of 0.95, indicating that it offered particularly precise projections, as seen in the line graph original versus prediction.

6 Future Scope

The sophisticated design will eventually be included in web-based and mobile-based apps for general usage. In addition, following the data gathering period, the pandemic-related shortages of semiconductors caused a rise in automobile costs and had a significant impact on the used-car market. As a result, frequent gathering and evaluating of information are necessary; perhaps we would have a program for immediate processing.

References

1. Samruddhi K, Kumar RA (2020) Used car price prediction using K-nearest neighbor based model. *Int J Innov Res Appl Sci Eng* 4:629–632
2. Kishor K (2023) Study of quantum computing for data analytics of predictive and prescriptive analytics models. In: Quantum-safe cryptography. De Gruyter, pp 121–146. ISBN 978-3-11-079800-5 e-ISBN (PDF) 978-3-11-079815-9 e-ISBN (EPUB) 978-3-11-079836-4 ISSN 2940-0112. <https://doi.org/10.1515/9783110798159-010>
3. Pudaruth S (2014) Predicting the price of used cars using machine learning techniques. *Int J Inf Comput Technol* 4(7):753–764. Available at: <http://www.irphouse.com>
4. Kishor K, Sharma R, Chhabra M (2022) Student performance prediction using technology of machine learning. In: Sharma DK, Peng SL, Sharma R, Zaitsev DA (eds) Micro-electronics and telecommunication engineering. Lecture notes in networks and systems, vol 373. Springer, Singapore. https://doi.org/10.1007/978-981-16-8721-1_53
5. Kishor K (2022) Communication-efficient federated learning. In: Yadav SP, Bhati BS, Mahato DP, Kumar S (eds) Federated learning for IoT applications. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-85559-8_9
6. Pal N et al (2019) How much is my car worth? A methodology for predicting used cars' prices using random forest. *Adv Intell Syst Comput* 886:413–422. https://doi.org/10.1007/978-3-030-03402-3_28
7. Sharma R, Maurya SK, Kishor K (2021) Student performance prediction using technology of machine learning (July 3, 2021). In Proceedings of the International conference on innovative computing & communication (ICICC) 2021. Available at SSRN: <https://ssrn.com/abstract=3879645> or <https://doi.org/10.2139/ssrn.3879645>

8. Jain A, Sharma Y, Kishor K (2021) Prediction and analysis of financial trends using MI algorithm (July 11, 2021). In: Proceedings of the International conference on innovative computing & communication (ICICC) 2021, Available at SSRN: <https://ssrn.com/abstract=3884458> or <https://doi.org/10.2139/ssrn.3884458>
9. Tyagi D, Sharma D, Singh R, Kishor K (2020) Real time ‘driver drowsiness’ & monitoring & detection techniques. *Int J Innov Technol Explor Eng* 9(8):280–284. <https://doi.org/10.35940/ijitee.H6273.069820>
10. Gecig E et al (2019) Car price prediction using machine learning techniques. *TEM J* 8(1):113–118. <https://doi.org/10.18421/TEM81-16>
11. Kishor K (2022) Personalized federated learning. In: Yadav SP, Bhati BS, Mahato DP, Kumar S (eds) Federated learning for IoT applications. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-85559-8_3
12. Gupta S, Tyagi S, Kishor K (2022) Study and development of self sanitizing smart elevator. In: Gupta D, Polkowski Z, Khanna A, Bhattacharyya S, Castillo O (eds) Proceedings of data analytics and management. Lecture notes on data engineering and communications technologies, vol 90. Springer, Singapore. https://doi.org/10.1007/978-981-16-6289-8_15
13. Dholiya M et al (2019) Automobile resale system using machine learning. *Int Res J Eng Technol (IRJET)* 6(4):3122–3125
14. Kishor K, Tyagi R, Bhati R, Rai BK (2023) Develop model for recognition of handwritten equation using machine learning. In: Mahapatra RP, Peddaju SK, Roy S, Parwekar P (eds) Proceedings of International conference on recent trends in computing. Lecture notes in networks and systems, vol 600. Springer, Singapore. https://doi.org/10.1007/978-981-19-8825-7_23
15. Kishor K, Saxena N, Pandey D (2023) Cloud-based intelligent informative engineering for society 5.0, 1st edn. Chapman and Hall/CRC, New York, pp. 1–234. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895>
16. Listiani M (2009) Support vector regression analysis for price prediction in a car leasing application, technology. Hamburg University of Technology
17. Kishor K, Nand P (2023) Wireless networks based in the cloud that support 5G. In: Cloud-based intelligent informative engineering for society 5.0, 1st edn. Chapman and Hall/CRC, New York, pp 23–40. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895-2>
18. Kishor K (2023) Cloud computing in blockchain. In: Cloud-based intelligent informative engineering for society 5.0, 1st edn. Chapman and Hall/CRC, New York, pp 79–105. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895-5>
19. Kishor K (2023) Impact of cloud computing on entrepreneurship, cost, and security. In: Cloud-based intelligent informative engineering for society 5.0, 1st edn. CRC Press, New York, pp 171–191. eBook ISBN: 9781003213895. <https://doi.org/10.1201/9781003213895-10>
20. Kishor K, Pandey D (2022) Study and development of efficient air quality prediction system embedded with machine learning and IoT. In Gupta D et al (eds) Proceeding International conference on innovative computing and communications. Lecture notes in networks, systems, vol 471. Springer, Singapore. https://doi.org/10.1007/978-981-19-2535-1_24
21. Samruddhi K, Kumar DR (2020) Used car price prediction using K-nearest neighbor based model. *Int J Innov Res Appl Sci Eng (IJIRASE)* 4(3):686–689

Effects of Material Deformation on U-shaped Optical Fiber Sensor



**Mohd Ashraf, Mainuddin, Mirza Tariq Beg, Ananta Sekia,
and Sanjai K. Dwivedi**

Abstract This work presents an evanescent wave-based U-shaped plastic optical fiber sensor. The effect of bend-induced material deformation on numerical aperture (N.A) and V-number has been thoroughly investigated. With sufficiently smaller bend radius, the numerical aperture of a U-bent fiber decreases toward zero near the inner curvature of the bending. The present scenario causes entire optical power loss at interface between core-clad (sample region) interfaces, resulting in no detectable power. It has been observed that as the bending radius increases, the local N.A value decreases while the refractive index of the surrounding region remains fixed. As a result, the fractional power in the surrounding region rises, resulting in increased absorbance and sensitivity of the EW absorption-based optical fiber sensor.

Keywords Evanescent wave · U-bent fiber · Numerical aperture · V-number · Plastic optical fiber · Sensor

1 Introduction

In the literature, a large variety of optical sensors have been reported [1–10] for various applications. Out of many optical sensor, glass fiber-based optical sensors require expensive and complex equipment for fiber processing steps because of the

M. Ashraf · Mainuddin (✉) · M. T. Beg

Department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi, India

e-mail: mainuddin@jmi.ac.in

M. T. Beg

e-mail: mtbeg@jmi.ac.in

A. Sekia

DRL, DRDO, Tezpur, Assam, India

S. K. Dwivedi

DOP, DRDO, Delhi, India

fragile nature of glass. Plastic optic fiber (POF) is preferred over glass fiber due to ease of fabrication and repeatability. POF sensors have become a promising candidate for quantitative chemical detection. Numerous researchers have used various geometries/structures to increase the sensitivity of fiber sensors. S-shaped [11] and D-shaped [12], tapered shapes [13], coiled shapes [14], taper-in-taper [15], core mismatch structures [16], fiber ball structures [17], convex tapered fiber structures [18], tapered U-shaped [19], double tapered U-shaped [20], and U-shaped [21, 22] are some of these geometries. Among these geometries, U-shaped probe sensors are found to have 10 times more sensitivity than straight fiber sensors [23]. Several studies have described the use of fiber optics with U-shaped geometry for the detection of many parameters, including relative humidity [24], pH value [25], ethanol [26], ammonia gas [27, 28], iron impurities in water [29], salinity [30, 31], and others. The important feature of plastic optical fiber is their high numerical aperture and large diameter [32]. In order to make a U-bend POF sensor, the bending region needs to be heated at proper temperature and bent with the right amount of force [24]. When a POF bends into a U-shape, it undergoes a change in refractive index (RI) within its core, resulting in a corresponding alteration in the fiber's local numerical aperture (N.A.). Thus, besides geometrical factors, the deformation effects [33, 34] also contribute to the variations observed in the local N.A. of polymer optical fiber (POF).

This paper presents the results of an in-depth investigation into the numerical aperture and V-number effects that are brought about by bend-induced material deformation.

2 Theory

The concentration of sample solution can be estimated using fiber optics-based absorption spectroscopy. When light penetrates the core of a fiber, a fraction of it enters the cladding and produces an evanescent wave (EW). EW amplitude is denoted as $E(y)$ that diminishes exponentially [35] with increasing radial distance y from the core. This behavior can be represented by the relation:

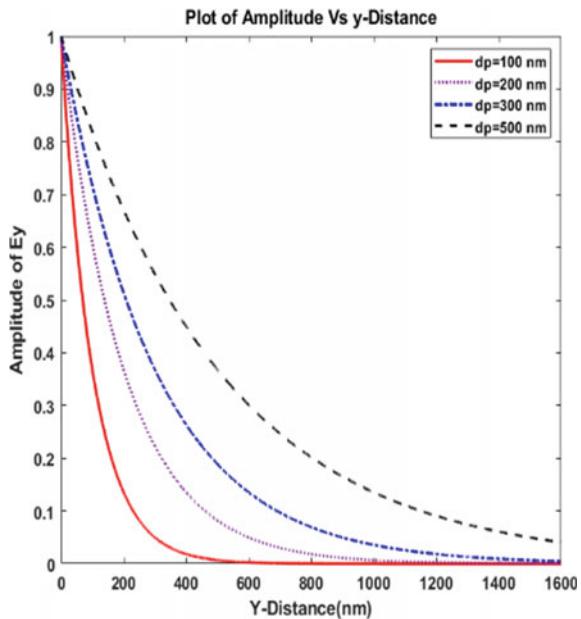
$$E(y) = E_0 \exp\left(-y/d_p\right) \quad (1)$$

where E_0 is the amplitude of the EW at the core-cladding interface and penetration depth is represented by d_p . The attenuation of waves is caused by this exponential decay of light with distance y , as shown in Fig. 1.

According to [36], the penetration depth (d_p) depends on the propagating light wavelength (λ), core (n_1) and cladding (n_2) RI, and incidence angle (θ_i).

$$d_p = \frac{\lambda}{2\pi n_1} \sqrt{\left[\sin^2 \theta_i - \left(\frac{n_2}{n_1}\right)^2\right]} \quad (2)$$

Fig. 1 Variation of amplitude of $E(y)$ with respect to y distance based on Eq. (1) at different penetration depth



The EW in a surrounding medium is normally very weak [37], which causes the sensor to have limited sensitivity. The sensitivity of POF can be enhanced by the various geometrical deformations [38] on the fiber optic detecting region (probe). Among available geometries, U-shape is an excellent geometry with increased evanescent wave penetration depth. Since, bending the fiber (Fig. 2) considerably maximizes the amount of light that leaks into the medium around it, which in turn enhances the d_p , and thus the sensitivity of sensor.

Apart from the penetration depth, the numerical aperture (N.A) also affects the sensor's sensitivity. The sensitivity of the sensor with a uniform detecting region is

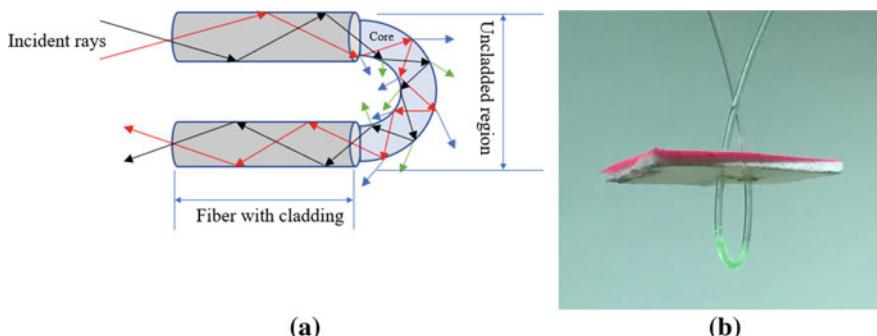


Fig. 2 **a** Ray model of U-bend de-cladded fiber, **b** light coming out from the U-bend fiber region

improved by an increase in the N.A. However, N.A, on the other hand changes with the bending of the fiber and this variation can be computed as [39].

$$N.A = \sqrt{\left(n_1^2 - n_m^2 \left(\frac{R+r}{R+a} \right)^2 \right)} \quad (3)$$

where n_1, n_m are the RI of core and cladding (surrounding region) respectively. Where $a(-r \leq a \leq r)$ is the position in the core, radius of fiber and bending radius are represented by r and R , respectively.

3 Design Considerations and Results

3.1 Sensor Characteristics

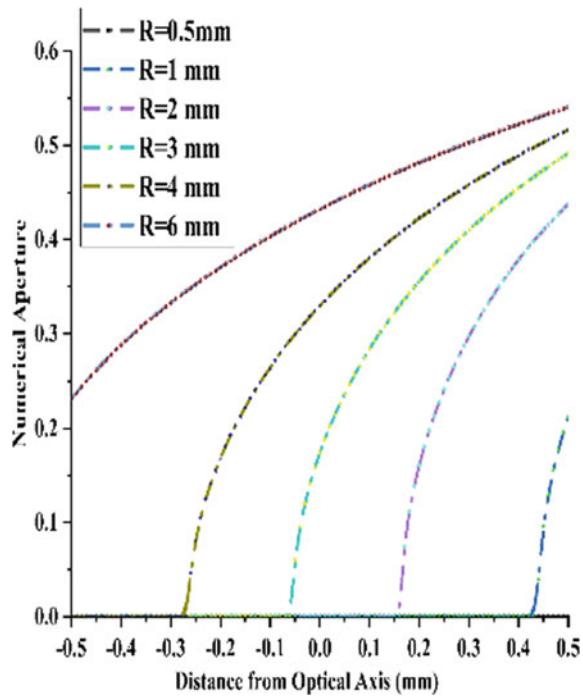
At sufficient small bend radius, the numerical aperture of a U-bend fiber decreases toward zero near the inner curvature of the bent resulting in positive non-zero N.A only for a portion of the fiber core. This means that if the fiber is acutely bend then, light travels only in a portion of the POF core, and most of the light waves escape the POF core. The escaped light waves will be available for absorption, which consequently leads to enhanced sensor sensitivity. However, the construction of a U-bend fiber sensor probe on the other hand required opposing force gradients across the cross-section of the POF along its optical axis. Since bare POF is an optically isotropic material and undergoes photoelastic strain deformation ranging from 2–3% and can even undergo plastic deformation of up to ~15% [29]. Bending the POF into a U-shape cause's material deformation that exceeds the elastic and plastic limitations, and affecting fiber core RI. The change in RI of the POF core has an effect on the N.A of the fiber and can be represented mathematically as [29].

$$N.A = \sqrt{\left\{ \left(1.49 \left(1 - 0.1391 \times \left[\frac{a}{R} \right] \right) \right)^2 - n_m^2 \left(\frac{R+r}{R+a} \right)^2 \right\}} \quad (4)$$

The N.A of a POF with a diameter of 1 mm was calculated using Eq. (4) and results are depicted in Fig. 3 for various bending radii (R).

The plot against the N.A and RI of the surrounding region at a constant bending radius of 1.5 mm is presented in Fig. 4. The simulation shows that more light leaks out of the fiber when bend-induced material deformation is considered. The simulations result demonstrate that the changes of the local N.A of POF is influenced by deformation effects as well as geometrical impacts. The fractional power given by Eq. (5) can be determined with the aid of the fiber's V-number.

Fig. 3 Variation of N.A with respect to the distance from the optical axis



$$V = \frac{2\pi r}{\lambda} N.A \quad (5)$$

By substituting the value of local numerical aperture from relation (4) into relation (5), we derive the effective V-number as given by

$$V = \frac{2\pi r}{\lambda} \sqrt{\left\{ \left(1.49 \left(1 - 0.1391 \times \left[\frac{a}{R} \right] \right) \right)^2 - n_m^2 \left(\frac{R+r}{R+a} \right)^2 \right\}} \quad (6)$$

It can be observed from relation (6) that there is a change in the V-number of the POF whenever there is a change in either the RI of the cladding region or the bending radius. This change can also be seen in Fig. 5. Since, the fractional power in the clad can be used to estimate the evanescent wave power (power in the cladding region).

$$\frac{P_{clad}}{P} = \frac{4\sqrt{2}}{3V} \quad (7)$$

By placing the effective V-number in place of V in the expression (7), an expression has been obtained for effective fractional power in cladding (relation 8), which integrates the simultaneous impact of sensor geometry and bend-induced material deformation in EW Fiber Sensors (EWFS).

Fig. 4 Variation of N.A and refractive index of medium for U-shaped POF at fixed bend radius ($R = 1.5$ mm)

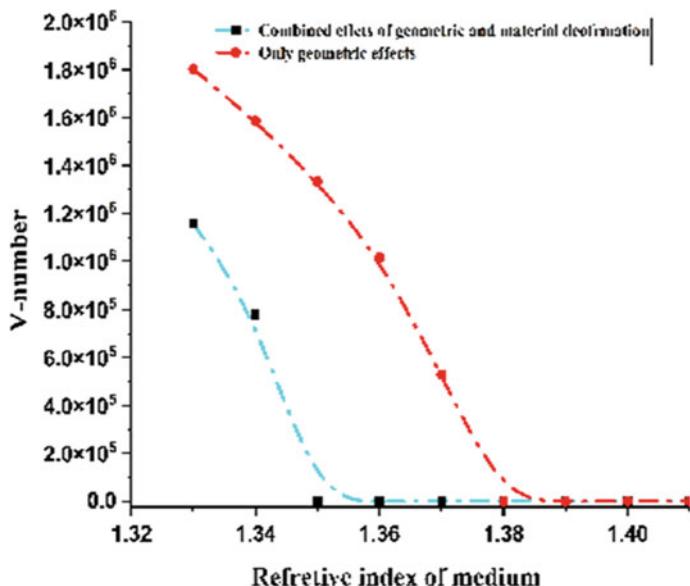
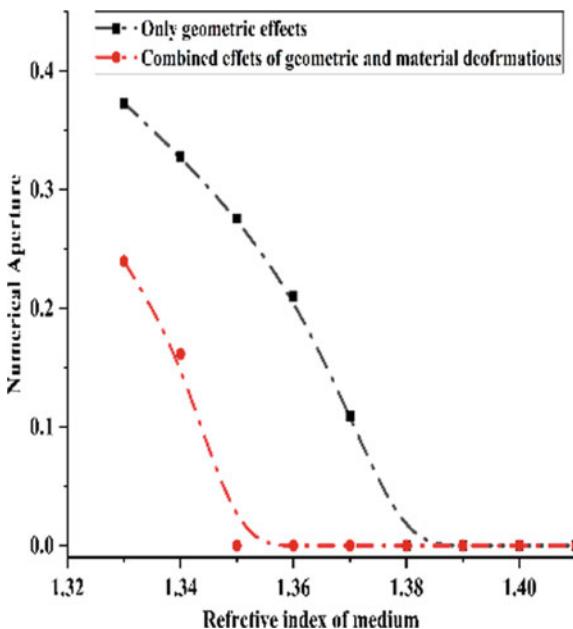


Fig. 5 Plot against V-number and refractive index of medium for U-shaped POF at fixed bend radius ($R = 1.5$ mm)

$$\frac{P_{clad}}{P} = \frac{2\sqrt{2}\lambda}{3\pi r \sqrt{\left\{ \left(1.49 \left(1 - 0.1391 \times \left[\frac{a}{R} \right] \right) \right)^2 - n_m^2 \left(\frac{R+r}{R+a} \right)^2 \right\}}} \quad (8)$$

where P_{clad} is the power in clad and P is the total power. From relation (8), it has been found that increasing the V-number reduces the fractional power in the clad. Therefore, selecting the V-number is critical for estimating the evanescent power of the POF sensor. Since the fiber deformation influences the local numerical aperture, which consequently changes the V-number and hence changes the fractional power contained in the POF cladding.

3.2 *Evanescent Wave Absorbance*

When the cladding is removed from a fiber, the EW that exists at the boundary of the fiber's core can interact with the external environment. The amount of power that is transmitted at the end of this de-cladded region can be determined by

$$P_l = P_0 \times e^{-\gamma l} \quad (9)$$

where the amount of optical power transmitted across a given length is denoted by P_l , while P_0 represents the power launched into the POF. The EW absorption coefficient (γ) directly depends on the fractional power present in the clad (which is the external environment surrounding the fiber's core) [37]. Additionally, the Beer-Lambert law states that EW absorbance directly depends on the absorption coefficient. Therefore, absorbance (A) is proportional to the cladding fractional power, and can be calculated using following relation.

$$A = \frac{P_{clad}}{2.303 P} \times \alpha C l \quad (10)$$

Substituting the value of fractional power from Eq. (8) into the Eq. (10), we derive the absorbance (A) for U-shaped EW absorption-based fiber optic sensor.

$$A = \frac{0.130 \times \alpha C l \lambda}{r \sqrt{\left\{ \left(1.49 \left(1 - 0.1391 \times \left[\frac{a}{R} \right] \right) \right)^2 - n_m^2 \left(\frac{R+r}{R+a} \right)^2 \right\}}} \quad (11)$$

Thus, evanescent wave absorbance depends on the mediums refractive index as well as material deformation effects. Additionally, the absorbance is directly related to factors such as the absorption coefficient (α) of the analyte medium, the concentration of molecule (C), the wavelength (λ), and the sensing region length (l).

3.3 Sensitivity

Sensitivity of a sensor refers to the degree to which the sensor can detect and respond to changes in the environment or physical quantity being measured. Sensitivity is typically expressed as [40]

$$S = \frac{\gamma \in l}{\alpha} \quad (12)$$

According to Eq. (12), EW absorption sensor's sensitivity is closely related to the EW absorption coefficient of the molecule (γ) [41]. It can be inferred that with the increase in γ values, the sensor's sensitivity improves. As we observed, fractional power is dependent on many factors such as the operating wavelength, geometry of the POF probe, bend-induced material deformation, fiber bending radius, and RI of the medium (n_m), etc. Hence, the EW absorption-based POF sensor sensitivity is proportional to cladding fractional power. Thus, Eq. (12) becomes

$$S = \frac{P_{\text{clad}}}{P} \times \frac{\in l}{\alpha} \quad (13)$$

Scholars have suggested a number of approaches to increase POF sensor's sensitivity. These designs have been successful in enhancing the amount of fractional power by increasing the number of total internal reflection events that occur over shorter distances.

4 Conclusion

The article has presented a thorough examination of how the deformation of the material caused by bending affects the numerical aperture (N.A) and light propagation within the core. The investigation reveals that when the bend radius is less than 1 mm, the numerical aperture of a bent plastic optical fiber (POF) falls to zero near the bend's inner curvature. This causes a substantial loss of directed optical power at the core-analyte interface (cladding) and therefore, there is no light detected at the end. Furthermore, changes in the RI of the surrounding medium can impact the N.A of the POF in the sensing region. As the bend radius increases, the local N.A value decreases at the fixed RI of the sample medium. This allows extra light waves to escape from the fiber and enter the medium, causing an increase in the fractional power in the surrounding medium. Consequently, this enhances the absorbance and sensitivity of the EW absorption-based POF sensor.

Acknowledgements This work is supported by the Life Science Research Board (LSRB, DRDO) under the DRDO project LSRB-393.

References

1. Fallah H, Asadishad T, Parsanasab GM, Harun SW, Mohammed WS, Yasin M (2021) Optical fiber biosensor toward E-coli bacterial detection on the pollutant water. *Eng J* 25(12):1–8
2. Pathak A, Gangwar R, Priyadarshni P, Singh V (2017) A robust optical fiber sensor for the detection of petrol adulteration. *Optik Int J Light Electron Opt* 149:43–48. <https://doi.org/10.1016/j.ijleo.2017.09.036>
3. Biswas N, Bewtra JK, Taylor KE (2005) A simple colorimetric method for analysis of aqueous phenylene diamines and aniline. *J Environ Eng Sci* 4(6):423–427. <https://doi.org/10.11139/s05-005>
4. Lin M, Hu X, Pan D, Han H (2018) Determination of iron in seawater: from the laboratory to in situ measurements. *Talanta* 188:135–144. <https://doi.org/10.1016/j.talanta.2018.05.071>
5. Mainuddin, Beg MT, Moinuddin, Tyagi RK, Rajesh R, Singhal G, Dawar AL (2005) Optical spectroscopic based in-line iodine flow measurement system—an application to COIL. *Sensors Actuators B: Chem* 109(2):375–380. <https://doi.org/10.1016/j.snb.2005.01.004>
6. Jadhav MS, Laxmeshwar LS, Akki JF, Raikar PU, Tangod VB, Raikar US (2017) Multi-mode fiber optic sensor for adulterant traces in edible oil using nanotechnology technique. *Mater Today: Proc* 4(11, Part 3):11910–11914. ISSN 2214-7853. <https://doi.org/10.1016/j.matpr.2017.09.111>
7. Mainuddin, Singhal G, Tyagi RK, Maini AK (2011) Diagnostics and data acquisition for chemical oxygen iodine laser. *IEEE Trans Instrum Meas* 16(6):1747–1756. <https://doi.org/10.1109/TIM.2011.2178727>
8. Fahim F, Mainuddin M, Mittal U, Kumar J, Nimal AT (2021) Novel SAW CWA detector using temperature programmed desorption. *IEEE Sens J* 21(5):5914–5922. <https://doi.org/10.1109/jsen.2020.3042766>
9. Singh L, Kumar G, Jain S, Kaushik BK (2021) A novel plus shaped cavity based optical fiber sensor for the detection of Escherichia-Coli. *Results Opt* 5:100156. ISSN 2666-9501. <https://doi.org/10.1016/j.rio.2021.100156>
10. Sabeena Begam S, Vimala J, Selvachandran G, Ngan TT, Sharma R (2020) Similarity measure of lattice ordered multi-fuzzy soft sets based on set theoretic approach and its application in decision making. *Mathematics* 8:1255
11. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Tien BD, Priyadarshini I, Sarkar M, Le T (2020) Crime rate detection using social media of different crime locations and Twitter part-of-speech tagger with brown clustering, pp 4287–4299
12. Nguyen PT, Ha DH, Avand M, Jaafari A, Nguyen HD, Al-Ansari N, Van Phong T, Sharma R, Kumar R, Le HV, Ho LS, Prakash I, Pham BT (2020) Soft computing ensemble models based on logistic regression for groundwater potential mapping. *Appl Sci* 10:2469
13. Jha S et al (2019) Deep learning approach for software maintainability metrics prediction. *IEEE Access* 7:61840–61855
14. Sharma R, Kumar R, Sharma DK, Son LH, Priyadarshini I, Pham BT, Bui DT, Rai S (2019) Inferring air pollution from air quality index by different geographical areas: case study in India. *Air Qual Atmos Health* 12:1347–1357
15. Sharma R, Kumar R, Singh PK, Raboaca MS, Felseghi R-A (2020) A systematic study on the analysis of the emission of CO, CO₂ and HC for four-wheelers and its impact on the sustainable ecosystem. *Sustainability* 12:6707
16. Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I, Le D-N (2020) Global forecasting confirmed and fatal cases of COVID-19 outbreak using autoregressive integrated moving average model. *Front Public Health* 8:580327. <https://doi.org/10.3389/fpubh.2020.580327>
17. Malik PK, Sharma R, Singh R, Gehlot A, Satapathy SC, Alnumay WS, Pelusi D, Ghosh U, Nayak J (2021) Industrial internet of things and its applications in industry 4.0: state of the art. *Comput Commun* 166:125–139. ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2020.11.016>

18. Sharma R, Kumar R, Satapathy SC, Al-Ansari N, Singh KK, Mahapatra RP, Agarwal AK, Le HV, Pham BT (2020) Analysis of water pollution using different physicochemical parameters: a study of Yamuna River. *Front Environ Sci* 8:581591. <https://doi.org/10.3389/fenvs.2020.581591>
19. Dansana D, Kumar R, Parida A, Sharma R, Adhikari JD et al (2021) Using susceptible-exposed-infectious-recovered model to forecast coronavirus outbreak. *Comput Mater Continua* 67(2):1595–1612
20. Vo MT, Vo AH, Nguyen T, Sharma R, Le T (2021) Dealing with the class imbalance problem in the detection of fake job descriptions. *Comput Mater Continua* 68(1):521–535
21. Sachan S, Sharma R, Sehgal A (2021) Energy efficient scheme for better connectivity in sustainable mobile wireless sensor networks. *Sustain Comput: Inf Syst* 30:100504
22. Ghanem S, Kanungo P, Panda G et al (2021) Lane detection under artificial colored light in tunnels and on highways: an IoT-based framework for smart city infrastructure. *Complex Intell Syst*. <https://doi.org/10.1007/s40747-021-00381-2>
23. Sachan S, Sharma R, Sehgal A (2021) SINR based energy optimization schemes for 5G vehicular sensor networks. *Wireless Pers Commun*. <https://doi.org/10.1007/s11277-021-08561-6>
24. Khijwania SK, Srinivasan KL, Singh JP (2005) An evanescent-wave optical fiber relative humidity sensor with enhanced sensitivity. *Sens Actuators, B Chem* 104(2):217–222. <https://doi.org/10.1016/j.snb.2004.05.012>
25. Gupta BD, Sharma NK (2002) Fabrication and characterization of U-shaped fiber-optic pH probes. *Sens Actuators, B Chem* 82(1):89–93
26. Gao SS, Qiu HW, Zhang C, Jiang SZ, Li Z, Liu XY, Yue WW, Yang C, Huo YY, Feng DJ, Li HS (2016) Absorbance response of a graphene oxide coated U-bent optical fiber sensor for aqueous ethanol detection. *RSC Adv* 6(19):15808–15815. <https://doi.org/10.1039/C5RA22211G>
27. Cao W, Duan Y (2005) Optical fiber-based evanescent ammonia sensor. *Sens Actuators, B Chem* 110(2):252–259. <https://doi.org/10.1016/j.snb.2005.02.015>
28. Korposh S, Okuda H, Wang T, James S, Lee S-W (2015) U-shaped evanescent wave optical fibre sensor based on a porphyrin anchored nanoassembled thin film for high sensitivity ammonia detection. *Proc SPIE Int Soc Opt Eng* 9655. <https://doi.org/10.1117/12.2184438>
29. Ashraf M, Mainuddin, Beg MT, Moin F, Rajesh R, Singhal G (2022) U-bent plastic optical fiber sensor for iron in iron supplements. *IEEE Sens J* 22(15):14921–14928. <https://doi.org/10.1109/JSEN.2022.3187829>
30. Stupar D, Bajic J, Joza A, Dakic B, Slankamenac M, Zivanov M, Cibula E (2012) Remote monitoring of water salinity by using side-polished fiber-optic U-shaped sensor. In: 15th International power electronics and motion control conference and exposition, EPE-PEMC 2012 ECCE Europe. LS4c.4-1. <https://doi.org/10.1109/EPEPEMC.2012.6397458>
31. Ashraf M, Mainuddin (2023) Simulation of optical FBG based sensor for measurement of temperature, strain and salinity. In: Tiwari M, Ismail Y, Verma K, Garg AK (eds) Optical and wireless technologies (OWT 2021). Lecture notes in electrical engineering, vol 892. Springer, Singapore. https://doi.org/10.1007/978-981-19-1645-8_3
32. Bilro L, Alberto N, Pinto JL, Nogueira R (2012) Optical sensors based on plastic fibers. *Sensors (Basel)* 12(9):12184–12207. Epub 2012 Sep 5. PMID: 23112707; PMCID: PMC3478834. <https://doi.org/10.3390/s120912184>
33. Beyler CL, Hirschler MM (2002) Thermal decomposition of polymers. In: Philip J, DiNenno PE (eds) SFPE handbook of fire protection engineering, 3rd edn. National Fire Protection Association Quincy, Massachusetts, pp 1–111
34. Ashraf M, Mainuddin, Beg MT, Moin F, Rajesh R, Singhal G (2023) Sensitivity enhancement in U-shaped evanescent wave fiber sensor. *IEEE Sens J*. <https://doi.org/10.1109/JSEN.2023.3262864>
35. Sheeba M et al (2005) Fibre optic sensor for the detection of adulterant traces in coconut oil. *Meas Sci Technol* 16:2247–2250. <https://doi.org/10.1088/0957-0233/16/11/016>
36. Gupta BD, Dodeja H, Tomar AK (1996) Fibre-optic evanescent field absorption sensor based on a U-shaped probe. *Opt Quant Electron* 28:1629–1639. <https://doi.org/10.1007/BF00331053>

37. Leung A, Mohana Shankar P, Mutharasan R (2007) A review of fiber-optic sensors. *Sens Actuators, B Chem* 125:688–703. <https://doi.org/10.1016/j.snb.2007.03.010>
38. Korposh S, James SW, Lee SW, Tatam RP (2019) Tapered optical fibre sensors: current trends and future perspectives. *Sensors (Basel)* 19(10). <https://doi.org/10.3390/s19102294>
39. Sai VVR, Kundu T, Mukherji S (2009) Novel U-bent fiber optic probe for localized surface plasmon resonance-based biosensor. *Biosens Bioelectron* 24(9):2804–2809. <https://doi.org/10.1016/j.bios.2009.02.007>
40. Punjabi N, Satija J, Mukherji S (2015) Evanescent wave absorption based fiber-optic sensor—cascading of bend and tapered geometry for enhanced sensitivity. In: Mason A, Mukhopadhyay S, Jayasundara K (eds) *Sensing technology: current status and future trends III*. Smart Sensors, Measurement and Instrumentation, vol 11. Springer, Cham. https://doi.org/10.1007/978-3-319-10948-0_2
41. Gupta BD, Singh CD (1994) Fiber-optic evanescent field absorption sensor: A theoretical evaluation. *Fiber Integr Opt* 13(4):443. <https://doi.org/10.1080/01468039408202251>

Classification of DNA Sequence for Diabetes Mellitus Type Using Machine Learning Methods



Lena Abed AL Raheim Hamza, Hussein Attia Lafta,
and Sura Zaki Al Rashid

Abstract High blood sugar levels in diabetes mellitus (DM) can cause cardiac arrest, nervous system damage, vision loss, foot problems, liver or kidney damage, and death if left untreated. Age, gender, family history, BMI, and glucose levels all contribute to diabetes. To increase diabetes detection and prevent health concerns, machine learning techniques are used for prediction. Identifying the type of diabetes and considering the risk of accompanying diseases can improve diabetes prediction accuracy. This study uses one-way analysis of variance, mutual information, and F-regressor with random forest, Gaussian Naive Bayes, support vector machine, and decision tree for feature selection. Results with and without selected algorithms are compared. They have been used to adjust diabetic care using clinical parameters like accuracy, precision, recall, and F1-score. Random forest (RF) using F-regressor (FR) or ANOVA feature selection and numerous iterations of N (75) and K (3–5) outperforms competitors with 0.9 accuracy. This proves the diabetes-related DNA sequence classification technique works.

Keywords Diabetes mellitus · Machine learning · Deep learning · DNA sequence

1 Introduction

Diabetes often causes issues. Disorders affect toddlers to seniors. The FID estimated 451 diabetics globally in 2017 [1]. Studies suggest the number affected may double by 2045. To prevent and manage diabetes, predict onset [2]. Diabetes is

L. A. AL Raheim Hamza (✉)

College of Science for Women, University of Babylon, Babylon, Iraq

e-mail: lena.alrahiem.gsci112@student.uobabylon.edu.iq

H. A. Lafta · S. Z. Al Rashid

College of Information Technology, University of Babylon, Babylon, Iraq

e-mail: wsci.husein.attia@uobabylon.edu.iq

S. Z. Al Rashid

e-mail: sura_os@itnet.uobabylon.edu.iq

curable with early detection and treatment [3]. To link asthma to heredity, Bubby and Chrisman [4] established DNA-Senet. It was explored whether high-dimensional vector representations of DNA sequences around SNPs might relate SNPs to asthma, disclose new correlations, and forecast SNP-disease relationships for other difficult diseases. We linked DNA to asthma with an average AUC of 0.81. Two contextually semantic asthma loci and SNP-asthma relationships were found by DNA-Senet. The DNA-Senet linked CHD, type 2 diabetes, and rheumatoid arthritis sequencing to DNA anomalies. This method can find unique disease-associated sequences for various diseases. Sequential deep learning for bacterial categorization is suggested by Lugo and Barreto-Hernández [5]. A neural network identification model for bacterial whole-genome sequences is obtained from enormous next-generation sequencing data. After identification model validation, the bidirectional recurrent neural network beat other classification approaches.

High blood sugar causes diabetes. Diabetes shots may be needed lifelong if untreated. Heart, liver, renal, vision, and other issues may exist. Type 2 diabetes accounts for 90% of US cases. T2D, T1D, and gestational diabetes prevail (GDM) [6]. Prediabetes lowers glucose. Without diabetes, high glucose. This becomes Type 2 diabetes without early detection and treatment. A broad waist, low HDL, high blood pressure, high triglycerides, and uncontrolled blood sugar may induce metabolic syndrome. OGTT, FPG, A1C detect prediabetes [7]. Childhood type-1 diabetes (T1DM): Diabetes and insulin dependency result from autoimmune damage to insulin-producing cells. This disorder, which can harm the heart, blood flow, gums, nerves, pregnancy, eyes, and kidneys, is more common in young people [8]. Detecting type-1 autoantibodies in urine or blood (ketones) is important. Family and youth diabetes raise type 2 risk. Management of type 2 diabetes is common in seniors, insulin or reaction low [9]. Type 2 diabetes affects adults. Lack of exercise, obesity, neurological, and immunological problems are major causes. Other risks include ocular, neurological, renal, cardiac, and brain issues, the A1C, FPG, and RPG tests (random plasma glucose). Additionally, oral glucose tolerance and glucose challenge are assessed [10]. Standard second-trimester GDM: Low insulin causes gestational diabetes [11]. Obesity, PCOS, inactivity, and family history increase risk. Tolerance and glucose challenge tests are crucial for observing the health of the patient. Postpartum diabetes is from hypoglycemia, preeclampsia, and cesarean. It usually fades after birth but can induce type 2 diabetes if ignored [12]. AI ML: Trend-following computers forecast new data. ML models require data mining, optimization, and statistics. Data retrieval and knowledge representation optimization are automated [13]. ML predicts SVM, RF, XGBoost, LGBM, DT, GBM, Naive Bayes, logistic, linear regression [14, 15]. ML diabetes prediction research is common. Many studies use ML because it predicts DM best. The most common ailment, diabetes, is dangerous. 90% of diabetes is T2D. T2D's transcriptome and genetics were revealed by high-throughput sequencing. A single-cell sequencing breakthrough may reveal intricate disease-related cellular heterogeneity [16, 17]. Insulin resistance causes metabolic syndrome and T2D. IR-T2D interactions are unknown. DNA microarrays contain metagenes' situation-changing genes. After finding metagenes, Saxena et al. [18] created five gene expression profile machine learning models utilizing LASSO,

SVM, XGBoost, random forest, and ANN advanced AI neural network that cleared 95%. Diabetes was found by ML.

2 Related Works

Li et al. [19] examined Type 2 diabetes' molecular processes in multiple cell types. Over 1600 cells showed unique gene expression profiles in transcriptomics (949 from T2D patients and 651 from healthy individuals). By transcriptional characteristics, Monte Carlo feature selection, SVM, and incremental pruning separated single-cell data by group. Sneha and Gangil [20] detect diabetes early using effective feature selection. The DT and RF algorithms have 98.20 and 98.00% specificity, 82.30% Naive Bayes. Generalizing traits improves classification. SVM, RF, NB, DT, KNN, and two others are compared. Try rapid-miner. Dataset features are investigated. Most accurate are random forest and decision tree. SVM has 77.73 and 73.48%, while NB has 82.30% and SVM 77% in the suggested method. Improvements are sought through research. Taffa et al. [21] developed SVM-NB diabetes prediction. Three datasets trained the model. Kosovo got it. Data has eight attributes. The 402 patients included 80 type 2 diabetics. Its dietary and activity focus is unique. Data for training and testing is identical. Models using multiple algorithms succeed 97.6%. SVM is 95.52% accurate, Naive Bayes 94.52. Other ML techniques for matrix testing and evaluation will improve the model.

Kandhasamy and Balamurali [22] tested KNNs, SVMs, RFs, and J48 on UCI data archive specificity, sensitivity, and accuracy. Five-fold cross-validation categorized raw and preprocessed datasets. J48 decision tree classifier had 73.82% accuracy after preprocessing, while KNN ($k = 1$) and random forest had 100%. Zhang et al. [23] found unique SARS-CoV-2 transcriptome signatures as clinical models or vaccine targets in acute upper respiratory tissues using qualitative biomarkers and quantitative criteria. Least redundancy maximum relevance was used to evaluate selected attributes after Boruta transcriptomics. Incremental feature selection and classification developed qualitative biomarker gene and quantitative rule attribute lists. SVM was most accurate, then KNN, RF, DT. Organization of the paper: Sect. 2 covers machine learning-based DM diagnosis studies. Section 3 lists numerous machine learning algorithms for the specified architecture, while Sect. 4 evaluates dataset properties and Sect. 5 preprocessing before presenting the results. Section 6 covers execution and Sect. 7 concludes.

3 Proposed System

The proposed system utilizes machine learning methods for the diagnosis of diabetic diseases. In both cases, we first subject the data to some form of preprocessing, then use the four different classification techniques we've discussed, and lastly compare

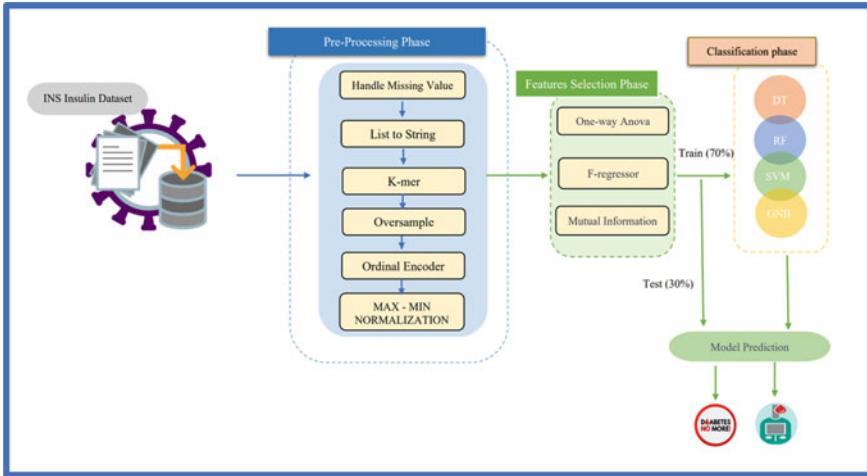


Fig. 1 Proposed system

the outcomes. Figure 1 is a schematic depiction of the proposed system, which provides even more detail.

4 Dataset

A famous health and biomedical research institute, the National Agency of Health, provided the INS Insulin Dataset. The dataset investigated diabetes mellitus insulin genetic variations. The extensive INS Insulin Dataset comprises DNA sequences and clinical data from T1D, T2D, and additional diabetic patients. Diabetics of many ethnicities make up this mix. 301-unit DNA from 14,571 persons is sampled. These sequences include 10,199 diabetes and 4371 non-diabetes [18].

5 Data Preprocessing

Effective preprocessing is needed to use machine learning to analyze DNA data for diabetes mellitus types. This study analyzes one-shot encoding, k-mer representation, min–max normalization, oversampling, and list-to-text conversion to improve algorithm efficiency and accuracy. This paper describes each technique's goals and methodology and its significance to diabetes mellitus disease study.

5.1 Handle Missing Values

Before diabetes mellitus, study machine learning DNA data analysis, missing values must be corrected. Deletion, imputation, or multiple imputations may be employed to maintain data integrity and quality for genetic pattern analysis and interpretation in diabetes mellitus. Successful management and handling improve reliability and contribute to key diabetic mellitus research discoveries utilizing machine learning models [19].

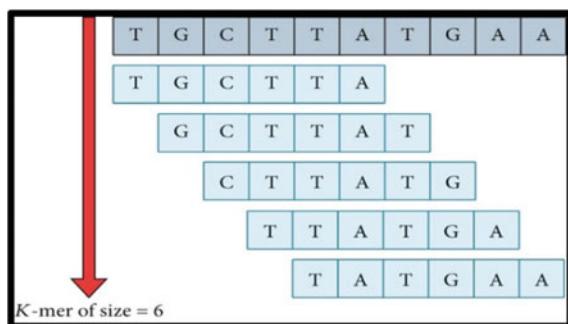
5.2 List to String

In DNA sequencing, genetic information is often expressed as a nucleotide list. Several machine learning methods need textual inputs. To fix this format conflict, convert lists to text. Through text preparation, numerous character-based approaches and algorithms can be used more efficiently [20].

5.3 K-mer

Analyzing long, complicated DNA sequences is difficult. To overcome this, the k-mer representation method splits the sequence into overlapping k-length subsequences with various patterns. By reducing dimensionality and collecting local patterns and motifs, machine learning algorithms can identify diabetic features [21]. A sequence's "k-mer" describes all K-length subsequences. Four monomers (A, G, A, and T), three dimers (AG, GA, and AT), two tetramers (AGA and GAT), and one tetramer is in AGAT (AGAT). Figure 2 shows 6-k-mer DNA sequence assembly.

Fig. 2 DNA sequencing with k-mer [22]



5.4 *Oversampling*

Genetic research commonly employs datasets with one class underrepresented. To balance classes, oversampling increases minority class samples. Random duplication, synthetic minority oversampling technique, or adaptive synthetic sampling provide minority classes with enough training data to prevent biased predictions and increase model performance. This strategy enhances machine learning algorithms' imbalanced scenario handling [23].

5.5 *Ordinal Encoding*

Ordinal encoding is a popular machine learning preprocessor. It is often used to process DNA for diabetes research. This converts categorical data like diabetes types into numerical representations while maintaining order. Ordinal encoding preserves variable order using unique integers. This makes representation and input feature handling easier for computer algorithms, facilitating classification exploration [24].

5.6 *Min–Max Normalization*

Normalization is essential for DNA data comparability. Min–max normalization, which rescales feature values to a range between 0 and 1, is a frequent normalization method.

$$X_{\text{norm}} = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (1)$$

where X is the original feature value, X_{norm} is the modified value, X_{\min} is the dataset minimum value, and X_{\max} is the dataset maximum value [25].

6 Feature Selection

Features must be selected for machine learning to categorize diabetes mellitus from DNA sequences [26]. Academic study compares one-way ANOVA, F-regressor, and mutual information feature selection. This approach finds type-specific diabetes genes, presenting principles, approaches, and their usefulness to this investigation.

6.1 ANOVA

One-way analysis of variance assesses group mean differences. One-way ANOVA analyzes DNA sequences to determine how each attribute or genetic variation affects the target variable, such as diabetes mellitus type. This method generates F-value and p-value metrics to determine if features and the outcome parameter are correlated. Next, traits with lower p-values and more importance are investigated and explained [27]. Equation gives the F-statistic formula.

$$F = \text{MSB}/\text{MSE} \quad (2)$$

where MSB: mean sum of squares between the groups and MSE: mean squares of errors.

6.2 F-Regressor

The F-regressor is a feature selection method based on ANOVA F-values. It examines the linear correlation between each feature and target variable independently and ranks them by F-value to discover strongly linked ones for diabetes mellitus type differentiation [28]. Higher values suggest stronger associations, making them the most significant features for extracting diabetes condition discrimination information.

6.3 Mutual Information

Mutual information evaluates statistical dependence. In feature selection, it evaluates target variable information. Complex DNA sequence dependencies benefit from its linear and nonlinear correlation detection. Mutual information values improve feature-target variable linkages, making them more useful and preferable for future analysis [29]. Think of x as a gene and y as a class. Use $H(x)$ and $H(y)$ from Eqs. (3) and (4) to calculate set entropy [19].

$$H(x) = \sum_{i \in x} P(x) \log p(x) \quad (3)$$

$$H(y) = \sum_{j \in x} P(y) \log(y) \quad (4)$$

Once we have calculated the entropies, mutual information can be calculated as:

$$\text{MI}(x, y) = H(x) + H(y) - H(x, y) \quad (5)$$

where x is the feature value and y is the target value.

7 Classification

Diabetes DNA analysis requires classification. This paper analyzes DNA sequence categorization using Gaussian NB, random forest, SVM, and decision tree [30]. The dataset is split into two subdatasets. The training dataset, with 80% of the original dataset. And testing dataset, with 20% of the original dataset.

7.1 Random Forest

Ensemble learning predicts using several decision trees. This approach trains decision trees on random dataset samples to create a forest. It handles high-dimensional data well, considering complex relationships and preventing overfitting. Multiple tree's joint judgments help random forest to classify DNA sequences in diabetes study [31]. This random forest design uses 100 trees.

7.2 Gaussian NB

GNB is a probabilistic ML classification algorithm using Gaussian distribution. Gaussian Naive Bayes assumes each feature or predictor can predict the output. Final prediction of all components yields group dependent variable categorization probability. Equation (6) assumes Gaussian feature probabilities:

$$p(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right) \quad (6)$$

These formulas determine the variance and mean of continuous variable x for each y class using σ and μ [32, 33].

7.3 Support Vector Machine

An effective supervised learning approach for classification. SVMs seek an ideal hyperplane that maximally separates data points by category. SVM can handle linear

and nonlinear separable data using kernel functions [34, 35]. RBF kernel calculation uses this equation:

$$k(x_i + x_j) = \exp\left(\gamma \|x_i + x_j\|^2\right) \quad (7)$$

A learnable parameter, gamma (γ), represents the RBF kernel in this scenario. SVM can classify complex DNA sequence patterns and correlations for diabetes mellitus research because it performs well with high-dimensional datasets and has robust generalization [36, 37].

7.4 Decision Tree

Machine learning decision trees analyze data for rules. Internal nodes make feature-based decisions and leaves label and forecast classes. It gives discrete and continuous variables feature importance. Reliable genetic differences are detected by DNA sequence analysis to classify diabetes, calculate performance matrices, and optimize predictive model parameters [38].

8 Results and Discussion

Analyzing DNA genes, dataset size, and feature representation affect diabetes classifiers. Science analyzes algorithms using accuracy, precision, recall, and F1-score model classification precision calculation. Comparing actual and expected labels assessed classifier performance. Comparing true positives to irrelevant, erroneous, and unclassified cases evaluates accuracy. This algorithm should consider dataset properties and research goals. Data prep cuts redundancy and inconsistency. This study examined diabetic DNA using machine learning, mutual information feature selection, ANOVA, and f-regressor. 75–20 feature sizes and 3–8 K-mer values were examined. A split 80/20 train-test assessed model accuracy. DNA data analytics and machine learning identified diabetes subtypes. Table 1 gives the accuracy, precision, recall, and F1-score for each machine learning model using different feature selection procedures.

Random forest has good accuracy values of 0.89 for one-way ANOVA, 0.88 for f-regressor testing, and 0.87 for mutual information technique application, despite competitive precision/recall/F1-score values. Gaussian NB performed worse than random forest in Table 2, only showing decent measurements during quick ANOVA trials and overall average scores with reasonable precise utility among test cases, making it suitable for many single hypothesis problems but lessening control where more dimensions exist, which can increase prediction errors or f. SVM and decision tree feature selection improved our classification algorithms tremendously. Each

Table 1 Comprehensive overview of performance metrics with $N = 20$

Methods	FR			FOA			MI					
	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score
<i>Feature selection with N = 20, K = 8</i>												
RF	0.88	0.93	0.85	0.88	0.87	0.9	0.85	0.88	0.87	0.9	0.84	0.87
GNB	0.69	0.69	0.69	0.69	0.66	0.65	0.67	0.66	0.69	0.67	0.7	0.68
SVM	0.78	0.8	0.77	0.78	0.76	0.78	0.75	0.77	0.78	0.79	0.77	0.78
DT	0.82	0.9	0.77	0.83	0.82	0.89	0.78	0.83	0.81	0.89	0.77	0.83
<i>Feature selection with N = 20, K = 7</i>												
RF	0.88	0.92	0.86	0.89	0.88	0.92	0.84	0.88	0.87	0.92	0.84	0.88
GNB	0.7	0.67	0.72	0.7	0.69	0.68	0.69	0.69	0.69	0.66	0.71	0.68
SVM	0.78	0.79	0.79	0.79	0.78	0.79	0.77	0.78	0.78	0.76	0.78	0.77
DT	0.81	0.9	0.77	0.83	0.81	0.9	0.76	0.82	0.82	0.91	0.77	0.83
<i>Feature selection with N = 20, K = 6</i>												
RF	0.87	0.91	0.84	0.88	0.87	0.9	0.85	0.88	0.87	0.91	0.85	0.88
GNB	0.68	0.66	0.68	0.67	0.68	0.64	0.7	0.67	0.71	0.69	0.72	0.7
SVM	0.77	0.77	0.77	0.77	0.77	0.77	0.78	0.77	0.77	0.8	0.8	0.8
DT	0.81	0.9	0.76	0.82	0.82	0.89	0.77	0.83	0.82	0.9	0.78	0.84
<i>Feature selection with N = 20, K = 5</i>												
RF	0.88	0.9	0.86	0.88	0.88	0.9	0.86	0.88	0.87	0.9	0.84	0.87
GNB	0.7	0.68	0.7	0.69	0.68	0.66	0.69	0.67	0.7	0.65	0.72	0.68
SVM	0.79	0.79	0.78	0.78	0.76	0.77	0.77	0.77	0.78	0.77	0.79	0.78
DT	0.81	0.89	0.76	0.82	0.82	0.89	0.78	0.83	0.8	0.89	0.76	0.82
<i>Feature selection with N = 20, K = 4</i>												

(continued)

Table 1 (continued)

Methods	FR			FOA			MI		
	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score	ACC
RF	0.88	0.93	0.85	0.89	0.89	0.92	0.86	0.89	0.87
GNB	0.69	0.67	0.71	0.69	0.69	0.66	0.69	0.67	0.7
SVM	0.79	0.8	0.78	0.79	0.79	0.79	0.78	0.78	0.79
DT	0.82	0.9	0.78	0.84	0.82	0.92	0.77	0.83	0.9
<i>Feature selection with N = 20, K = 3</i>									
RF	0.88	0.91	0.85	0.88	0.88	0.9	0.87	0.89	0.88
GNB	0.69	0.67	0.68	0.68	0.68	0.64	0.71	0.67	0.69
SVM	0.78	0.79	0.77	0.78	0.78	0.79	0.79	0.78	0.78
DT	0.81	0.9	0.76	0.82	0.82	0.9	0.78	0.84	0.81

model was assessed using one-way ANOVA, f-regressor, and mutual information. One-way ANOVA, f-regressor, and mutual information SVM accuracy were 0.84, 0.84, and 0.83. SVM, like random forest, exhibited continuous recall and F1-score over feature selection. Decision tree had an accuracy rating of 0.80–0.83 on one-way ANOVA, good precision values, and recall determined by F1-score, but lower than random forest and SVM classifier algorithm through the mutual information scheme. Combining many classifiers in deep neural networks or ensemble techniques improves prediction reliability and academic progress.

9 Conclusion

This study critically evaluated and compared DNA sequence data classification techniques random forest, Gaussian NB, SVM, and decision tree for diabetes mellitus type detection. The integration of one-way ANOVA, F-regressor, and mutual information helped reveal genetic features that inform classification models. Performance indicators, like accuracy scores, precision recall rates, and F1 scores were essential for assessing each algorithm's capabilities. This study has important significance for current research by giving dependable data to help machine learning model selection based on DNA analysis for correct categorization of diabetes mellitus types with high confidence. Data analysis algorithms that can accurately interpret vast volumes of diabetic data without human bias were examined in this study. This disease's stages were diagnosed and forecasted using machine learning due to its intricacy. The recommended solution employing random forest with feature selection consistently achieved 90% accuracy for various feature selection methods (ANOVA or F-regressor), N (50 and 75), and K. (3–5). These findings help researchers choose machine learning models and effectively categorize diabetes mellitus types using DNA sequence analysis.

Table 2 Comprehensive overview of performance metrics with N = 75

Methods	FR			FOA			MI					
	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score
<i>Feature selection with N = 75, K = 8</i>												
RF	0.88	0.9	0.87	0.89	0.89	0.9	0.88	0.89	0.89	0.88	0.89	0.89
GNB	0.71	0.69	0.72	0.71	0.71	0.69	0.72	0.7	0.7	0.68	0.73	0.7
SVM	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.83	0.82	0.84	0.83
DT	0.82	0.91	0.78	0.84	0.82	0.89	0.77	0.83	0.8	0.89	0.77	0.82
<i>Feature selection with N = 75, K = 7</i>												
RF	0.89	0.9	0.88	0.89	0.89	0.89	0.89	0.89	0.88	0.9	0.87	0.88
GNB	0.7	0.68	0.72	0.7	0.71	0.69	0.72	0.71	0.71	0.7	0.7	0.7
SVM	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.82	0.83
DT	0.81	0.9	0.77	0.83	0.81	0.9	0.76	0.82	0.8	0.88	0.76	0.81
<i>Feature selection with N = 75, K = 6</i>												
RF	0.89	0.9	0.88	0.89	0.89	0.89	0.9	0.89	0.89	0.9	0.88	0.89
GNB	0.7	0.67	0.71	0.69	0.71	0.7	0.73	0.71	0.7	0.69	0.71	0.7
SVM	0.82	0.82	0.82	0.82	0.84	0.84	0.85	0.84	0.83	0.83	0.83	0.83
DT	0.82	0.9	0.77	0.83	0.83	0.9	0.79	0.84	0.82	0.9	0.77	0.83
<i>Feature selection with N = 75, K = 5</i>												
RF	0.89	0.89	0.89	0.89	0.9	0.9	0.89	0.9	0.88	0.9	0.87	0.88
GNB	0.72	0.69	0.72	0.7	0.71	0.7	0.73	0.71	0.71	0.7	0.7	0.7
SVM	0.84	0.83	0.84	0.84	0.84	0.84	0.83	0.84	0.83	0.83	0.82	0.82
DT	0.82	0.89	0.77	0.83	0.82	0.92	0.77	0.84	0.81	0.9	0.76	0.82
<i>Feature selection with N = 75, K = 4</i>												

(continued)

Table 2 (continued)

Methods	FR			FOA			MI		
	ACC	Pre	Recall	F1-score	ACC	Pre	Recall	F1-score	ACC
RF	0.9	0.91	0.89	0.9	0.9	0.89	0.91	0.9	0.89
GNB	0.71	0.68	0.73	0.7	0.68	0.71	0.69	0.7	0.68
SVM	0.84	0.84	0.83	0.84	0.83	0.83	0.83	0.84	0.84
DT	0.82	0.91	0.78	0.84	0.82	0.89	0.77	0.83	0.77
<i>Feature selection with N = 75, K = 3</i>									
RF	0.9	0.89	0.91	0.9	0.9	0.87	0.92	0.89	0.9
GNB	0.7	0.68	0.71	0.69	0.71	0.69	0.72	0.7	0.71
SVM	0.83	0.83	0.83	0.83	0.84	0.84	0.83	0.84	0.83
DT	0.8	0.89	0.75	0.82	0.81	0.91	0.76	0.83	0.76

References

1. Li X, Zhang J, Safara F (2023) Improving the accuracy of diabetes diagnosis applications through a hybrid feature selection algorithm. *Neural Process Lett* 55(1). <https://doi.org/10.1007/s11063-021-10491-0>
2. Arora A, Shoeibi N, Sati V, González-Briones A, Chamoso P, Corchado E (2021) Data augmentation using Gaussian mixture model on csv files. *Adv Intell Syst Comput*. https://doi.org/10.1007/978-3-030-53036-5_28
3. Mirza S, Mittal S, Zaman M (2018) Decision support predictive model for prognosis of diabetes using SMOTE and decision tree
4. Bubby S, Chrisman B (2021) DNA-SEnet: a convolutional neural network for classifying DNA-asthma associations. *J Emerg Investig* 4
5. Lugo L, Hernández EB (2021) A recurrent neural network approach for whole genome bacteria identification. *Appl Artif Intell* 35(9):642–656. <https://doi.org/10.1080/08839514.2021.1922842>
6. Chaki J, Thillai Ganesh S, Cidham SK, Ananda Theertan S (2022) Machine learning and artificial intelligence based diabetes mellitus detection and self-management: a systematic review. *J King Saud Univ Comput Inf Sci* 34(6). <https://doi.org/10.1016/j.jksuci.2020.06.013>
7. Fregoso-Aparicio L, Noguez J, Montesinos L, García-García JA (2021) Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology Metab Syn* 13(1). <https://doi.org/10.1186/s13098-021-00767-9>
8. Ramadhan NG, Adiwijaya, Romadhony A (2021) Preprocessing handling to enhance detection of type 2 diabetes mellitus based on random forest. *Int J Adv Comput Sci Appl* 12(7). <https://doi.org/10.14569/IJACSA.2021.0120726>
9. Naz H, Ahuja S (2020) Deep learning approach for diabetes prediction using PIMA Indian dataset. *J Diab Metab Disord* 19(1). <https://doi.org/10.1007/s40200-020-00520-5>
10. Butt UM, Letchmunan S, Ali M, Hassan FH, Baqir A, Sherazi HHR (2021) Machine learning based diabetes classification and prediction for healthcare applications. *J Healthc Eng* 2021. <https://doi.org/10.1155/2021/9930985>
11. Deng Y et al (2021) Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients. *NPJ Digit Med* 4(1). <https://doi.org/10.1038/s41746-021-00480-x>
12. Kaur H, Pannu HS, Malhi AK (2019) A systematic review on imbalanced data challenges in machine learning: applications and solutions. *ACM Comput Surv* 52(4). <https://doi.org/10.1145/3343440>
13. Zhuang F et al (2021) A comprehensive survey on transfer learning. *Proc IEEE* 109(1). <https://doi.org/10.1109/JPROC.2020.3004555>
14. Li K, Daniels J, Liu C, Herrero P, Georgiou P (2020) Convolutional recurrent neural networks for glucose prediction. *IEEE J Biomed Health Inform* 24(2). <https://doi.org/10.1109/JBHI.2019.2908488>
15. Recurrent neural network and convolutional network for diabetes blood glucose prediction. *Int J Mach Learn Comput* 12(6). <https://doi.org/10.18178/ijmlc.2022.12.6.1115>
16. Tasin I, Nabil TU, Islam S, Khan R (2022) Diabetes prediction using machine learning and explainable AI techniques. *Healthc Technol Lett*. <https://doi.org/10.1049/htl2.12039>
17. Al-Bermary HM, Al-Rashid SZ (2021) Microarray gene expression data for detection Alzheimer's disease using k-means and deep learning. In: Proceedings of the 7th International engineering conference "research and innovation amid global pandemic", IEC 2021. <https://doi.org/10.1109/IEC52205.2021.9476128>
18. National Library of Medicine. <https://www.ncbi.nlm.nih.gov/>. Accessed 14 May 2023
19. Es-Sabery F et al (2021) A MapReduce opinion mining for COVID-19-related tweets classification using enhanced ID3 decision tree classifier. *IEEE Access* 9. <https://doi.org/10.1109/ACCESS.2021.3073215>

20. Valsalan P, Hasan NU, Farooq U, Zghaibeh M, Baig I (2023) IoT based expert system for diabetes diagnosis and insulin dosage calculation. *Healthcare (Switzerland)* 11(1). <https://doi.org/10.3390/healthcare11010012>
21. Ye H, Tang S, Yang C (2021) Deep learning for chlorophyll-a concentration retrieval: a case study for the pearl river estuary. *Remote Sens (Basel)* 13(18). <https://doi.org/10.3390/rs13183717>
22. Gunasekaran H, Ramalakshmi K, Rex Macedo Arokiajraj A, Kanmani SD, Venkatesan C, Dhas CSG (2021) Analysis of DNA sequence classification using CNN and hybrid models. *Comput Math Methods Med* 2021. <https://doi.org/10.1155/2021/1835056>
23. Ibraheem EMA, El-sisy AME (2019) Comparing the effect of three denture adhesives on the retention of mandibular complete dentures for diabetic patients (randomized clinical trial). *Bull Natl Res Cent* 43(1). <https://doi.org/10.1186/s42269-019-0052-7>
24. Kabakuş AT (2020) The data science met with the COVID-19: revealing the most critical measures taken for the COVID-19 pandemic. *Sakarya Univ J Comput Inf Sci*. <https://doi.org/10.35377/saucis.03.03.771501>
25. Asfaw TA (2019) Prediction of diabetes mellitus using machine learning techniques. *Int J Comput Eng Technol* 10(4). <https://doi.org/10.34218/ijcet.10.4.2019.004>
26. Ahn CH, Lee S, Song HM, Park JR, Joo JC (2019) Assessment of water quality and thermal stress for an artificial fish shelter in an urban small pond during early summer. *Water (Switzerland)* 11(1). <https://doi.org/10.3390/w11010139>
27. Al-Sarem M et al (2021) An improved multiple features and machine learning-based approach for detecting clickbait news on social networks. *Appl Sci (Switzerland)* 11(20). <https://doi.org/10.3390/app11209487>
28. Kim SK, Yeun CY, Yoo PD (2019) An enhanced machine learning-based biometric authentication system using RR-interval framed electrocardiograms. *IEEE Access* 7. <https://doi.org/10.1109/ACCESS.2019.2954576>
29. Xuegang L, Junrui L, Juan W (2021) Missing data reconstruction based on spectral k-support norm minimization for NB-IoT data. *Math Probl Eng* 2021. <https://doi.org/10.1155/2021/1336900>
30. Aminah R, Saputro AH (2019) Diabetes prediction system based on iridology using machine learning. In: 2019 6th International conference on information technology, computer and electrical engineering, ICITACEE 2019. <https://doi.org/10.1109/ICITACEE.2019.8904125>
31. Rani A, Kumar N, Kumar J, Sinha NK (2022) Machine learning for soil moisture assessment. *Deep Learn Sustain Agric*. <https://doi.org/10.1016/B978-0-323-85214-2.00001-X>
32. Vishwakarma DK, Dhiman C (2019) A unified model for human activity recognition using spatial distribution of gradients and difference of Gaussian kernel. *Vis Comput* 35(11). <https://doi.org/10.1007/s00371-018-1560-4>
33. Raizada RDS, Lee YS (2013) Smoothness without smoothing: why Gaussian Naive Bayes is not naive for multi-subject searchlight studies. *PLoS ONE* 8(7). <https://doi.org/10.1371/journal.pone.0069566>
34. Barman M, Dev Choudhury NB (2020) A similarity based hybrid GWO-SVM method of power system load forecasting for regional special event days in anomalous load situations in Assam, India. *Sustain Cities Soc* 61. <https://doi.org/10.1016/j.scs.2020.102311>
35. Alimjan G, Sun T, Liang Y, Jumahun H, Guan Y (2018) A new technique for remote sensing image classification based on combinatorial algorithm of SVM and KNN. *Intern J Pattern Recogn Artif Intell* 32(7):1–23. <https://doi.org/10.1142/S0218001418590127>
36. Aziz FA, Al-Rashid SZ (2022) Prediction of DNA binding sites bound to specific transcription factors by the SVM algorithm. *Iraqi J Sci* 63(11). <https://doi.org/10.24996/ijss.2022.63.11.37>
37. Muzzammel R, Raza A (2020) A support vector machine learning-based protection technique for MT-HVDC systems. *Energies (Basel)* 13(24). <https://doi.org/10.3390/en13246668>
38. Hafeez MA, Rashid M, Tariq H, Abideen ZU, Alotaibi SS, Sinky MH (2021) Performance improvement of decision tree: a robust classifier using Tabu search algorithm. *Appl Sci (Switzerland)* 11(15). <https://doi.org/10.3390/app11156728>

Unveiling the Future: A Review of Financial Fraud Detection Using Artificial Intelligence Techniques



Sankalp Goel and Abha Kiran Rajpoot

Abstract Financial fraud is the illegal use of mobile platforms for transactions when credit card or identity theft is exploited to create fake money. With the spread of smartphones and online transaction services, financial fraud and credit card fraud have become rapidly growing issues. Accurate detection of financial fraud in this context is crucial, as it can result in significant financial losses. Therefore, we conducted a survey of machine learning, deep learning, and data mining methodologies for financial fraud detection. In our study, we evaluated our methodology for detecting fraud and handling vast financial data, contrasting it with artificial neural networks. Our reviewed process encompassed variable selection, sampling, and the utilisation of supervised and unsupervised algorithms. This allowed us to effectively identify financial fraud and process extensive financial datasets.

Keywords Artificial intelligence · Financial fraud · Machine learning · Neural networks

1 Introduction

Businesses, online services, and internet users have all grown significantly in recent years. To make their lives simpler and easier, people today rely on internet banking systems for money transfers, utilise debit and credit cards for purchasing, and use online bill payment services. Despite the positive aspects of online transactions, financial establishments and customers face substantial losses from fraudulent activities such as Illegitimate borrowing transactions, deceptive use of credit cards, fraudulent attempts to obtain personal information, creating counterfeit documents, dishonest creation of unauthorised accounts, and illicit activities targeting online banking systems. These fraud crimes result in significant financial losses and affect customer confidence and the financial stability of establishments. The fact that fraudulent

S. Goel · A. K. Rajpoot (✉)

Department of Computer Science Engineering, Sharda University, Greater Noida, India
e-mail: abhakiran.rajpoot@sharda.ac.in

activity is continually changing is one of the biggest obstacles to the identification of financial crime. Financial institutions struggle to keep up with the evolving fraud environment as fraudsters continuously develop new strategies and methods to get around detection systems. This requires constant updates and improvements to fraud detection systems to stay effective.

Additionally, there is a need for advanced artificial intelligence (AI) methodologies to improve the accuracy of fraud detection. Traditional rule-based approaches may not be effective in detecting complex and subtle fraud patterns and may also generate false positives, leading to unnecessary investigations and increased operational costs. Financial institutions may create more complex and adaptable fraud detection models that can continually learn from fresh data and react to changing fraud trends by utilising machine learning and other AI approaches. Furthermore, the integration of big data presents both opportunities and challenges in financial fraud detection. Big data can provide valuable insights and patterns that can help to detect fraud, but it also presents challenges in terms of processing and analysing large volumes of data in real-time. This requires advanced data processing and analysis techniques that can handle the velocity, variety, and volume of big data to enable timely and effective fraud detection.

FDS is a complex and evolving field that requires continuous updates and improvements to address the challenges posed by changing fraudulent behaviours, a lack of comprehensive transaction information tracking methods, and the need for advanced AI methodologies. By leveraging data mining, machine learning, big data, and other advanced technologies, financial institutions can enhance their fraud detection capabilities and better protect themselves and their clients against financial fraud.

2 Literature Review

2.1 *Machine Learning Techniques for Financial Fraud Detection*

Various machine learning algorithms have been suggested in numerous research as a means of detecting financial fraud. These techniques include logistic regression, isolation forests, decision trees, neural networks, and ensemble methods. For instance, Liu et al. [1] utilised decision trees and Support Vector Machines, among other machine learning techniques, in their study to spot false financial statements. Decision trees performed better than other methods in terms of accuracy, according to the results.

2.2 Deep Learning for Financial Fraud Detection

Zhang et al. [2] put forward a deep learning-based model that combined CNNs and RNNs to detect credit card fraud. Their model achieved high accuracy and recall rates, outperforming other traditional methods. The use of deep learning techniques, such as Gradient Boosting Machine and Support Vector Machines, has produced promising results in the identification of financial fraud.

2.3 Ensemble Methods for Financial Fraud Detection

Ensemble methods, which combine multiple models to make predictions, have gained popularity in financial fraud detection. Studies have shown that ensemble methods, such as bagging and boosting, can enhance the accuracy and robustness of fraud detection models. For example, Li et al. (2019) proposed an ensemble approach that combined multiple classifiers, including decision trees, isolation forests, and hidden Markov models, to detect online financial fraud. Their results showed that the ensemble approach achieved higher accuracy compared to individual classifiers.

2.4 Unsupervised and Semi-supervised Learning for Financial Fraud Detection

Unsupervised and semi-supervised learning techniques have been explored for financial fraud detection, where labelled data may be scarce or expensive to obtain. For instance, Zhang et al. [2] proposed an unsupervised learning-based method that utilised self-organising maps and clustering techniques to detect fraudulent transactions in credit card data. Their results showed that the unsupervised learning approach achieved competitive performance even without labelled data.

2.5 Explainable AI for Financial Fraud Detection

Since it offers clear and comprehensible insights into the decision-making process of AI models, explainable AI (XAI) has drawn interest in recent years. For the purpose of detecting financial fraud, a number of studies have looked at the use of XAI approaches such as rule-based models and interpretable machine learning algorithms. For example, Dimitrios [3] proposed an XAI-based approach that combined a rule-based model with a deep learning model to detect insider trading. Their results demonstrated that the XAI approach provided explainable and accurate fraud detection results.

2.6 Feature Selection and Feature Engineering

Financial fraud may be detected using AI approaches, although feature engineering and selection are essential. In order to increase the precision of fraud detection models, several studies have concentrated on discovering pertinent features and devising efficient feature engineering procedures. For instance, Wu et al. (2020) suggested a feature selection method based on mutual knowledge and a feature engineering technique that used temporal information to detect insider trading. In accordance with their findings, the fraud detection model's performance was greatly enhanced by feature selection and engineering.

3 Models and Methodologies

3.1 FDS of Bayesian Learning and Dempster–Shafer Theory

For the purpose of detecting credit card fraud, a fusion strategy employing Bayesian learning and Dempster–Shafer theory is suggested [4]. Four elements make up the fraud detection system (FDS): a rule-based filter, a Dempster–Shafer algorithm, a transaction history database, and a Bayesian learner [5]. The rule-based filter analyses incoming data according to predetermined criteria to spot probable fraud trends. To determine the overall chance of fraud, the Dempster–Shafer adder aggregates data from many sources. Past transaction data is kept in the transaction history database for study and reference. For better fraud detection, the Bayesian learner continually updates its knowledge based on fresh information. Together, these parts make up a complete system that employs a range of strategies to identify and stop fraudulent activity.

The rule-based component assigns the suspicion level of each incoming transaction depending on how much it deviates from a desirable pattern. Then, using a combination of many pieces of evidence, the DS theory generates an initial belief value [6]. The initial belief values are then further combined using the DS theory to provide an overall belief that labels the transaction as either legitimate or suspect. Bayesian learning is used to increase or reduce the hypothesis that a transaction is suspicious depending on how similar it is to previous fraudulent or legitimate transactions. This method has proven to be highly accurate and quick, which has increased detection rates and decreased false alarms. However, it should be noted that this approach may be expensive and may have a lower processing speed compared to other methods (Fig. 1).

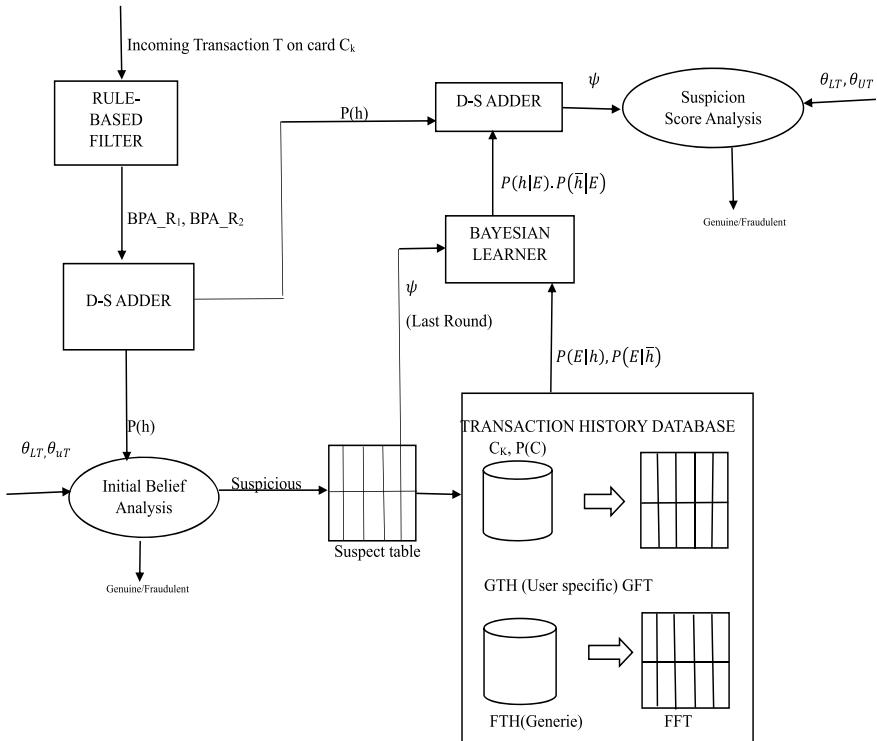


Fig. 1 Block diagram of DST and BL

3.2 The Evolutionary-Fuzzy System

By using genetic algorithms and fuzzy logic to generate a set of rules that can accurately distinguish fraudulent and non-fraudulent transactions, evolutionary fuzzy systems can be utilised in the detection of fraud. These methods create a collection of fuzzy rules that characterise the traits of fraudulent and non-fraudulent transactions using past transaction data. The Fuzzy Darwinian Detection system mentioned in the original text is an illustration of an evolutionary fuzzy system used in fraud detection [7]. The system creates a collection of variable-length fuzzy rules that capture the differences between several kinds of data using a combination of fuzzy logic and genetic programming. The algorithm, which was created expressly for detecting insurance fraud, classifies data as “safe” or “suspicious” depending on whether client payments are past due or how many payments are past due and longer than three months. The use of evolutionary fuzzy systems in fraud detection can have several benefits, including high accuracy in detecting fraudulent transactions, adaptability to novel fraud patterns, and proficiency with huge, complicated datasets. However, they may also be slow processors and expensive to implement (Fig. 2).

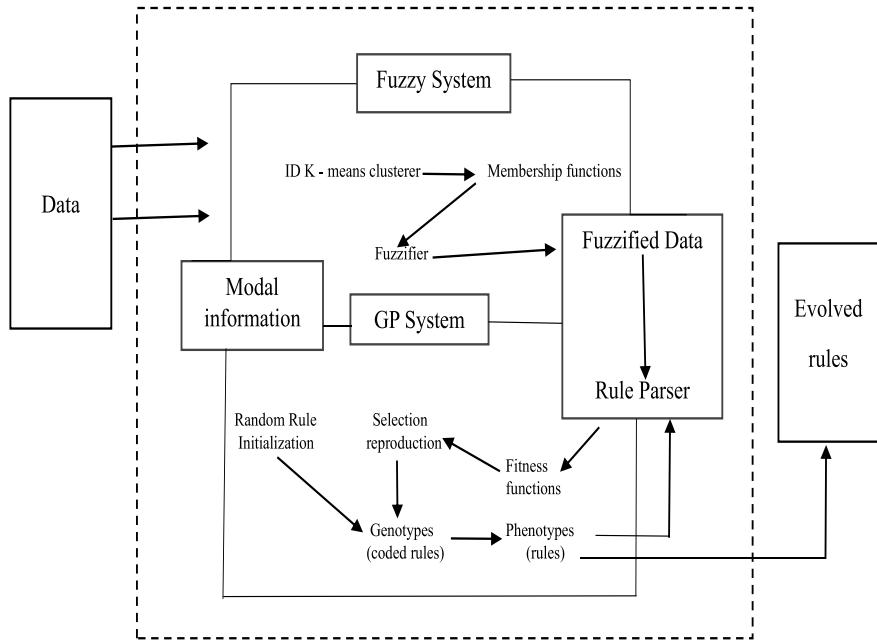


Fig. 2 Block diagram of the EFS

3.3 Deep Artificial Neural Networks

A class of machine learning models known as deep artificial neural networks is based on the anatomy and operation of the human intellect. By using unsupervised learning approaches to find patterns even in unstructured or unlabelled data, these models can analyse and understand the data. Deep learning is built on artificial neural networks (ANNs), commonly referred to as neural networks or multilayer perceptrons [8]. A single-neuron model called a perceptron existed before more complex neural networks. Because of their hierarchical or multilayered structure, neural networks are able to anticipate the future [9]. A multilayer perceptron typically consists of one or more intermediate layers, called concealed layers, between the input and output layers [10]. The connections between the input layer, concealed layers, and output layer, all of which point in the direction of the intake layer, create an anticipatory network. Each layer's lack of an unmediated link from the production layer to the intake layer identifies it. Backpropagation learning algorithms are commonly used to train most multilayer perceptrons. These algorithms update the weights and biases of the neural network during training by minimising the error between the predicted output and the actual output based on the labelled training data. Deep artificial neural networks have been widely used in various financial fraud detection methods due to their ability to learn complex patterns from large amounts of data and achieve state-of-the-art performance in many tasks.

3.4 BLAST–SSAHA Hybridization

The technique of financial fraud detection known as BLAST–SSAHA hybridization employs a fusion of the Basic Local Alignment Search Tool (BLAST) and the Sequence Search and Alignment by Hashing Algorithm (SSAHA) principles [11]. BLAST is a widely used tool in bioinformatics for comparing DNA or protein sequences, while SSAHA is a sequence alignment algorithm used for rapid comparison of DNA sequences [12]. In the context of financial fraud detection, BLAST–SSAHA hybridization can be applied to analyse patterns in financial transaction data, such as credit card transactions or wire transfers, to identify potentially fraudulent activities [13]. Here is how it works (Fig. 3).

Sequence Generation. Financial transaction data can be represented as a series of characters or numbers and includes information such as transaction amounts, kinds, timestamps, and relevant properties. The crucial information required for analysis, processing, and storage is included in these sequences. This data is organised and made easily available through encoding for a variety of uses, such as transactional analysis, fraud detection, and financial forecasts. This kind of transaction encoding makes it possible to manipulate and use the data effectively, supporting insightful financial decision-making.

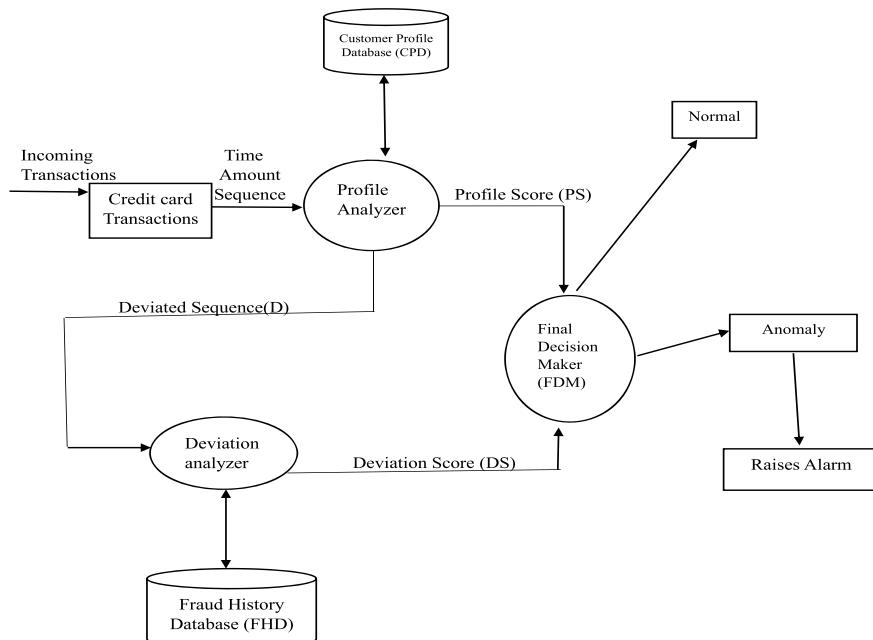


Fig. 3 Architecture of the BLAST and SAHAA fraud detection systems

Database Creation. A database can be built to record recognised patterns of fraudulent transactions by using previous data or other trustworthy sources. By cross-referencing incoming transactions, this invaluable resource allows the detection of suspected fraudulent operations. Organisations may improve their capacity to spot and stop fraudulent acts, protecting their systems and assets, by analysing new transactions against the recognised patterns of fraud.

BLAST Alignment. The sequences generated from the financial transaction data can be aligned against the sequences in the fraud pattern database using the BLAST algorithm [14]. Transactional data frequently includes trends or themes related to fraud. BLAST may identify and locate well-known fraud patterns by analysing the data. Its capacity to spot these themes aids in the detection of probable fraudulent actions and the risk mitigation associated with them.

SSAHA Hashing. The sequences may be processed with the SSAHA method to produce a hash table of fragmented sequences and financial transaction data. This table provides effective comparison and comparable sequence recognition, assisting in the discovery of probable matches to recognised fraud patterns. Financial organisations are better able to identify fraudulent activity and take the necessary action when they use SSAHA hashing to quickly analyse massive amounts of data.

Hybridization. Utilising the advantages of both techniques, combining the findings of SSAHA hashing with BLAST alignment improves fraud detection. With the help of the hybrid technique, possible matches between financial transaction data and the fraud pattern database are successfully found, increasing the efficacy and accuracy of fraud detection. This method offers an all-inclusive approach for identifying fraudulent activity by combining the benefits of BLAST and SSAHA.

Fraud Detection. Based on the potential matches identified through the hybridization approach, a fraud detection system can generate alerts or trigger further investigation to identify and prevent financial fraud. It is important to note that the effectiveness of BLAST–SSAHA hybridization in financial fraud detection would depend on the quality and size of the fraud pattern database, the accuracy of the sequence generation process, and other factors related to the specific use case and data quality [15]. Additionally, financial fraud detection typically involves a multilayered approach that combines various techniques and strategies, and BLAST–SSAHA hybridization can be a part of a larger fraud detection system.

3.5 Decision Tree

A well-liked machine learning technique for categorization problems, such as detecting financial fraud, is the decision tree. It has a tree-like structure where each leaf node represents a result or a class label (such as “fraud” or “non-fraud”) and each inside node reflects a judgement based on a characteristic (e.g. decline or approve). Decision trees are helpful for finding patterns and making judgements in complicated

data because they are simple to understand and intuitive [16]. Here are the steps to building a decision tree model for financial fraud detection:

Data Preparation. The process of data analysis must start with data preparation. It entails gathering the data and carrying out preprocessing operations. This includes cleaning the data by eliminating mistakes or inconsistencies, addressing missing values using strategies like imputation, and converting categorical variables to numerical representations for analysis. These procedures guarantee that the data is in an appropriate format for additional analysis and modelling.

Feature Selection. Identify the relevant features (also known as predictors or attributes) that may affect fraud occurrence. This can be done using domain knowledge or feature selection techniques such as information gain or Gini impurity.

Data Split. To create training and test datasets, divide the data. The decision tree model will be trained using the training dataset, and its performance will be tested using the testing dataset.

Decision Tree Construction. To create the decision tree model, use the training dataset. The characteristic that best categorises the data into various classifications (fraud or non-fraud) will be used by the algorithm to recursively partition the data at each node [17]. This is done by maximising a criterion, such as information gain or Gini impurity, which measures the purity of the class labels in each branch.

Model Evaluation. Utilising standard evaluation criteria for classification tasks, it is possible to evaluate the decision tree model's performance on the testing dataset. These measurements include F1-score, recall, accuracy, and precision.

Model Optimisation. Fine-tune the decision tree model by pruning or limiting the tree size to prevent overfitting, which occurs when the model is too complex and may not generalise well to new data. Model interpretation environment for real-time financial fraud detection. Monitor its performance. Interpret the decision tree to gain insights into the patterns and rules learned by the model. This can help to understand the decision-making process and provide explanations for fraud detection decisions.

Model Deployment. Implement the trained and fine-tuned decision tree model in a production environment, regularly enhancing and modifying the model as required to ensure precise and efficient identification of financial fraud.

4 Conclusion

The latest machine learning and artificial intelligence techniques used for FDS are reviewed in this study. The machine learning strategy featured assembling, categorization, and a characteristic nomination procedure based on filters. However, since the success of these methods depends on the input data, more research is required to identify the best fusion of assembly and categorization algorithms for this strategy.

Future study will concentrate on confirming the results using other financial datasets. Further research is required to solve the slow processing speed of the existing machine learning method. Machine learning will be used with deep artificial neural networks, data mining, and other approaches to improve the precision and speed of real-time financial fraud detection. These developments are anticipated to speed up the detection process and offer more effective ways to prevent financial crime in real-time circumstances.

References

1. Liu et al (2018) Machine learning approach in FDS
2. Zhang S, et al (2019) Deep learning-based recommender system: a survey and new perspectives. *ACM Comput Surv (CSUR)* 52(1):1–38
3. Dimitrios K (2019) Can artificial intelligence replace whistle-blowers in the business sector. *Int J Technol Policy Law* 3(2):160–171
4. Panigrahi S, Kundu A, Sural S, Majumdar AK (2009) Credit card fraud detection: a fusion approach using Dempster-Shafer theory and Bayesian learning. *Inf Fusion* 10(4):354–363
5. Kundu A, Panigrahi S, Sural S, Majumdar AK (2009) Credit card fraud detection: a fusion approach using Dempster-Shafer theory and Bayesian learning. *Special Issue Inf Fusion Comput Secur* 10(4):354–363
6. Maes S, Tuyls K, Vanschoenwinkel B, Manderick B (1993) Credit card fraud detection using Bayesian and neural networks. In: Interactive image-guided neurosurgery, pp 261–270
7. Bentley PJ, Kim J, Jung G-H, Choi J-U (2000) Fuzzy Darwinian detection of credit card fraud. In: 14th Annual fall symposium of the Korean information processing society, 14th October 2000
8. Haykin S (1999) Neural networks: a comprehensive foundation, 2nd edn, p 842
9. Castelli M, Manzoni L, Popović A (2016) An artificial intelligence system to predict quality of service in banking organizations. *Comput Intell Neurosci*
10. Francois C (2017) Deep learning with Python
11. Brause R, Langsdorf T, Hepp M (1994) Neural data mining for credit card fraud detection. In: International conference on tools with artificial intelligence, pp 621–630
12. Kundu A, Sural S, Majumdar AK (2006) Two-stage credit card fraud detection using sequence alignment. In: Proceedings of the International conference on information systems security. Lecture notes in computer science, vol 4332, pp 260–275. Springer Verlag
13. Ning Z, Cox AJ, Mullikin JC (2001) SSAHA: a fast search method for large DNA databases. *Genome Res* 11(10):1725–1729
14. Madden T (2003) The BLAST sequence analysis tool
15. Altschul SF, Gish W, Miller W, Myers W, Lipman J (19990) Basic local alignment search tool. *J Mol Biol* 215:403–410
16. Sahin Y, Duman E (2011) Detecting credit card fraud by decision trees and support vector machines. In: Proceeding of International multi-conference of engineering and computer statistics, vol 1
17. Dileep MR, Navaneeth AV, Abhishek M (2021) A novel approach for credit card fraud detection using decision tree and random forest algorithms. In: 2021 Third International conference on intelligent communication technologies and virtual mobile networks (ICICV). IEEE

Remodeling E-Commerce Through Decentralization: A Study of Trust, Security and Efficiency



Adnan Shakeel Ahmed, Danish Raza Rizvi, Dinesh Prasad,
and Amber Khan

Abstract While there has been a massive growth in online shopping platforms. There data has been mostly stored in centralized storages which mean they can be an easy target to fraud and theft. It is important to preserve the data of the product and the user in a secure and decentralized manner. With the advent of blockchain technology across the horizon, it is necessary to build a marketplace model that works in a reliable way on this technology while combating the shortcomings of the previous models. Aim of this paper is to address the issues and propose a model that can get reduced fees costs while uploading the data to the blockchain that is uploading the item for sale and while purchasing the item, we also as well as find a solution to collusion from different parties. The end result of this paper should be a fully functional application that is capable of reducing gas fees and increases user reliability on the platform.

Keywords Blockchains · Decentralized Marketplace · Collusion attack

A. S. Ahmed (✉) · D. Prasad · A. Khan

Department of Electronics and Communications Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi 110025, India
e-mail: adnanshakeel@pm.me

D. Prasad

e-mail: dprasad@jmi.ac.in

A. Khan

e-mail: akhan1@jmi.ac.in

D. R. Rizvi

Department of Computer Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi 110025, India
e-mail: drizvi@jmi.ac.in

1 Introduction

In today's world, the boom of information and network communication technologies has led to the emergence of diverse digital commerce business models and firms that have become an integral part of our daily routines. The rising popularity of E-commerce can be attributed to its efficiency and ease of use, leading to a growing preference among consumers for online purchasing.

One of the primary concerns for buyers in the realm of purchasing goods online pertains to the selection of items that are of superior quality, while also being at an affordable price. In practical terms, when buyers have the intention of procuring a specific type of item, like "trainers," they typically browse numerous pages featuring that item. Every section contains descriptions of trainers, which include title, cost, images and promotional material provided by the vendor. Online shoppers have the ability to look over the merchandise offered by vendors, along with related descriptions of the goods, in order to make informed purchasing decisions. Centralized Systems are being implemented on popular E-commerce sites like eBay and Amazon. These systems are designed for processing and storing information in a centralized manner on the platforms' servers.

This paper aims to explore the following research questions:

- Can decentralized E-commerce platforms effectively safeguard buyers and sellers by protecting information and ensuring the optimal performance of a secure system?
- How can decentralized technologies promote convergence between buyers and sellers by preventing collusion attacks and reliable management?

The remainder of this paper is organized as follows, Sect. 2 introduces previous research to establish context of our research, Sect. 3 describes the research methods used to collect and analyze data. Section 4 presents the findings of our research and conclusion is drawn in Sect. 5.

2 Background and Related Work

There have been many attempts to create a mature application that has an architecture for a decentralized market which can provide immutability, confidentiality and is able to safeguard the product information and reputation for consumers; however, most designed architectures tend to focus on single transactions rather than an entire E-commerce platform, while others created a decentralized ecosystem with the marketplace acting as the middlemen. The establishment of trust and reputation plays a crucial role in facilitating the success of E-commerce. Marketplace middlemen are viewed with skepticism by buyers due to their experienced exploitative behavior. According to previous research, the assessment of risk is a crucial factor that holds equal significance for consumers when making purchases in the digital marketplace.

Pongnumkul and Chaiyaphum [1] conducted a comprehensive analysis of the performance of private blockchain systems. The study presents a comprehensive evaluation of the performance of private blockchain in diverse scenarios, thereby providing significant insights for organizations contemplating the integration of private blockchain technology. The observations can assist entities in making well-informed choices regarding the most appropriate setups and agreement mechanisms, resulting in streamlined and effective deployments of private blockchains that are tailored to their particular needs and limitations. The findings made by the authors offer significant perspectives for entities contemplating the adoption of private blockchain technology, enabling them to make informed choices regarding the most suitable arrangements and agreement mechanisms. Therefore, entities have the ability to enhance the performance and efficacy of their private blockchain implementations by customizing them to their individual requirements and limitations.

Lopez and Farooq proposed and novel approach to overcome the challenges of the growing ecosystem for a secure, transparent and efficient data exchange in the context of smart mobility systems. There architecture used the blockchain technology to propose a distributed, safe and a protected platform for data exchange between various stakeholders in the transport ecosystem. This multilayered framework allowed to separately focusing on enhancing the scalability as well as addressing the broader needs of different stakeholders such as data producers. Consumers and service providers. The authors argued that the proposed architecture has the ability to revolutionize the way data is shared, managed and monetized [2].

Various systems [3–8] for assessing credibility have been suggested for diverse security contexts. The most famous reputation systems for E-commerce platforms are the following. Hasan et al. [3, 4] conducted a thorough examination aimed at identifying and analyzing diverse methodologies for designing and implementing reputation systems, while simultaneously ensuring the preservation of privacy and security. The author's findings provide a comprehensive comprehension of the current status of privacy-preserving reputation systems, elucidating the merits and demerits of various approaches. The paper provides a comprehensive examination and evaluation, rendering it a valuable asset for scholars and professionals seeking to construct and implement secure, privacy-preserving reputation systems utilizing blockchain and cryptographic methodologies.

Vágújhelyi [9] investigates the utilization of decentralized blockchain technology for the purpose of implementing time-lock encryption. Time-lock encryption is a security mechanism that restricts access to encrypted data until a pre-established time frame has elapsed or certain predetermined conditions have been satisfied. The utilization of blockchain technology is suggested by the author as a means to establish a medium that is both transparent and resistant to tampering, for the purpose of storing and verifying time-lock conditions. The author highlights the significance of time-lock encryption, which can be executed through decentralized blockchain applications, in augmenting data security and generating novel applications for secure and timed data exchange. This can potentially make a valuable contribution to the field of cryptography and secure communication.

Compared to the existing literature, the goal of this paper is to build an architecture which is capable of storing the data of users in a truly decentralized manner without the involvement of a third party intermediary which can act maliciously and find a way to address collusion attacks.

3 Research Approach

3.1 Study Design

We conducted our study on the Ethereum blockchain [10, 11] which is a collection of nodes that operate an Ethereum client. We adopted an experimental approach, the study is divided in to two sections, in the first section we discuss the details of the proposed framework (see Fig. 1), for our E-commerce platform. In the next section we work on the design of the smart contract [12–14] and the methodology used to address our problems.

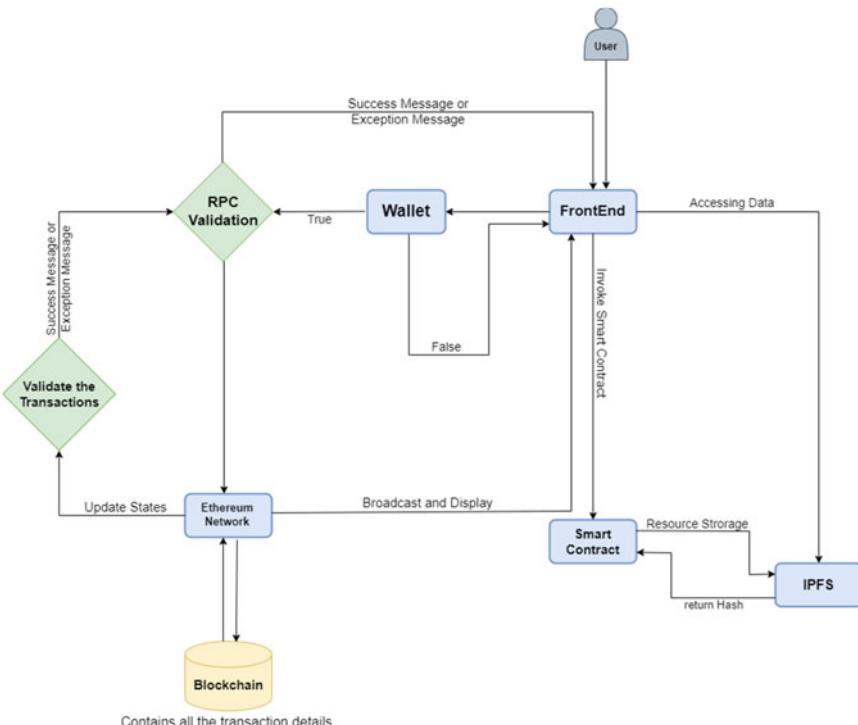


Fig. 1 Proposed architecture for an E-commerce platform

The previous models in existence have worked on creating a system that is semi decentralized, they store the available item information and the user data on to IPFS [15, 16] while they have a centralized database which is used to query all the incoming requests used to fetch the information on the marketplace. In our model, we propose a way of querying as well as storing the necessary information on the blockchain in a manner that does not cost much gas fee [17, 18] to the user while fetching them from the blockchain.

3.2 Proposed System Architecture

In order to facilitate an explanation of the proposed architecture, we need to know that the overall structure has been divided into 3 core parts with multiple other roles involved in the proposal. The application does not involve the use of any centralized databases [19] like Mongo DB [20] for optimization which can act as a central point of failure. We use IPFS [21] to store the resources in a decentralized and secure fashion.

The 3 parts of the architecture and their design is discussed below.

Front End

The Front End uses tools like React JS to build up our UI and we use Tailwind CSS for the styling. The purpose of the Front End is to streamline the process of interacting with blockchain for the user.

Smart Contract

Every time we interact with the blockchain we use Alchemy as our node service provider and this is where our smart contract makes all the Remote Control Procedural calls to validate transactions and make changes in the Ethereum network.

Decentralized Data Storage

The third most important part of this decentralized application is the Decentralized Data Storage that we use, in our case we use IPFS where we will upload the metadata about our item and helps in reducing the cost to fetch items from the blockchain each time.

3.3 Implementation Methodology

We aim to design a unique and a verifiable digital token for the item that is transparent and can be viewed, stored or even sent over the blockchain (Table 1).

The design of the smart contract can be divided into three main steps:

Table 1 System variables in our smart contract

msg.sender()	Individual who initiates the execution of a smart contract
msg.value()	The amount of monetary value within the executed smart contract
this.balance()	Used to get the balance of the current contract instance

Item Creation

The seller of the item creates a unique item by providing the necessary details like title, price, description and other relevant information.

Input: IPFS URL that has metadata of the new item created.

Output: It assigns a ItemId to the item and saves the metadata.

// First check if the seller has enough ETH to list and create an item

```
require(msg.value == listPrice)
```

checks if the seller has sent enough ETH as compared to the listing price.

require(price > 0) we also make sure that the price of the item is not in negative.

//Now we use the safeMint function in solidity to create the new item

```
safeMint(msg.sender, newItem)
```

setItemURI(newItem, ItemURI) we are able to able to map this newly created item to the IPFS URL that also contains all the item metadata.

ItemId.increment we then increment the item counter by 1 for each new item created in the marketplace. We keep track of the newly generated item by using the Counters library from OpenZeppelin.

Execute Sale Between the Seller and the Buyer

In this we first need to verify that the item that is being sold has been uploaded and is available on the system.

Input: The input is going to be the ItemId of the item that is going to be sold.

Output: Nothing is returned.

//First we get the price of the item and the address of the seller.

//It then checks if the value sent with the transaction is equal to the price of the item.

```
require(msg.value == price);
```

//If the value sent with the transaction is equal to the price of the item, it updates the details of the item by setting currently Listed to true.

```
idListedItem[ItemId].currentlyListed = true;
```

It also increments the number of items sold counter so that we can keep track of how many items have been sold.

//Using the transfer function we successfully transfer the ownership of the item from the seller to the buyer.

```
transfer(address(this), msg.sender, ItemId);
```

Function to Safeguard from Collusion

In this we transfer the item token to the marketplace smart contract and it is done to avoid any peer-to-peer transactions and create a escrow like mechanism which is not centralized. We create a function which has the following conditions implemented in the contract.

//This calculates the marketplace fee.

```
uint fee = (item.price * feeRate) / 100;
```

Add a condition to see if the amount submitted covers the marketplace fee.

```
require(msg.value >= fee, "Payment does not cover marketplace fee.");
```

//Transfer the fee to the marketplace.

```
owner.transfer(fee);
```

//Transfer the rest of the amount to the seller.

```
item.seller.transfer(msg.value - fee);
```

This way we are able to ensure that the marketplace gets its due fees, reduces direct peer-to-peer and provides a level of trust and security.

4 Result

Metrics were collected after testing and validating all of the transactions on the Goerli Test Network. The network furnishes comprehensive statistical data that can be utilized to monitor the duration of transaction execution and the quantity of gas expended during the transaction. Ganache [22, 23] was utilized to produce many customer accounts to replicate legitimate worldwide transactions. We analyzed that model was able to create a platform that optimally safeguard the buyers and sellers by protecting the item information by uploading it to a decentralized IPFS storage system.

4.1 Gas Fees and Time Cost Analysis

We tested the gas fees for the two main functions that were carried out as the number of blocks in the blockchain was increased.

Figure 2 indicates that the gas expenses of the two functions which exhibit a marginal rise in tandem with the augmentation of block quantity, owing to the

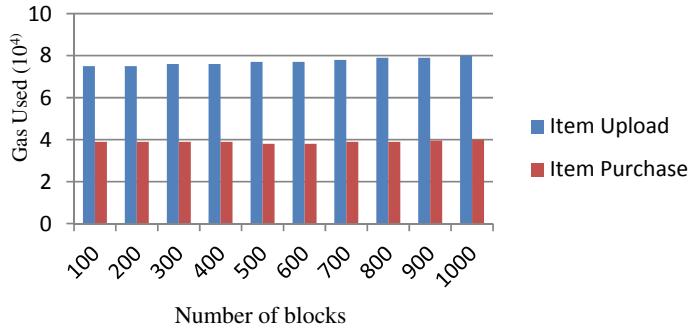


Fig. 2 Gas fees analysis

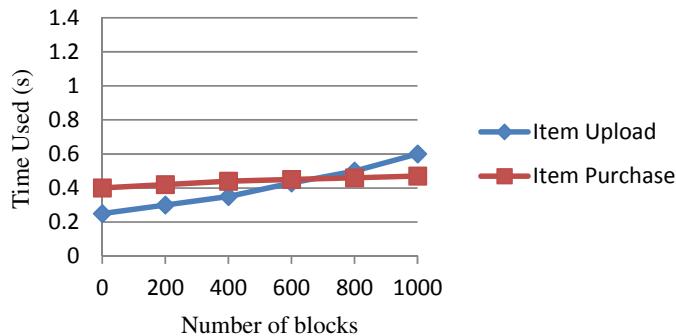


Fig. 3 Time cost analysis

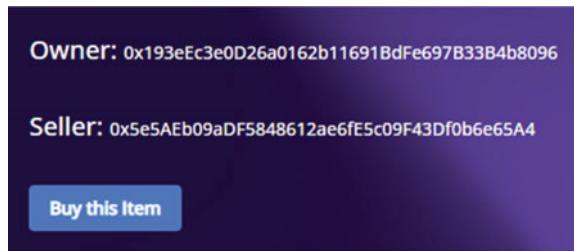
subsequent factors. Firstly, even searching for an address has a complexity of $O(1)$. Secondly, since we will be storing only the URL address to the blockchain, there will be no escalation in gas costs and will remain stable.

We also calculate the time cost [24] in both functions (see Fig. 3), they show a modest increase as the quantity of the blocks grows. As the amount of blocks grows, there is a small rise in the time required for obtaining and maintaining data within IPFS.

4.2 Reliability Analysis

We were able to solve the issue of collusion attacks [25, 26] (see Fig. 4) between the seller and buyer, we programmed the smart contract to transfer the item from the seller address to the marketplace address, and when the buyer would purchase the item, the marketplace would ensure that the seller receives its amount once the marketplace fees has been deducted from the selling price of the item.

Fig. 4 A view from the purchase item



In case of collusion between the seller and the marketplace or the buyer and the marketplace, it is important that we choose a decentralized blockchain and make sure to go through the smart contract which will be responsible for enforcing the rules and regulations of the market. This process should be automated or else it could undermine the integrity of the marketplace. We can also see the identity of the seller and the buyer also remains confidential as the attacker cannot identify the identity of the addresses, hence protecting the trade from any manipulation.

5 Conclusions and Future Scopes

5.1 Conclusions

In regards with the problem definition of our research we can say that our application stands successful. We implemented smart contracts that help in ensuring that the rules of the marketplace are enforced on everyone regardless of which party it is be it the seller, buyer or the marketplace itself.

We also see that we were able to successfully carry out our trading function of item upload by the seller and item purchase by the buyer in a very low cost. It shows that the system is feasible. The only cost that the seller has to bear is the listing fee and the marketplace fee for carrying out the trade between the two parties. We also came to a solution for the collusion attack between the seller and the buyer by transferring the item to the marketplace and releasing the item only once the marketplace fees has been submitted. We see that our model is also tamper proof as most of the item data is stored off chain in a distributed manner. In our case we also see that the marketplace does not hold back any of the proceeds of the trade.

5.2 Future Scopes

It would be interesting to see decentralized marketplaces having advancement in cross-chain protocols which can facilitate the integration of decentralized marketplaces with multiple blockchain systems, resulting in enhanced cooperation, liquidity and market expansion opportunities. In addition, integrate IoT [27], artificial intelligence and machine learning tools into our ecosystem to help the user identify fraud. AI tools could also enhance the user experience by optimizing the search for the user and returning better recommendations based on their preferences. There can be more work in the domain of protecting user privacy by adding zero knowledge proofs to our system and can help in maintaining the anonymity of the users.

References

1. Pongnumkul S, Siripanpornchana C, Thajchayapong S (2017) Performance analysis of private blockchain. In: 26th International conference on computer communication and networks (ICCCN)
2. Lopez D, Farooq B (2020) A multi-layered blockchain framework for smart mobility data-markets. *Transp Res Part C Emerg Technol* 111:588–615
3. Hasan O, Brunie L, Bertino E (2022) Privacy-preserving reputation systems based on blockchain and other cryptographic building blocks: a survey. *ACM Comput Surv* 55(2):1–37. <https://doi.org/10.1145/3490236>
4. Schaub A, Bazin R, Hasan O, Brunie L, A trustless privacy-preserving reputation system, pp 398–411
5. Mukherjee A, Liu B, Glance N, Spotting fake reviewer groups in consumer reviews, pp 191–200
6. Buechler M, Eerabathini M, Hockenbrocht C, Wan D (2015) Decentralized reputation system for transaction networks. Technical Report, University of Pennsylvania
7. Dennis R, Owen G, Rep on the block: a next generation reputation system based on the blockchain, pp 131–138
8. Pavlov E, Rosenschein JS, Topol Z (2004) Supporting privacy in decentralized additive reputation systems. In: Jensen C, Poslad S, Dimitrakos T (eds) Trust management. Springer, Berlin, Germany, pp 108–119
9. Vágújhelyi F (2018) Time-lock encryption by decentralized blockchain application
10. Sahu L, Sharma R, Sahu I, Das M, Sahu B, Kumar R (2021) Efficient detection of Parkinson's disease using deep learning techniques over medical data. *Expert Syst* e12787. <https://doi.org/10.1111/exsy.12787>
11. Sharma R, Kumar R, Sharma DK et al (2021) Water pollution examination through quality analysis of different rivers: a case study in India. *Environ Dev Sustain*. <https://doi.org/10.1007/s10668-021-01777-3>
12. Ha DH, Nguyen PT, Costache R et al (2021) Quadratic discriminant analysis based ensemble machine learning models for groundwater potential modeling and mapping. *Water Resour Manage*. <https://doi.org/10.1007/s11269-021-02957-6>
13. Dhiman G, Sharma R (2021) SHANN: an IoT and machine-learning-assisted edge cross-layered routing protocol using spotted hyena optimizer. *Complex Intell Syst*. <https://doi.org/10.1007/s40747-021-00578-5>
14. Sharma R, Gupta D, Polkowski Z, Peng S-L (2021) Introduction to the special section on big data analytics and deep learning approaches for 5G and 6G communication networks (VSI-5g6g). *Comput Electr Eng* 95:107507. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2021.107507>

15. Singh PD, Dhiman G, Sharma R (2022) Internet of things for sustaining a smart and secure healthcare system. *Sustain Comput Inf Syst* 33:100622. ISSN 2210-5379. <https://doi.org/10.1016/j.suscom.2021.100622>
16. Sharma R, Arya R (2021) A secure authentication technique for connecting different IoT devices in the smart city infrastructure. *Cluster Comput.* <https://doi.org/10.1007/s10586-021-03444-8>
17. Sharma R, Arya R (2021) Secure transmission technique for data in IoT edge computing infrastructure. *Complex Intell Syst.* <https://doi.org/10.1007/s40747-021-00576-7>
18. Rai M, Sharma R, Satapathy SC et al (2022) An improved statistical approach for moving object detection in thermal video frames. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11548-x>
19. Verma R, Sharma R (2022) Dual notched conformal patch fed 3-D printed two-port MIMO DRA for ISM band applications. *Frequency.* <https://doi.org/10.1515/freq-2021-0242>
20. Sharma N, Sharma R (2022) Real-time monitoring of physicochemical parameters in water using big data and smart IoT sensors. *Environ Dev Sustain.* <https://doi.org/10.1007/s10668-022-02142-8>
21. Anandkumar R, Dinesh K, Obaid AJ, Malik P, Sharma R, Dumka A, Singh R, Khatak S (2022) Securing e-health application of cloud computing using hyperchaotic image encryption framework. *Comput Electr Eng* 100:107860. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2022.107860>
22. Sharma R, Arya R (2022) UAV based long range environment monitoring system with industry 5.0 perspectives for smart city infrastructure. *Comput Ind Eng* 168:108066. ISSN 0360-8352. <https://doi.org/10.1016/j.cie.2022.108066>
23. Truffle Suite Configuration. <https://truffleframework.com/docs/advanced/configuration>
24. Ahmed AS, Ahmad MO (2020) Performance comparison of MPLS and ATM based networks. In: Proceedings of the 2nd International conference on ICT for digital smart and sustainable development
25. Singh S, Hosen ASM, Yoon B (2021) Blockchain security attacks, challenges, and solutions for the future distributed IoT network. *IEEE Access* 1–1. <https://doi.org/10.1109/ACCESS.2021.3051602>
26. Thakur S, Breslin J (2020) Collusion attack from hubs in the blockchain offline channel network. https://doi.org/10.1007/978-3-030-37110-4_3
27. Akhtar MM, Rizvi DR (2020) IoT-chain: security of things for pervasive, sustainable and efficient computing using blockchain. *EAI Endorsed Trans. Energy Web*

Estimation of Wildfire Conditions via Perimeter and Surface Area Optimization Using Convolutional Neural Network



R. Mythili[✉], K. Abinav, Sourav Kumar Singh, and S. Suresh Krishna

Abstract Wildfires are a major natural disaster that can cause significant damage to ecosystems and human communities. Wildfire behavior must be predicted accurately for effective emergency response and evacuation planning. The proposed system suggests a novel Convolutional Neural Networks (CNNs) method to estimate wildfire conditions via optimization of perimeter and surface area. Extraction of the required features is from the historical wildfire data which is performed before preprocessing and training. The trained CNN is then validated and optimized for performance, with the goal of accurately predicting wildfire behavior in real-time. The proposed system results show the effectiveness of the method, improving the wildfire prediction ability. The outcome of the prediction determines the probability of a wildfire. The results can be used to monitor areas where wildfires are expected. This helps to implement strategies that can be used to mitigate the effects of wildfires. The combination of perimeter and surface area optimization with CNNs represents a promising new approach to wildfire prediction and management, with broad applications in land management, sustainability, and emergency response.

Keywords Machine learning · Convolutional neural network · Google earth engine

1 Introduction

Wildfires are a major environmental challenge worldwide, causing significant damage to ecosystems, homes, and infrastructure, as well as human injury and loss of life [1, 2]. Early detection and accurate prediction of wildfire conditions are essential for effective prevention, management, and containment of these disasters [3]. Traditional methods of estimating wildfire behavior involve manual analysis of satellite

R. Mythili (✉) · K. Abinav · S. K. Singh · S. S. Krishna

Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India

e-mail: mythilir2@srmist.edu.in

imagery or ground observations, which can be time-consuming and error-prone. Convolutional neural networks (CNNs) are showing significant potential in the field of wildfires due to developments in deep learning [4].

In recent years, there have been many studies that used CNNs for wildfire detection and analysis [5, 6]. However, most of these studies have focused on detecting the location of a wildfire in satellite images, while less attention has been paid to estimating the wildfire conditions, of fire spread and intensity [7]. The proposed work suggests a novel approach to estimate wildfire conditions using CNNs, based on perimeter and surface area optimization.

The proposed method utilizes the changes in perimeter and surface area of a wildfire over time to estimate its spread and intensity. The outcome of the prediction determines the probability of a wildfire. The results can be used to monitor areas where wildfires are expected. This helps to implement strategies that can be used to mitigate the effects of wildfires [8–10].

To validate the proposed method, real-world wildfire datasets are used, including ground observations, from several regions around the world. The results are compared with existing methods for wildfire condition estimation, hence proved in terms of accuracy and computational efficiency [11–14].

Overall, the proposed method provides a novel approach to estimate wildfire conditions using CNNs, which can significantly improve the accuracy and efficiency of wildfire analysis. The proposed methodology can provide valuable insights for policymakers, land managers, and first responders, enabling them to take proactive measures to mitigate the impact of wildfires [15].

2 Existing Systems

Existing systems for estimating wildfire conditions use a range of methods and technologies to gather and analyze data. These systems can include remote sensing, weather data, ground-based observations, mathematical models, machine learning, and GIS. By integrating and analyzing these data sources, these systems can provide valuable information to help manage and mitigate the impacts of wildfires. Although the present-day wildfire prediction systems have advanced significantly in recent years [16–19], yet there are number of issues, need to be resolved in order to increase their efficiency and accuracy. The limitations include the data quality and data availability, the lack of integration between different systems, limited spatial and temporal resolution, difficulty in modeling complex interactions, and limited understanding of long-term trends. In recent years, there have been many studies that used CNNs for wildfire detection and analysis [20–26]. However, most of these studies have focused on detecting the location of a wildfire in satellite images, while less attention has been paid to estimating the wildfire conditions, such as the fire's intensity and spread. The study includes a novel approach for estimating wildfire conditions using CNNs, based on perimeter and surface area optimization [27–30].

Implementation of new strategies is required to minimize the wildfire effects. The proposed system is implemented as data collection, preprocessed missing/null values, and applying Convolutional Neural Network algorithm to predict the accuracy and visualize the same. The modular structure of this code makes it simple to alter the date, data source, location, or historical and spatial sampling. The proposed model can be used to enhance the data collection from restricted data sources, which are accessible at GEE platform. Newer datasets might include information on an international scale or wider span time frame. Due to more fires and larger burned regions, recent data is considered to be heavily weighted than a decade data. The model is able to predict the wildfires using the attributes like rain, humidity, the burned area of the forest, etc. The system advantages include the use of CNN technology and accuracy, choice of prediction attributes.

3 Proposed System Architecture

The proposed wildfire detection system as in Fig. 1 includes the various below-mentioned modules and the relevant functional implementation.

Figure 1 shows the proposed system architecture on estimation of wildfire conditions via perimeter and surface area optimization using convolutional neural network (CNN) which involve the below-mentioned six key steps:

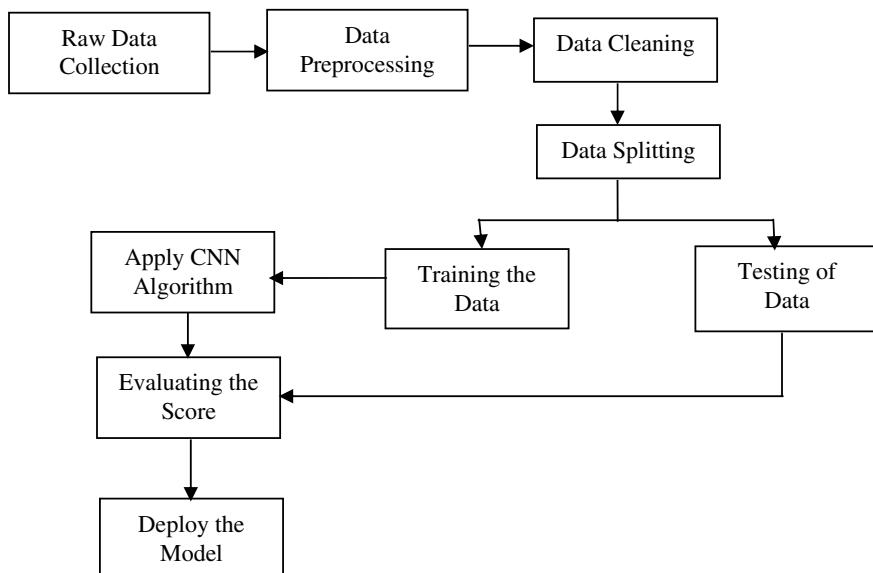


Fig. 1 Proposed system architecture

- (i) **Collecting the Raw Data:** The primary step involves collecting raw data from various sources, such as satellite and aerial imagery, weather data, and ground-based observations. The raw data will be in different formats and structures, so it needs to be collected, organized, and integrated into a common database.
- (ii) **Preprocessing the Data:** The raw data will undergo various preprocessing steps to remove noise, filter out irrelevant information, and convert data into a common format. Preprocessing steps can include image resizing, color correction, and normalization to ensure that the data is consistent and compatible with the CNN.
- (iii) **Cleaning and Preparing the Data:** By eliminating outliers and missing values, the preprocessed data will then be further cleaned and prepared by being converted into the correct input format for the CNN. For the data to be reliable, consistent, and reflective of the underlying population, this stage is essential.
- (iv) **Data Splitting to Train and Test:** Both the training and testing sets are created from the cleaned and prepared data. The CNN is trained through training set and effectiveness assessing by the test set.
- (v) **Applying CNN:** The CNN algorithm is applied on the training data to learn the correlation between input attributes and output labels. Convolutional, pooling, and fully connected layers constitute the components that make up the CNN architecture. Using hyperparameter tweaking, the number of layers, filters, and activation functions will be optimized.
- (vi) **Evaluating the Score:** On the basis of the testing data, the performance of CNN is assessed by the major performance measures, include accuracy, precision, recall, and F1-score. The evaluation metrics will provide insights into how well the CNN is performing and whether it is accurately predicting the behavior of wildfires.

4 Module Implementation

The proposed system modules are functionally implemented and executed as below mentioned to detect the wildfire in an effective manner.

4.1 Collection of Data

Data collection is a process that gathers information on wildfire conditions from a variety of sources, which is then utilized to create machine learning models. A set of wildfire conditions data with features is the type of data used in the proposed work. For the purpose of the proposed system, Forest Fire dataset, Kaggle [31] is used. The shape of the dataset is (517 rows * 13 columns) and it has 13 attributes including fires, rain, drought index, etc.

4.2 Preprocessing the Data

To arrive at the meaningful data, data preprocessing is done as cleaning and validating. Data preprocessing often involves the below-mentioned processes:

Formatting: It is expected that the data format which is selected prevents from data handling. Data may be from the relational database or text file if it is in a proprietary file format or from a flat file if it is in a relational database. The proposed model involves the change of the month and days from text to numeric quantity because it will be more convenient for machine learning models to deal with numbers, and ‘fire’ attribute is also renamed with output.

Cleaning: Data cleansing is the process of replacing missing data. There may be times when the data is either incomplete or missing. There may be a need to put a halt to these instances. In the proposed dataset, there are no blank or missing values.

Sampling: Algorithms may take a lot longer to handle bigger volumes of data and may require more computational power and memory to operate. It may be much faster to explore and test ideas by choosing a smaller representative sample of the specified data rather than the complete dataset.

4.3 Extraction of Features

This phase of extraction of features involves adding/removing, transforming and combining features to accomplish the final set of features. For example, the feature selection is used to identify the major features for predicting wildfire conditions, or feature extraction technique is used to extract features from data of wildfires. Data standardization is done by eliminating the mean and scaling to unit variance.

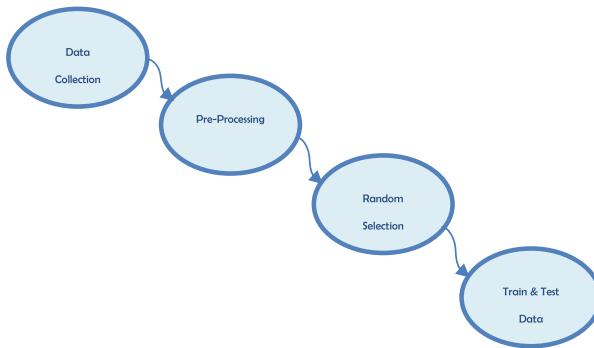
$$\text{Score} = (x - \text{mean}) / \text{std} \quad (1)$$

Dataset is fitted to scaler by removing the attribute output.

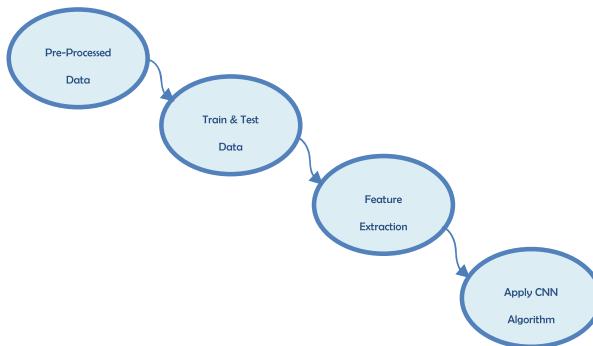
4.4 Evaluating the Model

Model evaluation is the next phase in the model development process. It deals with the model finding, that most closely matches the data and predicts how the model will perform in future. In data science, it is unacceptable to use training data to evaluate model performance because doing so can quickly lead to optimistic and fitting models. Significant metrics of recall, accuracy, R^2 score, and precision are used to evaluate the CNN performance on the test data. The evaluation metrics will shed light on how well CNN is functioning and whether it is successful in foreseeing wildfire activity. Figure 2 describes the data flow of the proposed model.

Phase 1



Phase 2



Phase 3

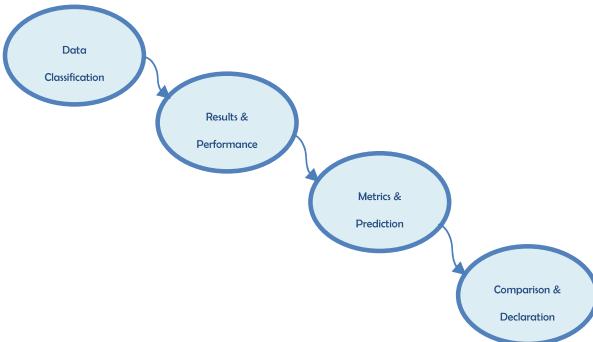


Fig. 2 Data flow diagrams

5 Result Analysis

The proposed system is executed with the dataset Forest Fire dataset from Kaggle [31] and the major metric results are derived as mentioned in Table 1.

The metrics evaluation approach for the proposed model, which is compared to R-Squared and Root Mean Square Error (RMSE), is shown in Table 1. Root Mean Square Error (RMSE) between the actual and anticipated values (R_{ai} and R_{pi}) is calculated using the RMSE score. RMSE of the proposed model is calculated as

$$\text{RMSE} = \sqrt{\left[\sum (R_{pi} - R_{ai})^2 / n \right]}, \quad (2)$$

which is numerically equal to the value of 0.4924266 for the proposed system model.

A measurement of how well a model fits the data is called R-Squared. \mathbf{R}^2 is calculated as

$$R^2 = 1 - \text{RSS/TSS}, \quad (3)$$

where R^2 is the determination coefficient.

RSS stands for the residual squared sum.

TSS stands for total square sum.

R^2 value of the model is 0.02990449.

The model accuracy and validation loss of our model are shown in Figs. 3 and 4. The proposed model's accuracy range is between 0.53 and 0.62, and its validation loss range is between 0.694 and 0.655.

On execution of the proposed system, Fig. 5 shows prediction of wildfires using the attributes of dataset. If the model outputs 1 means, wildfire is detected else wildfire is not detected.

Table 1 Metrics results of the proposed system

Metrics	Existing system	Proposed system
R^2 score	0.1126990	0.02990449
RMSE value	0.6629467	0.4924266

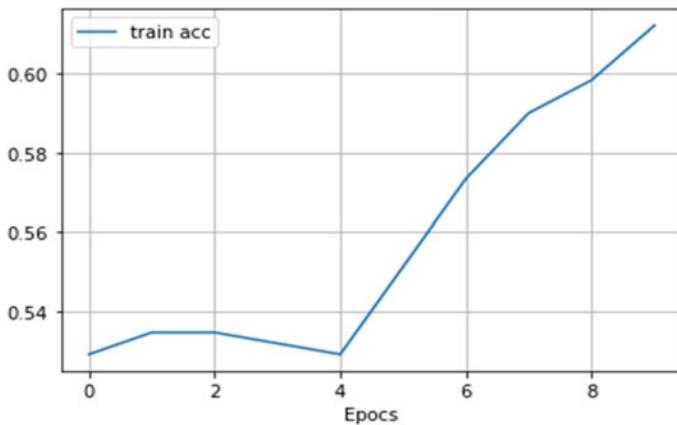


Fig. 3 Model accuracy

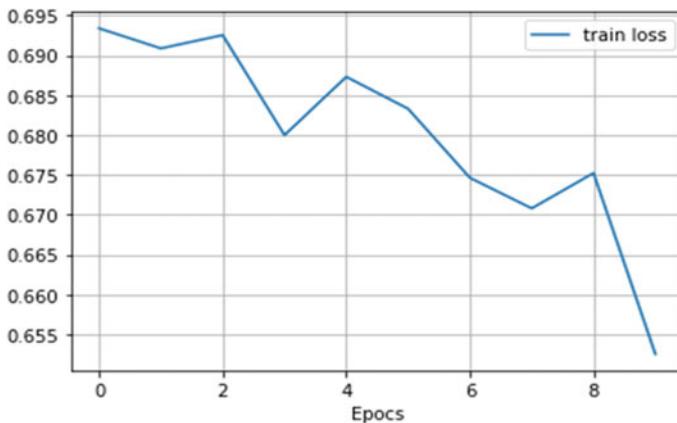


Fig. 4 Validation loss

```
1 test=model3.predict(np.array([[1,4,8,6,90.2,96.9,624.2,8.9,20.3,39,4.9,0.0]]))
2 print(test)
3 if test[0]==1:
4     print("wildfire detected")
5 else:
6     print("wildfire not Detected")
[[1.0151881e-11]]
wildfire not Detected
```

Fig. 5 Detection of wildfires

6 Conclusion

In conclusion, estimation of wildfire conditions via perimeter and surface area optimization using convolutional neural network is a promising solution for predicting the behavior of wildfires. By leveraging the power of machine learning and image processing, the system can provide accurate predictions of wildfire perimeters and surface areas, which are crucial for effective firefighting and resource allocation.

The system architecture involves the data of publicly available Forest Fire dataset [31] for data preprocessing, after that model is trained using the CNN, and optimization of the parameters is done before evaluating the performance. The model is evaluated against RMSE value and R^2 value with visualization graphs and reports. There are also several future enhancements that can be considered, such as incorporating additional data sources, implementing real-time data processing, and developing predictive models for fire behavior.

Overall, the proposed approach of estimation of wildfire conditions via perimeter and surface area optimization using convolutional neural network, significantly improves the accuracy and speed in a better way in view of wildfire predictions and ultimately helps protecting the human life and property assets. As wildfire events continue to increase in frequency and severity, the development and deployment of such systems are becoming more critical than ever.

7 Future Enhancements

There are several future enhancements that can be considered for estimation of wildfire conditions via perimeter and surface area optimization using convolutional neural network. Some of them are:

Incorporating additional data sources: The system relies on satellite and aerial imagery, weather data, and ground-based observations. In the future, other sources of data such as social media and drone imagery can be incorporated to improve the accuracy of the CNN.

Implementing real-time data processing: The system relies on batch processing of data, which can be time-consuming and limit the speed of predictions. Implementing real-time data processing can reduce the response time for firefighting efforts and help to prevent the spread of wildfires.

Integrating with other fire management systems: Other fire management systems, such as fire modeling and suppression systems, can be integrated with the system, to provide a comprehensive wildfire management solution.

References

1. Westerling AL et al. (2006) Warming and earlier spring increase western US forest wildfire activity. *Science* 313.5789:940–943
2. Petkovic M et al. (2020) Optimization of geographic information systems for forest fire risk assessment. In: 2020 21st international symposium on electrical apparatus and technologies (SIELA), IEEE
3. Gure M et al. (2009) Use of satellite images for forest fires in area determination and monitoring. In: 2009 4th international conference on recent advances in space technologies, IEEE
4. Guang Y, Di X (2011) Adaptation of Canadian forest fire weather index system and it's application. In: 2011 IEEE international conference on computer science and automation engineering, vol 2. IEEE
5. Devadevan V, Suresh S (2016) Energy efficient routing protocol in forest fire detection system. In: 2016 IEEE 6th international conference on advanced computing (IACC), IEEE
6. Gao X, Fei X, Xie H (2011) Forest fire risk zone evaluation based on high spatial resolution RS image in Liangyungang Huaguo mountain scenic spot. In: Proceedings 2011 IEEE International conference on spatial data mining and geographical knowledge services, IEEE
7. Sullivan AL (2009) Wildland surface fire spread modelling, 1990–2007. 2: empirical and quasi-empirical models. *Int J Wildland Fire* 18(4):369–386
8. Van Der Werf GR et al. (2017) Global fire emissions estimates during 1997–2016. *Earth Syst Sci Data* 9.2:697–720
9. Jain P et al. (2020) A review of machine learning applications in wildfire science and management. *Environ Rev* 28.4:478–505
10. Sahu L, Sharma R, Sahu I, Das M, Sahu B, Kumar R (2021) Efficient detection of Parkinson's disease using deep learning techniques over medical data. *Expert Syst e12787*. <https://doi.org/10.1111/exsy.12787>
11. Sharma R, Kumar R, Sharma DK et al. (2021) Water pollution examination through quality analysis of different rivers: a case study in India. *Environ Dev Sustain*. <https://doi.org/10.1007/s10668-021-01777-3>
12. Ha DH, Nguyen PT, Costache R et al. (2021) Quadratic discriminant analysis based ensemble machine learning models for groundwater potential modeling and mapping. *Water Resour Manage*. <https://doi.org/10.1007/s11269-021-02957-6>
13. Dhiman G, Sharma R (2021) SHANN: an IoT and machine-learning-assisted edge cross-layered routing protocol using spotted hyena optimizer. *Complex Intell Syst*. <https://doi.org/10.1007/s40747-021-00578-5>
14. Sharma R, Gupta D, Polkowski Z, Peng S-L (2021) Introduction to the special section on big data analytics and deep learning approaches for 5G and 6G communication networks (VSI-5g6g). *Comput Electri Eng* 95:107507. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2021.107507>
15. Singh PD, Dhiman G, Sharma R (2022) Internet of things for sustaining a smart and secure healthcare system, *Sustain Comput: Inform Syst* 33:100622. ISSN 2210–5379. <https://doi.org/10.1016/j.suscom.2021.100622>
16. Sharma R, Arya R (2021) A secure authentication technique for connecting different IoT devices in the smart city infrastructure. *Cluster Comput*. <https://doi.org/10.1007/s10586-021-03444-8>
17. Sharma R, Arya R (2021) Secure transmission technique for data in IoT edge computing infrastructure. *Complex Intell Syst*. <https://doi.org/10.1007/s40747-021-00576-7>
18. Rai M, Sharma R, Satapathy SC et al. (2022) An improved statistical approach for moving object detection in thermal video frames. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-021-11548-x>
19. Verma R, Sharma R (2022) Dual notched conformal patch fed 3-D printed two-port MIMO DRA for ISM band applications. *Frequenz*. <https://doi.org/10.1515/freq-2021-0242>

20. Sharma N, Sharma R (2022) Real-time monitoring of physicochemical parameters in water using big data and smart IoT sensors. Environ Dev Sustain. <https://doi.org/10.1007/s10668-022-02142-8>
21. Anandkumar R, Dinesh K, Obaid AJ, Malik P, Sharma R, Dumka A, Singh R, Khatak S (2022) Securing e-Health application of cloud computing using hyperchaotic image encryption framework. Comput Electri Eng 100:107860. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2022.107860>
22. Sharma R, Xin Q, Siarry P, Hong W-C (2022) Guest editorial: deep learning-based intelligent communication systems: Using big data analytics. IET Commun. <https://doi.org/10.1049/cmu2.12374>
23. Sharma R, Arya R (2022) UAV based long range environment monitoring system with Industry 5.0 perspectives for smart city infrastructure. Comput Indus Eng 168:108066. ISSN 0360–8352. <https://doi.org/10.1016/j.cie.2022.108066>
24. Russakovsky O et al. (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115:211–252
25. Bennett J, Stan L (2007) The netflix prize. In: Proceedings of KDD cup and workshop. vol 2007
26. Alonso-Betanzos A et al. (2003) An intelligent system for forest fire risk prediction and fire fighting management in Galicia. Expert Syst with Appl 25.4:545–554
27. Sakr GE, Elhajj IH, Mitri G (2011) Efficient forest fire occurrence prediction for developing countries using two weather parameters. Eng Appl Artif Intell 24(5):888–894
28. Dutta R et al. (2013) Deep cognitive imaging systems enable estimation of continental-scale fire incidence from climate data. Scient Reports 3.1: 1–4
29. Hodges JL, Lattimer BY (2019) Wildland fire spread modeling using convolutional neural networks. Fire Technol 55:2115–2142
30. Radke D, Hessler A, Ellsworth D (2019) FireCast: leveraging deep learning to predict wildfire spread. IJCAI
31. Cortez P, Morais A (2007) A data mining approach to predict forest fires using meteorological data. In Neves J, Santos MF, Machado J (eds), New trends in artificial intelligence

A Framework Provides Authorized Personnel with Secure Access to Their Electronic Health Records



Kanneboina Ashok and S. Gopikrishnan

Abstract Cloud-based data storage has become ubiquitous in e-commerce, education, research, and health care. Among these industries, health care has the potential to leverage cloud technology to store electronic health records (EHRs) and reduce costs associated with maintaining manual paper-based records. The storage, retrieval, and maintenance of electronic health records for healthcare providers have been challenging due to the influx of patients and the requirement to retain their records for extended periods. Consequently, many healthcare providers are keenly interested in transitioning their EHRs to cloud-based platforms. The risks mentioned above stem from storing electronic health records in distant geographical locations that may not be familiar to healthcare providers. Despite the existing proposals in the literature to address security concerns, there remains a demand for a proficient security framework capable of managing all security issues associated with using cloud storage for EHRs. The present study introduces a security framework that incorporates mechanisms to address confidentiality, integrity, and access control. The framework ensures secure storage and retrieval of electronic health records (EHRs) in cloud-based environments. The research makes a significant contribution by proposing a confidentiality mechanism named Segregation and Preserve Privacy of Sensitive Data in Cloud Storage (SPPSICS). This encryption technique is specifically designed to safeguard the privacy of sensitive data in electronic health records (EHRs) that are stored in cloud storage. The proposed mechanism endeavors to partition the sensitive and insensitive attributes of electronic health records (EHRs) and employ strong encryption techniques to safeguard the sensitive attributes.

Keywords Electronic health records · Health service providers · Healthcare providers · SPPSICS

K. Ashok (✉) · S. Gopikrishnan

School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh 522237, India

e-mail: ashok.21phd7048@vitap.ac.in

S. Gopikrishnan

e-mail: gopikrishnan.s@vitap.ac.in

1 Introduction

Digital technology has significantly transformed our daily routines and communication methods. Upholding paper-based patient records within healthcare establishments can be arduous over time. Electronic health records (EHRs) are utilized in healthcare systems to facilitate the seamless exchange of information. Electronic health record systems have the potential to fully digitize the healthcare system by utilizing clinical information to assist healthcare providers in delivering superior patient care. EHRs can automate and optimize the physician workflow while producing a comprehensive patient record, thereby improving the standard of care. Electronic health records enable ubiquitous access to health information, allowing users to retrieve data from any location at any time. Primary issue-related EHRs pertain to safeguarding the confidentiality and integrity of medical records. Unauthorized disclosure of sensitive information can stigmatize and embarrass the affected patient. Manipulating health record data can significantly alter its significance, constituting a violation and potentially degrading the patient's sense of self-worth. When these records are transferred to a publicly accessible platform, the responsibility to ensure their security is heightened. Preserving confidentiality necessitates promptly implementing appropriate safeguards to secure medical records, limiting access solely to authorized personnel. Diverse encryption methods are employed to safeguard the confidentiality of EHRs [1].

The employment of cloud-based storage for electronic health records has garnered increasing attention recently, owing to its advantageous features such as lowered expenses and enhanced availability. Nonetheless, this phenomenon also gives rise to diverse security apprehensions, such as the divulgence of data, violation of integrity, and unapproved entry to electronic health records. The challenges associated with implementing sufficient security measures to safeguard electronic health records in the cloud stem from their remote storage locations, which are not readily discernible to healthcare providers [2].

The present study puts forth a security framework that incorporates mechanisms aimed at safeguarding the confidentiality, integrity, and access control of electronic health records stored in cloud-based systems to tackle the security concerns associated with such systems. The framework endeavors to offer a holistic resolution that effectively manages all the security concerns associated with utilizing cloud storage for electronic health records. The initial contribution of this research work is the proposed confidentiality mechanism, SPPSICS, which segregates sensitive and insensitive attributes of EHRs and encrypts the sensitive attributes with strong encryption [2, 3].

The proposed framework also includes mechanisms for access control, data integrity, and secure communication. Access control mechanisms ensure that only authorized individuals can access EHRs, while data integrity mechanisms ensure that EHRs are not tampered with or modified without proper authorization. Secure communication mechanisms are also implemented to ensure the confidentiality of data in transit between healthcare providers and the cloud storage provider.

Overall, this research aims to provide a robust security framework that addresses the various security issues in storing EHRs in the cloud. This framework will help healthcare providers to migrate their EHRs to the cloud without compromising the security of sensitive patient information.

2 Literature Survey

The use of cloud-based storage for electronic health records has been extensively studied in the literature. Several studies have focused on the benefits of cloud technology in health care, including increased accessibility, cost savings, and scalability. However, these studies also acknowledge the security risks associated with cloud-based storage of EHRs.

Other studies have focused on specific security issues, such as data leakage and integrity breaches. For example, Ali et al. [27] proposed a secure data leakage prevention mechanism for EHRs that employs machine learning algorithms to detect and prevent unauthorized access to patient medical data. On the other hand, Wang et al. [26] proposed a blockchain-based solution for ensuring the integrity and traceability of EHRs in the cloud.

Akram et al. [1] conducted a literature review of the current state-of-the-art in cloud-based storage for EHRs. The review identified several benefits of cloud technology for health care, including increased accessibility, cost savings, and scalability. However, the review also highlighted several security risks associated with cloud-based storage of EHRs, including data leakage, integrity breaches, and unauthorized access. The review concluded that there is a need for comprehensive security frameworks to ensure the secure storage and retrieval of EHRs in the cloud.

Kaur and Singh [2] surveyed existing security mechanisms for EHRs in the cloud. The survey identified several mechanisms for ensuring the confidentiality, integrity, and access control of EHRs, including encryption, access control, and data masking. The survey also highlighted the importance of ensuring the privacy of patient medical data in the cloud-based storage of EHRs.

Alzahrani et al. [3] systematically reviewed the literature on cloud-based storage for EHRs. The review identified several benefits of cloud technology for health care, including increased accessibility, cost savings, and scalability. However, the review also highlighted several security risks associated with cloud-based storage of EHRs, including data leakage, integrity breaches, and unauthorized access. The review concluded that comprehensive security frameworks are required to secure the cloud storage and retrieval of EHRs.

Aljawarneh et al. [4] conducted a survey of existing security challenges in cloud-based storage for EHRs. The survey identified several challenges, including data privacy, ownership, segregation, and residency. The survey also proposed several mechanisms for addressing these challenges, including encryption, access control, and data masking.

Chen et al. [5] surveyed existing security mechanisms for the cloud-based storage of EHRs in China. The survey identified several mechanisms for ensuring the confidentiality, integrity, and access control of EHRs, including encryption, access control, and data masking. The survey also identified several challenges to adopting cloud technology for healthcare in China, including regulatory compliance, data privacy, and data security.

The present study endeavors to bridge the existing void in the scholarly literature by presenting a holistic security framework that guarantees the confidentiality, integrity, and access control of electronic health records in the cloud. The framework under consideration presents mechanisms aimed at segregating attributes of electronic health records (EHRs) into categories of sensitivity and insensitivity. Additionally, the framework employs robust encryption techniques to secure sensitive attributes. The framework incorporates access control mechanisms to guarantee that exclusively authorized users can gain entry to electronic health records (EHRs).

3 Proposed Security Framework

Figure 1 depicts the security framework proposed to safeguard electronic health records in cloud-based storage. The framework involves several steps for the secure storage and retrieval of EHRs [12].

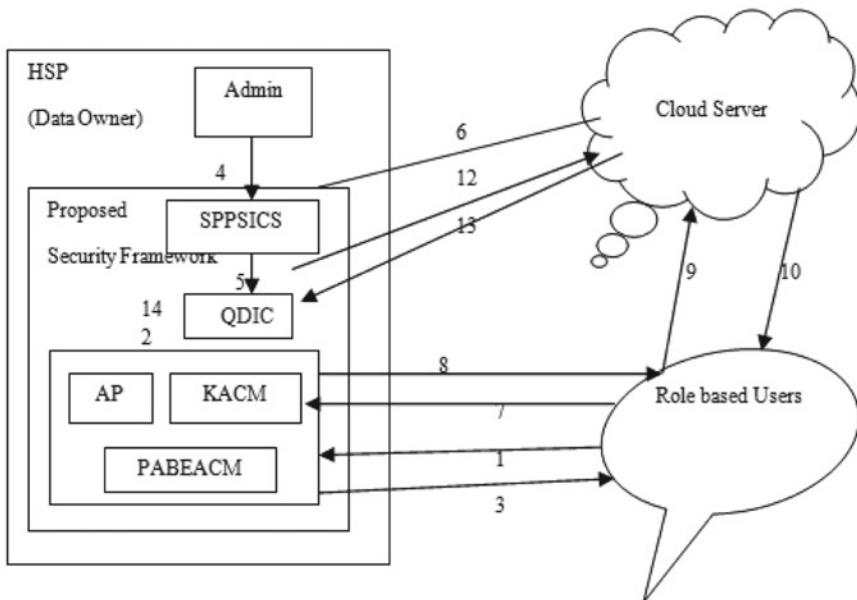


Fig. 1 Verily, a framework of security for electronic health records in the storage of cloud security

During the initial phase, individuals create their profiles by enrolling with the Health Service Provider (HSP). During the second stage, the Home Service Provider (HSP) transmits the user's profile data to the Access Provider (AP). The Access Point (A.P.) produces a series of authentication details for the user. It ascertains the attribute I.D.s relevant to the consumer's role list and the Key Access Control Matrix (KACM). The Access Point (A.P.) generates the identification numbers for attributes and the corresponding hash function.

Following this, the Access Point disseminates to the users with role-based privileges the username, password, hash function, hash parameter, eligible attribute identifiers, and the secret key K. During the fourth stage, the HSP/administrator transmits the accumulated data, characteristics, and electronic health records to the suggested security architecture. The framework comprises three distinct mechanisms: SPPSICS, QDIC, and PABEACM (Partial_Grained_Attribute-Based Encryption for Secure Data Access in the Cloud).

One of the mechanisms employed is SPPSICS, which categorizes electronic health records attributes into two distinct groups: high-sensitive and less-sensitive attributes. The information that has been segregated is subjected to encryption by utilizing Advanced Encryption Standard (AES) and Data Encryption Standard (DES) algorithms. The AES algorithm is utilized to encrypt the attributes that are deemed highly sensitive. In contrast, the DES algorithm is employed to encrypt the corresponding attributes considered less sensitive. The encrypted information is tagged and passed to the query-based data integrity checker (QDIC) for hash generation.

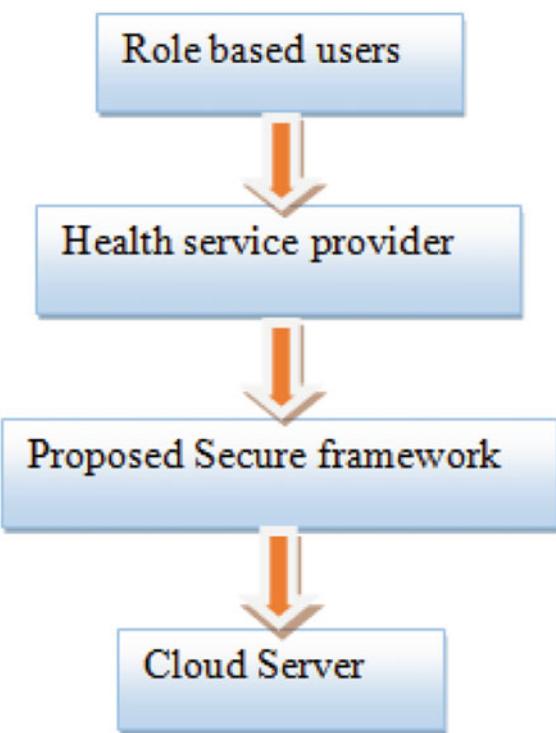
The encrypted electronic health records are sent from SPPSICS to QDIC in the fifth step. The QDIC utilizes the Message Digest 5 (MD5) algorithm to generate a hash value stored and maintained in the hash table. Sixth, SPPSICS places the encrypted EHRs in a safe location on a cloud server. The HSP or data owner's administrator may now talk to the cloud server safely, thanks to this measure.

Upon successful authentication, the PABEACM system verifies the user's assigned roles and generates hash parameters based on the identified attributes. This process involves consulting the KACM system to determine the user's eligibility for specific attributes. Subsequently, the hash parameters corresponding solely to the eligible attributes are transmitted to the user. In the ninth stage of the procedure, the user submits a request to the cloud server, providing the patient ID, to retrieve the EHR.

The methodology presented in this research comprises four distinct stages, as depicted in Fig. 2. During the initial phase, individuals input their personal and medical data into the system via either the data owner or the administrator of the healthcare service provider. During the second phase, the system gathers pertinent data according to the users' designated roles. Phase 3 is the proposed security framework, which includes mechanisms such as SPPSICS, QDIC, and PABEACM, as explained already. Finally, in phase 4, the encrypted information or electronic health records (EHRs) are stored and retrieved from the cloud server.

The security framework under consideration comprises three primary security mechanisms, as depicted in Fig. 3. One of the mechanisms employed is SPPSICS, which involves segregating sensitive and insensitive attributes of electronic health

Fig. 2 Different levels of the proposed model



records (EHRs) and their subsequent encryption using symmetric key algorithms. QDIC is the second mechanism that guarantees the reliability of electronic health records (EHRs) by validating their credibility. The suggested structure's final component is PABEACM, a system for managing access that operates on the basis of user roles.

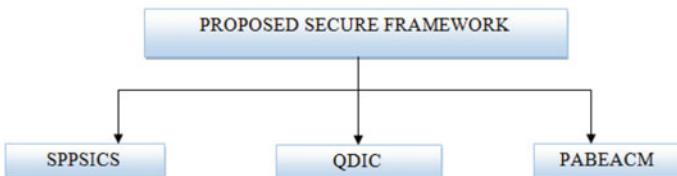


Fig. 3 Proposed security model

4 Key Features of the Proposed Security Framework

The proposed security framework for electronic health records (EHRs) is founded on a web-based system that guarantees secure login and enrollment. The online system facilitates ubiquitous access for users, enabling them to connect to the system from any location and at any time, contingent upon the availability of a reliable internet connection. The system's design aims to guarantee that exclusively authorized users are granted access to pertinent information. The system mandates the registration of role-based users, such as patients, doctors, nurses, technicians, chemists, insurance companies, and administrators. Upon registration, the administrator furnishes the user with a secret code mandatory for accessing the data [6, 7].

Patient records stored on a cloud are susceptible to unauthorized access, making security a crucial issue. Encryption is the most effective way to safeguard these records. The proposed EHR security framework employs an ontology-based model to classify EHR attributes based on their sensitivity levels. Highly sensitive information is encrypted with a stronger encryption algorithm like AES, while less-sensitive data is encrypted with a weaker algorithm like DES. The utilization of DES and AES algorithms is employed to ensure the confidentiality of electronic health records based on their respective sensitivity levels. This ensures that only authorized users can access the EHRs and that sensitive information is hidden from potential attackers [8].

Data Integrity Checking: To ensure the integrity of EHRs, the proposed framework utilizes the MD5 hashing technique. The resulting integrity verification report and digital signatures are sent to the relevant users [8].

Access Control: Different privileges or access rights are maintained to ensure that the system maintains the appropriate levels of authority for role-based users. For example, a patient has the authority to view all attributes of their EHR. Still, they may only modify limited attributes such as their address or mobile number with similar attributes. However, they are not provided append and delete rights [10].

5 Experimental Results and Discussion

The security framework in question has been successfully instantiated by utilizing the freely available software known as phpMyAdmin. The system's user interface is constructed utilizing a combination of PHP, JavaScript, and HTML, while the MySQL database management system supports the system's underlying functionality. The information proprietor employs encryption techniques to secure the electronic health records, and subsequently, the encrypted data is deposited onto a cloud server via a CSP. To assess the efficacy of the proposed framework, a series of experiments were carried out on a dataset comprising EHRs encompassing various data types.

The present study conducts a comparative analysis of the efficacy of the proposed security framework vis-à-vis the security systems introduced by Prerna et al. [29] and Sombir et al. [28]. EHR files of different input file sizes (1, 7, 15, 50, and 100 KB) were utilized for comparison. The encryption period of seconds for various input file sizes is shown in Fig. 3.

The findings suggest that the system introduced by Prerna et al. [29] requires 0.0104 s to encrypt a 1-kilobyte input file, whereas for a file size of 15 kilobytes, the processing time amounts to 0.114 s. When the file size is increased to 100 KB, it takes 1.04 s for encoding. On the other hand, Sombir et al. [28] method takes 0.3733 s to encrypt 1 KB, and it requires 37.33 s to encrypt a file of number 100 KB.

The framework under consideration was evaluated, wherein an input file of 1 KB was utilized. The encryption process was observed to take 0.0302 s, among other findings. The findings indicate that the framework put forth requires a shorter duration in comparison to the system developed by Sombir et al. [28] while necessitating a longer duration than the system devised by Prerna et al. [29].

Figure 4 presents a comparative analysis of the decryption duration in seconds across varying input file magnitudes. Prerna et al. [29] state that their system can decrypt a 1 KB input file in 0.0065 s. The execution time of 0.0975 s was observed for a 15 KB input file. The decryption process for a file size of 100 KB takes approximately 0.65 s. According to Sombir et al. [28], their system requires 0.3733 s to decrypt a 1 KB input file, while it takes 37.33 s to decrypt a file that is 100 KB in size. The framework under consideration exhibits a decryption time of 0.019 s when processing a 1 KB input file, while a file of 100 KB size requires 1.9 s for decryption.

The QDIC under consideration has been formulated utilizing the open-source software phpMyAdmin. The front end of this system employs PHP, JavaScript, and HTML, while MySQL supports the back end. The data owner will encrypt the electronic health records utilizing an encryption algorithm. The resultant encrypted data/

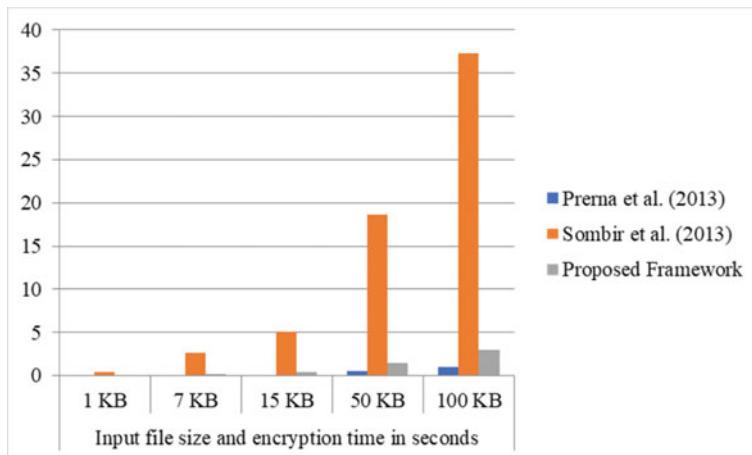
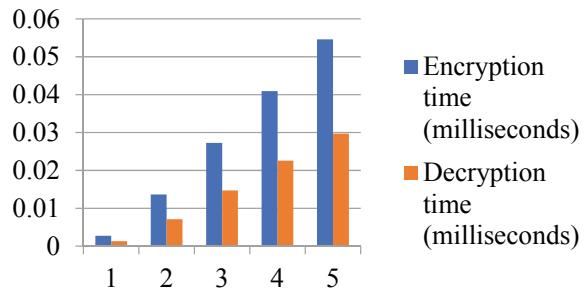


Fig. 4 Comparative analysis of the duration of encryption in seconds

Fig. 5 Time taken for encryption and decryption in seconds



EHRs will then be stored in a cloud server furnished by a cloud service provider. The QDIC, which has been suggested, has been implemented to store and manage electronic health records. The system above comprises a variety of attributes that exhibit distinct data types. The electronic health records under consideration in this study comprise five attributes. The encryption and decryption times, measured in milliseconds (ms), are presented in Fig. 5 while varying the number of attributes. The encryption process for a single attribute utilizing the AES algorithm resulted in a computational time of 0.00273 ms. The execution time of 0.02805 ms was observed for a system with five attributes. The duration of encryption was determined by altering the number of attributes. The time required for the decryption of a single attribute is 0.0015 ms. The duration required to process five attributes is 0.0073 ms. The decryption duration was evaluated for a maximum of 45 attributes and yielded a maximum time of 0.0675 ms.

The proposed Quality-Driven Integrity Check (QDIC) mechanism aims to safeguard the electronic health records' integrity by enabling the data owner to furnish the tag of an EHR to the QDIC to verify the safety of the original encrypted EHR. The report generated by the proposed data integrity checker for the EHR serves as the ultimate means of verifying the accuracy and consistency of the information that has been stored.

A comparison is made between the number of sub-keys necessary for the proposed PABEACM and the ABE access control mechanism [16]. Figure 6 presents a comparison. This study examines the number of keys needed based on different combinations of sensitive and insensitive attributes. Within the context of ABE, it is necessary to utilize distinct sub-keys for the encryption and decryption of individual attributes, regardless of their level of sensitivity. The proposed PABEACM scheme mandates that each attribute with a high level of sensitivity necessitates a distinct sub-key. However, only a single sub-key is necessary for the encryption/decryption of all insensitive attributes. Figure 6 shows that the ABE mechanism necessitates a total of 20 sub-keys to accommodate ten Hierarchical Secret Authorities (HSAs) and ten Independent Secret Authorities (ISAs). Conversely, the proposed PABEACM requires a mere 11 sub-keys. The ABE requires 40 sub-keys for 20 HSAs and 20 ISAs, while the PABEACM requires only 21 sub-keys. The data presented in Fig. 6 indicates a noticeable discrepancy in the number of sub-keys required by the ABE and

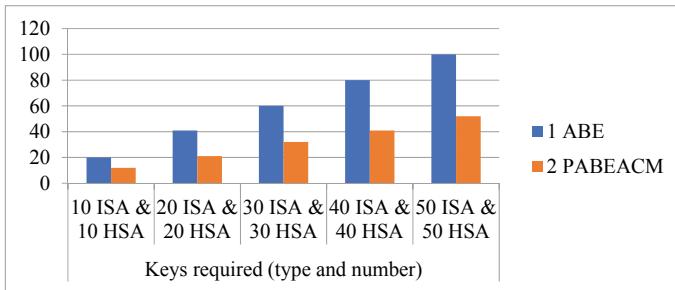


Fig. 6 Sub-keys generation

the proposed PABEACM as the quantity of ISAs rises. The utilization of PABEACM results in a notable reduction in the number of sub-keys required, thereby decreasing the duration of both key generation and rekeying processes.

Access control: This involves restricting access to EHRs to authorized personnel only.

Authentication: This involves verifying the identity of users accessing EHRs, such as through passwords or biometric authentication.

Audit trails: This involves keeping a record of all access to EHRs, which can help to identify unauthorized access attempts.

Data backup and recovery: This involves keeping copies of EHRs in case of data loss, corruption, or other issues.

Network security: This involves securing the network infrastructure used to transmit EHRs, such as through firewalls, intrusion detection systems, and other security measures.

EHRs have transformed the healthcare industry and enhanced the collaboration between doctors and patients. However, the susceptibility of health records to security breaches is significant in the absence of robust security measures. The proposed framework aims to augment the security measures of cloud-based electronic health records by incorporating security mechanisms catering to the three core security aspects: confidentiality, integrity, and access control. The effectiveness of the proposed framework was evaluated and compared with existing frameworks. This study presents the results and potential directions for future enhancements, as indicated by previous research [14–18].

6 Conclusion and Future Enhancement

To ensure that EHRs stored in the cloud are not tampered with, a new data integrity mechanism called QDIC has been developed. This is important because intentional or unintentional modifications may result in incorrect diagnoses. In addition, long-term

storage of medical records on disk drives can lead to bit rot, which can be misinterpreted during diagnosis. QDIC periodically retrieves encrypted EHRs stored in the cloud randomly and checks their hash values to ensure their integrity. Subsequently, a verification report is produced and provided to the data proprietor. The efficacy of QDIC was assessed on a cohort of authentic and manipulated electronic health records (EHRs), yielding a precision rate exceeding 90% for detection. QDIC eliminates needing a third-party auditor and guarantees data integrity without compromising efficiency.

Fine-grained access control mechanisms restrict access at the attribute level, resulting in high execution time and key usage; coarse-grained mechanisms offer access restrictions at a more general level of granularity, such as the level of files, which could result in the accidental exposure of confidential information. The PABEACM, currently under the proposal, is a mechanism for access control that operates at a partially grained level. Its primary function is to enforce fine-grained access control over highly sensitive attributes while simultaneously implementing coarse-grained access control over less-sensitive attributes. The ultimate goal of this mechanism is to maintain the security of highly sensitive data contained within EHRs. The PABEACM system effectively limits entry to data of a highly confidential nature, contingent upon the specific roles of users.

Meanwhile, attributes that are deemed insensitive are made accessible to all users, thereby resulting in a marked reduction in the time required for sub-key generation and distribution. A comparative analysis was conducted between the sub-key requirements of PABEACM and ABE, revealing that PABEACM exhibited superior efficiency. Furthermore, an efficient key management protocol could be added to enhance the proposed framework further.

References

1. Akram R, Mahmood K, Shahzad M (2020) A comprehensive review on cloud-based storage of electronic health records. *J Med Syst* 44(7):1–11
2. Kaur M, Singh P (2019) Security mechanisms for electronic health record storage in cloud: a survey. *Int J Appl Eng Res* 14(20):4045–4049
3. Alzahrani SH, Alharthi A, Almohammadi A, Alshahrani A (2021) Cloud-based electronic health records: a systematic review of security risks and privacy issues. *J Healthcare Eng* 2021:1–12
4. Aljawarneh S, Al-Jarrah OY, Al-Azzam A (2019) Cloud-based electronic health records: a review of security challenges and solutions. *J Med Syst* 43(7)
5. Chen J, Xu H, Yu Y, Gong Y (2018) A survey of security mechanisms for electronic health record storage in cloud in China. *J Med Syst* 42(11):1–11
6. Al-Mamun MA, Al-Fuqaha A (2018) Securing electronic health records in the cloud: a literature review. *IEEE Access* 6:31931–31947
7. Shrivastava R, Dwivedi AK (2019) A survey on security issues in cloud-based electronic health records. *Int J Eng Adv Technol* 8(6):1001–1007
8. Albakri SH, Alenezi A (2019) A review of security and privacy issues in cloud-based electronic health records. *Int J Comput Sci Netw Secur* 19(11):59–65

9. Dang X, Yang J, Wang Y, Chen J (2019) Security issues and solutions of electronic health records in cloud computing: a survey. *J Med Syst* 43(7):1–11
10. Abbas F, Gani A, Khan SU (2019) Cloud-based electronic health record system: a review. *IEEE Access* 7:50795–50810. <https://doi.org/10.1109/access.2019.2905767>
11. Alharthi A, Alzahrani SH, Mohammadi A, Alshahrani A (2020) Security issues in cloud-based electronic health records: a review. *Healthcare Technol Lett* 7(1):1–5
12. Fauzi A, Rosyid RH, Handayani PW, Permana H (2020) Security mechanisms in electronic health record systems: a review. In: 2020 international seminar on application for technology of information and communication (iSemantic), pp 1–5
13. Mohamed ME, Shahin A (2020) Cloud-based electronic health record systems: a systematic review. *Int J Healthcare Inform Syst* 15(1):15–33
14. Liu X, Zhang Z, Wang S (2020) A survey of security and privacy protection of electronic health record systems in cloud computing. *J Healthcare Eng* 2020:1–12
15. Azeez RA, Adamu AA (2020) A systematic review of security and privacy issues in cloud-based electronic health records. *J Healthcare Eng* 2020:1–9
16. Prasad B, Durbakula KRK, Mudunuri D (2021) Privacy-preserving secure electronic health records storage in cloud computing. *IEEE Access* 9:99431–99444
17. Rizwan M, Al-Muhtadi J (2021) Merging proxy re-encryption with attribute-based encryption for secure EHR sharing in the cloud. *IEEE Trans Cloud Comput* 9(1):326–335
18. Zhang C, Li X, Wei X, Zhang Y (2021) Blockchain-based health information management system with efficient and secure data sharing. *IEEE Trans Industr Inf* 17(2):1417–1425
19. Lu X, Jiao L, Guo Y, Li S (2021) An attribute-based and provable secure access control scheme for cloud-based EHR systems. *IEEE Trans Dependable and Secure Comput* 18(2):812–825
20. Zhang C, Liu W, Hu W, Wei X (2021) A privacy-preserving and efficient scheme for sharing health data in cloud-assisted mobile healthcare. *IEEE Trans Industr Inf* 17(4):2464–2474
21. Elsalamouny F, Mostafa N, ElBatt T (2021) Secure and efficient storage of electronic health records in cloud computing using proxy re-encryption and homomorphic encryption. *IEEE Access* 9:40952–40968
22. Hu W, Shen J, Liu Z, Zhang J (2021) Efficient and secure attribute-based access control for EHRs sharing in cloud-assisted healthcare systems. *IEEE Trans Industr Inf* 17(7):4684–4694
23. Qiu W, Liu Z, Shen J, Zhang J (2021) Efficient and secure decentralized access control scheme for sharing EHRs in mobile healthcare. *IEEE Trans Industr Inf* 17(8):5729–5739
24. Alassafi AY, Khattak A, AlZain MA (2021) A privacy-preserving secure data storage scheme for electronic health records in the cloud. *IEEE Access* 9:111193–111210
25. Pandey SK, Gupta PC, Singh A (2020) Securing electronic health records in cloud using encryption techniques. In: 2020 11th international conference on computing, communication and networking technologies (ICCCNT), Kharagpur, India, 2020, pp 1–6
26. Wang D et al (2020) Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *jama* 323(11):1061–1069
27. Ali F et al (2018) An assessment of the use of partial least squares structural equation modeling (PLS-SEM) in hospitality research. *Int J Contem Hospitality Manag* 30(1):514–538
28. Sombir S, Sunil KM, Sudesh K (2013) A performance analysis of DES and RSA cryptography. *Int J Emerg Trends & Technol Comput Sci (IJETTCS)* 2(3)
29. Preerna M, Abhishek S (2013) A study of encryption algorithms AES, DES and RSA for security. *Glob J Comput Sci Technol*

Explainable Artificial Intelligence for Deep Learning Models in Diagnosing Brain Tumor Disorder



Kamini Lamba and Shalli Rani

Abstract Deep neural networks (DNNs) have shown great potential in diagnosing brain tumor disorder, but their decision-making processes can be difficult to interpret, leading to concerns about their reliability and safety. This paper presents overview of explainable artificial intelligence techniques which have been developed to improve the interpretability and transparency of DNNs and have been applied to diagnostic systems for such disorders. Based on the utilized framework of explainable artificial intelligence (XAI) in collaboration with deep learning models, authors diagnosed brain tumor with the help of convolutional neural network and interpreted its outcomes with the help of numerical gradient-weighted class activation mapping (numGrad-CAM-CNN), therefore achieved highest accuracy of 97.11%. Thus, XAI can help healthcare professionals in understanding how a DNN arrived at a diagnosis, providing insights into the reasoning and decision-making processes of the model. XAI techniques can also help to identify biases in the data used to train the model and address potential ethical concerns. However, challenges remain in implementing XAI techniques in diagnostic systems, including the need for large, diverse datasets, and the development of user-friendly interfaces. Despite these challenges, the potential benefits for improving patient outcomes and increasing trust in AI-based medical systems make it a promising area of research.

Keywords Explainable artificial intelligence · Brain tumor · Health · Increased life expectancy · Deep neural network

K. Lamba · S. Rani (✉)

Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India

e-mail: shalli.rani@chitkara.edu.in

K. Lamba

e-mail: kamini.2400@chitkara.edu.in

1 Introduction

Being complex and intricate organs having more than 100 billion nerve cells [1], brain plays major role in the human body. Any kind of damage to it can result in loss of life for an individual if it is not identified and treated timely. For doing so, deep neural networks (DNNs) demonstrated tremendous potential in analyzing medical data, including the detection and diagnosis of brain tumors as these can automatically learn and extract complex features from medical images and make accurate predictions. Deep learning models are capable enough to perform analysis of medical data in large amount due to their inherited features as deep learning is a kind of machine learning which emphasizes on developing artificial neural networks having numerous layers [2] for extraction of high-level features while performing processing functions, whereas conventional approaches have limited capacity for processing raw data due to the presence of single transformation of inputs [3].

Moreover, deep neural networks can assist radiologists in decision-making process while diagnosing disorders at an early stage to provide timely treatment to the patients and increase their life expectancy. Other than this, deep neural networks [4] have ability to analyze chemicals, data associated with pharmacological in pharmaceutical industries for identification of people who consume drugs and optimization of formulation of drugs. Thus, such networks can assist pharmaceutical experts in gaining information of process in regard with drug discovery which could lead to generating its respective treatment terminologies. Due to the features of generating hypothesis based on extracted data by deep neural networks after identification of hidden patterns while performing analysis of input data, it has potential to disclose the interconnectivity of various devices to provide improved and accurate outcomes along with risk factors associated with respective stages of disorders for better understanding of complicated processes in regard with biological activities. To enable powers of data acquired from various devices such as wearables, sensors for monitoring real-time applications while diagnosing disorders, deep neural networks also have capacity to deploy multiple algorithms such as k-means clustering, fuzzy networks, Hough pixel transformations, random walk to detect any abnormalities lying within input MRI images to predict accurate and stable responses in the health-care sector as compared to the conventional approaches [5]. The working principle of deep neural networks generally considers a few stages in the health care. At initial stage, deep neural networks give significant contribution in the health care to acquire data such as medical images, genomic sequences, electronic health records from various devices such as wearables, sensors, digital machines. After acquiring data from distinct devices, data need preprocessing and cleaning with the help of filters to make sure that acquired data is free from any deterioration to achieve accurate and efficient outcomes in the health care. This is the high time where data augmentation can also be performed to balance the respective classes for avoiding compatibility issues if required. Such preprocessed form of data is applied to the architecture of deep neural networks for extraction of hidden patterns. At this point, various

networks such as convolutional, recurrent neural networks can be considered as per the requirement.

Training is provided to the considered architecture of deep neural networks to make the model convenient for recognizing images based on large scale of input data. It also help the model in reduction of differences occur between predicted output and ground truth labels for adjustment of its respective parameters such as weights as well as biases. Such training approach can also deploy backpropagation to compute its gradients responsible for modifying various parameters with the help of optimization algorithm for improving performance of model with respect to loss functions. After training model, evaluation is done based on test data to make sure that the predicted response matches with actual data which is possible via deploying fine-tuning approach to the hyperparameters of model for achieving efficient outcomes in health care.

Thus, deep learning has shown great potential to offer various advantages as well as probabilities to transform existing approaches in the healthcare sector. For instance, models comprising convolutional neural networks showed amazing results during analysis of medical images. These automated systems can help in diagnosing numerous diseases such as tumor, lung cancer, skin cancer [6], Alzheimer, Parkinson, skin lesions etc. an early stage due to which experts can consider these systems while making any decision regarding patients' health. Even algorithms based on deep neural networks [7] can also consider vast amount of clinical data for processing as well as analyzing purpose which further may include clinical notes, electronic health records, etc. Due to this property, such systems are capable enough to extract features from the given data to give contribution in decision-making process which may ultimately result in efficient as well as accurate outcomes.

This paper is organized as follows: Organization of this paper is done in the following way such as Sect. 2 presents the literature review following XAI approaches given in Sect. 3. Section 4 shows results of existing models accompanied by discussion, and Sect. 5 concludes the paper.

2 Literature Review

Integration of deep learning with explainable artificial intelligence (XAI) can aid radiologists in providing more efficient and accurate outcomes as compared to the existing approaches. Most of the researches proposed computer-assisted system to provide significant contributions in the medical field which often include development and evaluation of deep learning models for accurately detecting brain tumors, while also providing interpretable explanations for their decision-making processes. Ahmed et al. [8] also proposed an efficient deep neural network based model to diagnose various stages of brain tumor based on magnetic resonance imaging data. They also integrated deep learning model with SHapley additive explanations for better interpretation of outcomes achieved from deep learning models, whereas Jin et al. [9] generated post hoc explanatory algorithm for interpreting blackbox nature

of deep learning models to provide causes behind achieving specific predictions by an automated system while diagnosing and classifying brain tumor disease on the basis of input data containing magnetic resonance imaging.

For example, Kamnitsas et al. [10] proposed a 3D CNN for brain tumor segmentation in MRI images. Their method yielded innovative results based upon dataset of BraTS 2015 dataset. Bechelli et al. [11] provided theoretical explanations for utilizing machine learning and deep learning to diagnose cancer disorder. They also discussed challenges associated with proposed techniques while diagnosing cancer in health care. Sharma et al. [12] also proposed brain tumor diagnosis techniques based on deep learning models. They deployed visual geometry group following 19 layers to extract significant features from the provided input data. To perform such task, authors followed transfer learning process where model is already trained with much knowledge to perform the desired operation as a result it could take comparatively less time in learning new parameters additionally and perform efficiently, whereas Kukreja et al. [13] suggested deep learning models for recognition and classification of mathematical equations.

Although deep neural networks (DNNs) have shown promising results in the detection of brain tumors, however certain issues still require addressing at the earliest. Lack of interpretability of DNNs is found to be one of the significant issues, which makes it difficult to understand how they make decisions and can limit their clinical application. Explainable artificial intelligence (XAI) is a field of research that aims to address this challenge by developing methods to explain the reasoning of AI models. Several studies have highlighted the challenges of DNNs in brain tumor detection and the need for XAI. For example, the lack of diverse and well-annotated datasets can also be a challenge in developing accurate and generalizable DNNs for brain tumor detection. Thapa et al. [14] provided insights into traumatic brain injury which worsens functionality of brain and results in failure of cognitive thinking and decision-making processes. They also discussed therapies for dealing with consequences of traumatic brain injury along with recommendations given to patients for providing timely treatment to degrade impact of their injuries. Rehni et al. [15] also discussed significant causes behind pharmacological preconditioning of brain in which authors provided interplay between opioid and calcitonin gene-related peptide transduction systems.

Therefore, numerous strategies of explainable artificial intelligence in collaboration with deep learning [16] have been adopted by various researchers for developing a computer-assisted method for diagnosing brain tumor disease and interpreting outcomes of the proposed network for better visualization of each and every predicted response.

3 XAI Approaches

Following figures represent requirement of emerging XAI approaches for interpreting the predictions made by deep neural networks while diagnosing brain tumor disorder in which Fig. 1 illustrates the basic methodology of diagnosing brain tumor with the help of deep learning models which further comprises database that contains healthy as well as tumor images which are further preprocessed for resizing to maintain uniformity among all the images to obtain normalization and then forwarded to convolutional neural networks for extraction of features which are passed to classifiers for differentiating images having tumor from healthy ones. To make this approach more reliable among medical experts, Fig. 2 represents the key requirement behind utilizing explainable artificial intelligence techniques.

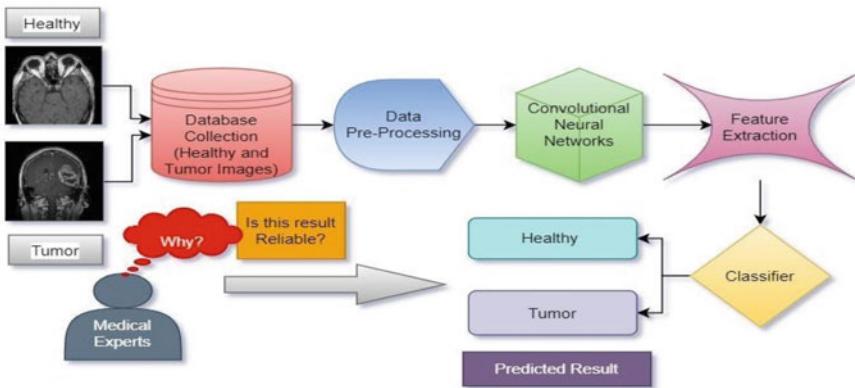


Fig. 1 Basic methodology of diagnosing brain tumor via deep learning models

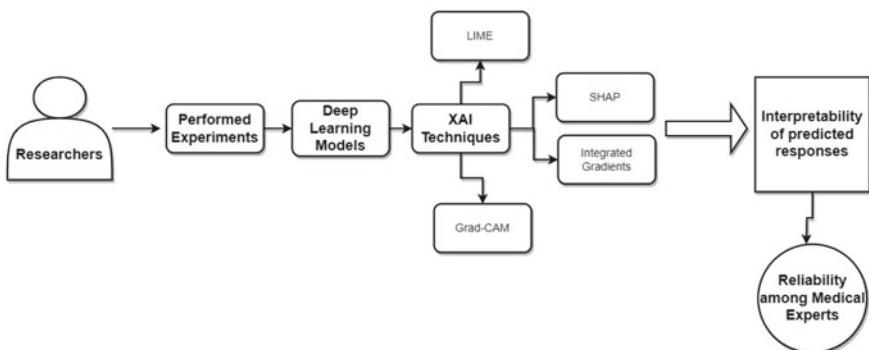


Fig. 2 Representation of XAI techniques and its necessity

3.1 Local Interpretable Model-Agnostic Explanations (LIMEs)

Ribeiro et al. [17] recommended this approach. It works by generating local surrogate models around a specific instance to explain its prediction. Given an input instance x and a trained machine learning model f , LIME approximates the prediction of f on x using a locally weighted linear regression model g as shown in Eq. 1:

$$g(x') = \arg \min_g L(f, g, pi'_x, omega). \quad (1)$$

Here, L is a loss function that measures the dissimilarity between f and g , pi'_x is a proximity kernel that measures the similarity between x and x' , and $omega$ is a regularization term that controls the complexity of g . The proximity kernel pi'_x is a kernel function that assigns higher weights to instances that are closer to x in some feature space. The weights of the linear regression model g are obtained by solving the following optimization problem as shown in Eq. 2:

$$\text{minimize}_g = L(f, g, pi'_x, omega) + lambda * \|g\|_1. \quad (2)$$

Here, $lambda$ is a hyperparameter that controls the sparsity of the weights, and $\|g\|_1$ is the $L1$ norm of the weight vector of g . Once the surrogate model g is trained, LIME generates explanations by analyzing the weights of the features in g that are most relevant for the prediction of f on x . The relevance of each feature is measured by its weighted contribution to the prediction of g on x . This can be done using a feature importance metric such as the absolute weight of the corresponding coefficient in the weight vector of g .

3.2 SHapley Additive ExPlanations (SHAPs)

It was proposed by Lundberg and Lee [18]. SHAP explains the prediction of a machine learning model f on input instance x as a linear combination of the contributions of its input features. Specifically, the SHAP value of feature j for instance x is defined as shown in Eq. 3:

$$phi_j(x, f) = 1/N * sum_{S \subseteq Z} [(|S| - 1)!(N - |S|)!/N! * (f(x_S) - f(x_{S-j}))]. \quad (3)$$

Here, Z represents set of all input features, $N = Z$, and x_S is the instance x with the features in S fixed to their background values. The background values are the values of the features in a reference dataset that are representative of the population. The notation Sj denotes the set S with feature j removed.

The SHAP value $\phi_{j|f}(x, f)$ represents the expected contribution of feature j to variance among f 's prediction on instance x and prediction of f on instance with feature j fixed to its background value. The expectation is taken over all possible subsets S of features, weighted by the number of ways to form each subset. The SHAP values can be computed efficiently using an algorithm called SHAPley value which is termed as a notion from collaborative game analysis which distributes a fair portion of a game's overall reward among every player proportional to their marginal contributions.

3.3 Integrated Gradients

It is a model interpretation technique that explains the predictions of any machine learning model by attributing the contribution of each feature to the prediction using the path integral of the gradients. It was proposed by Sundararajan et al. [19]. Here, Integrated Gradients compute the attribution of feature j for predicting machine learning model f on an input instance x as shown in Eq. 4:

$$IG_j(x, f) = (x_j - x'_j) * \text{integral}[\alpha = 0^1 \text{grad}_x' f(x + \alpha * (x' - x))] d\alpha. \quad (4)$$

Here x' is a baseline instance with the same shape as x , and $\text{grad}_x' f(x + \alpha * (x' - x))$ is the gradient of f in comparison with x_j at point $x + \alpha * (x' - x)$. The baseline instance x' is usually chosen to be a black image or a white noise image. The integral is approximated using numerical integration methods such as the trapezoidal rule.

The Integrated Gradients value $IG_j(x, f)$ represents the contribution of feature j to the prediction of f on instance x , relative to the baseline instance x' . The path integral of the gradients measures the change in the output of the model as feature j is varied along a straight line between x and x' , weighted by the gradient of the output with respect to feature j at each point along the line.

3.4 Gradient-Weighted Class Activation Mapping (Grad-CAM)

It was proposed by Selvaraju et al. [20]. It calculates gradients of predicted class score in comparison with feature maps of last convolutional layer in network. Formally, given an input image x , a deep neural network f , and a target class c , Grad-CAM computes the class activation map $M_c(x)$ as follows:

1. Let $A(k)$ be activation map of k -th convolutional layer in $f(x)$, and let y^c be the predicted score of class c before the softmax activation.
2. Compute the gradient of y^c with respect to $A(k)$ as shown in Eq. 5:

$$\alpha^c = 1/|A(k)| * \sum_i \sum_j [dy^c / dA(k)ij]. \quad (5)$$

Here, $dy^c / dA(k)ij$ is the gradient of y^c with respect to the activation at location (i, j) in $A(k)$.

3. Compute the weight w^c for each activation map as the global average pooling of the corresponding α^c as shown in Eq. 6:

$$w_k^c = \sum_i \sum_j \alpha^c / |A(k)|. \quad (6)$$

4. Generate the class activation map $M_c(x)$ via linearly combining activation maps $A(k)$ with the weights w^c as shown in Eq. 7:

$$M_c(x) = \text{relu}(\sum_k w^c A(k)). \quad (7)$$

Here, $\text{relu}(x) = \max(0, x)$ is the rectified linear activation function. The resulting class activation map $M_c(x)$ highlights the regions in the input image that are relevant for the predicted class c .

4 Results and Discussion

Following Table 1 represents the results of existing state-of-the-art approaches comprising explainable artificial intelligence techniques and deep neural networks for diagnosing brain tumor disorder.

Figure 3 represents the accuracy achieved by utilized XAI frameworks while diagnosing brain tumor for deep learning models. Thus, Pertzborn et al. [21] utilized matrix-assisted laser desorption/ionization (MALDI) imaging technique having ability to diagnose single-layer masses to diagnose brain tumor with the help of deep learning model followed by explainable artificial intelligence algorithms for clear explanations of results achieved by deployed model and achieved 80% accuracy. Gaur et al. [22] utilized convolutional neural network for detection of brain tumor and deployed explainable artificial intelligence techniques such as local interpretable model-agnostic explanations (LIME) and SHapley additive exPlanations

Table 1 State-of-the-art approaches for diagnosing brain tumor via XAI on DNN

References	Model	Techniques	Performance metrics
Pertzborn et al. [21]	Deep neural network with XAI	MALDI-XAI	80% accuracy
Gaur et al. [22]	Dual-input CNN with XAI	CNN, LIME, SHAP	94.64% accuracy
Park et al. [23]	Deep convolutional neural networks (DCNNs)	Transfer learning	97.01% accuracy
Marmolejo et al. [24]	XAI-CNN	Numgrad-CAM-CNN	97.11% accuracy

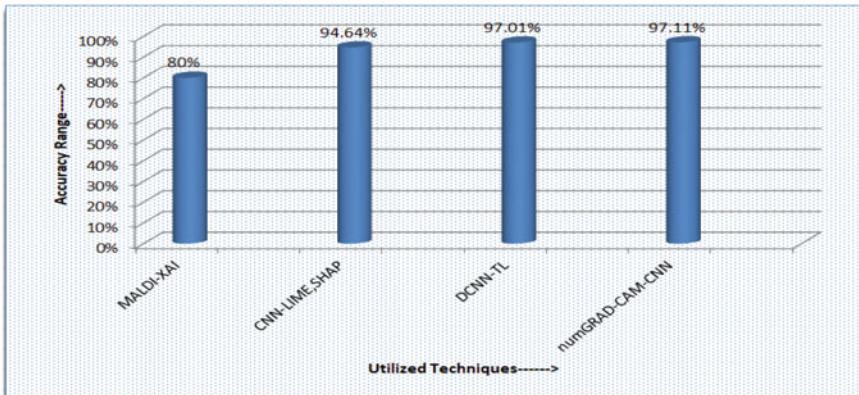


Fig. 3 XAI framework-based accuracy for diagnosing brain tumor

(SHAP) for providing interpretations of decisions made by convolutional neural network and achieved 94.64% accuracy.

Part et al. [23] developed a deep neural network system for extraction of features based on input data following transfer learning process while diagnosing cancer disease and achieved 97.01% accuracy. However, Marolejo et al. [24] proposed numerical gradient-weighted class activation mapping (numGrad-CAM-CNN) for interpreting outcomes of convolutional neural network and achieved 97.11% accuracy while diagnosing brain tumor disorder. However, Montavon et al. [25] discussed various XAI techniques for interpreting deep neural networks, including feature visualization, activation maximization, and saliency maps. They argued that these techniques can provide valuable insight into how deep learning models make predictions and help to identify potential sources of bias or error. Similarly, Doshi-Velez and Kim [26] argued that XAI is essential for building trust in machine learning models. They suggest that by providing insight into how a model works, XAI can help clinicians and researchers make more informed decisions and identify potential areas for improvement.

Therefore, gradient-weighted class activation mapping localized brain MRI region containing tumor to provide significant visualization of features with respect to the each layer of the proposed model for better and accurate interpretation while diagnosing and classifying brain tumor disease in the health care which can further aid radiologist in their decision-making procedures to enhance outcomes of patients.

5 Conclusion

Explainable artificial intelligence (XAI) is an important aspect of developing and deploying deep learning models for brain tumor detection. XAI techniques allow us to understand how a model makes predictions and identify potential sources of bias

or error, improving the accuracy and reliability of the model's predictions. XAI has been explored and discussed by several authors in the field of medical imaging and brain tumor detection. These authors have highlighted the importance of XAI for building trust in machine learning models, providing insight into how models work, and identifying areas for improvement. Overall, applying XAI techniques to deep learning models for brain tumor detection is a promising area of research that has the potential to improve patient outcomes and advance our understanding of brain tumors. However, challenges such as data privacy, complexity of XAI techniques, and lack of standardization need to be addressed in order to ensure the safe and effective deployment of these models in clinical settings. Therefore, further research is needed to refine and standardize XAI techniques, validate their effectiveness in clinical settings, and develop guidelines for their deployment in medical imaging and brain tumor detection.

References

1. Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW (2016) The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica* 131:803–820
2. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
3. Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 35(8):1798–1828
4. Bengio Y et al. (2009) Learning deep architectures for AI, Foundations and trends® in Machine Learning 2(1):1–127
5. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(6088):533–536
6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639):115–118
7. Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2018) Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform* 19(6):1236–1246
8. Ahmed S, Nobel SN, Ullah O (2023) An effective deep CNN model for multiclass brain tumor detection using mri images and shap explainability. In: 2023 International conference on electrical, computer and communication engineering (ECCE), IEEE, 2023, pp 1–6
9. Jin W, Li X, Fatehi M, Hamarneh G (2023) Generating post-hoc explanation from deep neural networks for multi-modal medical image analysis tasks. *MethodsX* 10:102009
10. Kamnitsas K, Ledig C, Newcombe VF, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B (2017) Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 36:61–78
11. Bechelli S (2022) Computer-aided cancer diagnosis via machine learning and deep learning: a comparative review, arXiv preprint [arXiv:2210.11943](https://arxiv.org/abs/2210.11943)
12. Sharma S, Gupta S, Gupta D, Juneja A, Khatter H, Malik S, Bitsue ZK (2022) Deep learning model for automatic classification and prediction of brain tumor. *J Sens*
13. Kukreja V, Ahuja S et al. (2021) Recognition and classification of mathematical expressions using machine learning and deep learning methods. In: 2021 9th International conference on reliability, infocom technologies and optimization (Trends and Future Directions) (ICRITO), IEEE, 2021, pp 1–5
14. Thapa K, Khan H, Singh TG, Kaur A (2021) Traumatic brain injury: mechanistic insight on pathophysiology and potential therapeutic targets. *J Mol Neurosci* 71(9):1725–1742

15. Rehni AK, Singh TG, Jaggi AS, Singh N (2008) Pharmacological preconditioning of the brain: a possible interplay between opioid and calcitonin gene related peptide transduction systems. *Pharmacol Reports* 60(6):904
16. Kamini, Rani S (2023) Artificial intelligence and machine learning models for diagnosing neurodegenerative disorders. In: Data analysis for neurodegenerative disorders, Springer, pp 15–48
17. Ribeiro MT, Singh S, Guestrin C (2016) why should i trust you?" explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp 1135–1144
18. Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. In: Advances in neural information processing systems, pp 30
19. Sundararajan M, Taly A, Yan Q (2017) Axiomatic attribution for deep networks. In: International conference on machine learning, PMLR, 2017, pp 3319–3328
20. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision, pp 618–626
21. Pertzborn D, Arold C, Ernst G, Lechtenfeld OJ, Kaesler J, Pelzel D, Guntinas-Lichius O, von Eggeling F, Hoffmann F (2022) Multi-class cancer subtyping in salivary gland carcinomas with maldi imaging and deep learning. *Cancers* 14(17):4342
22. Gaur L, Bhandari M, Razdan T, Mallik S, Zhao Z (2022) Explanation-driven deep learning model for prediction of brain tumour status using MRI image data. *Front Genet* 448
23. Park KH, Batbaatar E, Piao Y, Theera-Umpoon N, Ryu KH (2021) Deep learning feature extraction approach for hematopoietic cancer subtype classification. *Int J Environ Res Public Health* 18(4):2197
24. Marmolejo-Saucedo JA, Kose U (2022) Numerical grad-cam based explainable convolutional neural network for brain tumor diagnosis. *Mobile Netw Appl* 1–10
25. Montavon G, Samek W, Mu'ller K-R (2018) Methods for interpreting and understanding deep neural networks. *Digital Signal Process* 73:1–15
26. Doshi-Velez F, Kim B (2017) Towards a rigorous science of interpretable machine learning, arXiv preprint [arXiv:1702.08608](https://arxiv.org/abs/1702.08608)

Pioneering a New Era of Global Transactions: Decentralized Overseas Transactions on the Blockchain



Khadeer Dudekula and Annapurani Panaiyappan K.

Abstract Nowadays, it has become harder to transfer money overseas. Overseas transactions have intermediaries and lengthy settlement times, leading to increased remittance taxes due to financial institutions and regulatory requirements. Based on the review, we propose a novel approach using blockchain technology for overseas transactions. By using a decentralized finance application in blockchain technology, we are creating an application that converts the fiat currency which is an actual currency into a stablecoin (USDC). This stablecoin is transferred using blockchain and the user receives it back as a fiat currency. We process overseas transactions with lower remittance taxes and in less duration efficiently and securely.

Keywords Blockchain · Remittance taxes · Decentralized finance · Overseas transaction · Encryption · Stable coin · Decryption · Fiat currency

1 Introduction

For any overseas transactions, there will be taxes imposed by third parties like banks or any other payment methods. The foreign transaction fees would incur charges of around 5%. And this will differ from bank to bank, country to country, etc. Sometimes they would be charged on both the ends of the sender and receiver banks. These charges by third parties have been increasing over the period of time. To minimize the charges imposed by third parties, we are using the concept of blockchain technology [1–3]. Blockchain is a distributed ledger that records information in blocks in a

K. Dudekula

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, India

e-mail: dk2158@srmist.edu.in

P. K. Annapurani

Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur, India

e-mail: annapook@srmist.edu.in

secured manner as a single unaltered chain. The information stored in a blockchain is organized in blocks that are connected together to create an unalterable chain, which is secure, and immutable with no single entity or authority owning the ledger [4].

Usually, a blockchain is mostly used to store the record of transactional information of both sender and recipient. But this is not until 2015; an entrepreneur named Vitalik Buterin introduced a groundbreaking update in blockchain technology by releasing Ethereum blockchain. This Ethereal has smart contracts which are self-executing programs that can automatically enforce the terms of a transaction. Smart contracts in Ethereum offer a great deal of flexibility [5, 6]. There are two types of digital assets in general, namely cryptocurrencies like Bitcoin, Ethereum etc. And the other one is stablecoins like USD coin (USDC), USD Tether (USDT), etc. The main difference lies in their volatility. Generally, stablecoins are very low volatile compared to cryptocurrency unlike cryptocurrencies, and stablecoins are designed to maintain a perfect peg in value to a fiat currency. For instance, USD coin is a stablecoin whose value is pegged to a US dollar i.e.,..., 1 USD coin equals 1 US dollar [7, 8]. Usually, cryptocurrencies are used for trading, so there will be profits or losses depending on the demand. But these stablecoins are actually non-profitable because their value is designed to be equal to USD value. So, the main purpose of these stablecoins is not for trading but for buying other cryptos. Not all cryptos are universally available for trading in every country. People cannot buy these unavailable cryptos with their native currency, but people may show interest to trade or to stake different types of cryptos, expecting that those crypto prices would increase over time. So, in this case, it is difficult for the people to buy crypto which is not available in their native region. Here is where the stablecoins can come in handy. Since stablecoins are non-profitable as they do not show any rapid increments or decrements in value over time like cryptos, there is no point in restricting or making them unavailable. So stablecoins such as USDT and USDC are widely spread throughout the world in almost all countries, and people first buy these stablecoins and then use these stablecoins to buy other cryptos which are not available in their native region without any price volatility and additional fees. So, using this type of digital asset, stablecoins whose value is designed to maintain a pegged value to other fiat currencies such as the US dollar could be an optimal solution in overseas transactions because of their low volatility.

2 Existing Solution

Usually for any overseas transactions, people generally use payment methods like PayPal, international wire transfer, foreign currency drafts, etc. Among all international transactions, wire transfer is the most common method.

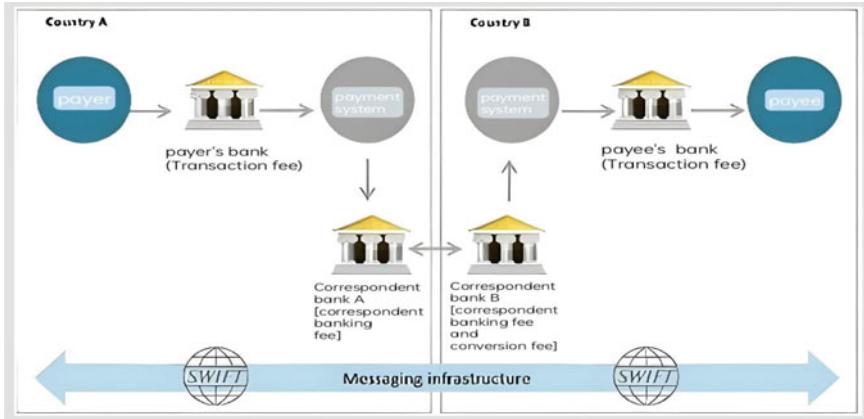


Fig. 1 Internal wire transaction system via SWIFT

2.1 International Wire Transfer

Bank-to-bank transactions are subjected to the constraints imposed by the underlying financial infrastructure as shown in Fig. 1. Due to the intricate network of intermediaries involved in settling a standard transfer, the process typically takes around three days to complete. In the case of a transfer from a State Bank of India (SBI) account to a Wells Fargo account in the USA, the transaction must traverse the SWIFT system, which entails the transaction of 33.7 million messages to more than eleven financial institutions on daily basis [9]. These payment orders are sent by the centralized SWIFT system which does not actually transmit the money. Actual money is processed through a network of intermediates as shown in Fig. 1. Each middleman would increase the transaction cost furthermore to 60% as shown in Fig. 2 of bank-to-bank payments [10].

These payment orders are sent by the centralized SWIFT system which does not actually transmit the money. Actual money is processed through a network of intermediates as shown in Fig. 1. Each middleman would increase the transaction cost furthermore to 60% as shown in Fig. 2 of bank-to-bank payments [11].

2.2 Transactions via Cryptocurrency

While it is an option to send money overseas using crypto, this method is not widely used by most people. This is likely due to the high volatility of cryptocurrencies, which can result in a loss of value for the recipient. Additionally, not everyone is comfortable using digital currencies, further limiting their adoption as a mainstream method of transferring funds. If we look at the international wire transfer, this method

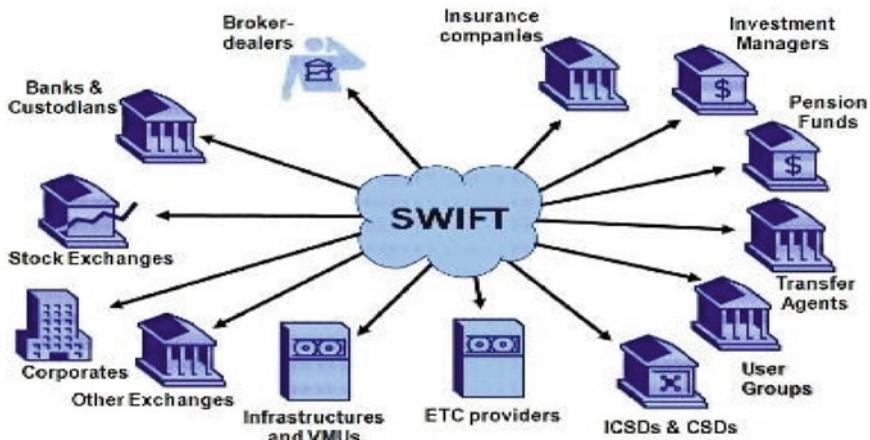


Fig. 2 Transaction system through SWIFT

comes with many potential drawbacks such as low transaction speed, higher remittance taxes, and involvement of third parties which make overseas transactions more complicate.

On the other hand, transaction via cryptocurrencies also comes with potential drawbacks for an instance, and Ripple is a company which transfer funds in form of XRP. XRP tokens are cryptocurrencies just like Bitcoin, Ethereum, etc., which are highly volatile for transaction. Moreover, Ripple is currently involved in a legal dispute with the US Securities and Exchange Commission (SEC) that dates back to 2020. This battle stems from Ripple's decision to release digital tokens onto market without registering with the SEC. As a result, there are concerns about the regulatory status of XRP and the potential implications for Ripple's operations. As a result of ongoing legal dispute with SEC, several major crypto exchanges such as Coinbase, Binance etc. have limited the XRP availability on their platforms [12].

3 Proposed Solution by Conversion of Fiat Currency

As we know that stable coins are more stable and less volatile, this would work as a base line in our solution to transfer funds overseas with less remittance charges.

3.1 A Unified Payment Interface Decentralized Finance App Works Globally

Unified Payment Interface or simply UPI is a real-time payment system that facilitates inter-bank transactions in India. So, UPI is only restricted to India. But there are some other apps like WeChat pay which have similar features to UPI, although it does not facilitate inter-banks in India, but instead it facilitates inter-banks in China. The main concept of our solution is to create a Decentralized Finance (DeFi) Unified Payment Interface app that works globally [14], which would be a messaging and settlement system designed to facilitate real-time cross-border payments.

- As of 7 May, 2023, 1INR = 0.012USD and 16,00,000INR = 19,576USD.
- And 1USD = 1USDC (a stablecoin), 19,576 USDC = 19,576 USD.

In an actual blockchain network when someone initiates a transaction, this transaction is broadcasted to all the nodes in blockchain for the validation. Nodes are individual digital devices that maintain a copy of the blockchain ledger and participate in the validation. Validation is the process in which the nodes check whether the transaction is valid or not. It also checks whether the sender has enough funds for the payment to perform a valid transaction. After the node validates a transaction, it will be broadcasted to other nodes on the network, and when a majority of the nodes validated the transaction, then it is confirmed and added to the blockchain [15]. After the validation is completed, this transaction will be added to the blockchain.

When a sender initiates a transaction in our proposed DeFi application, firstly the transaction is sent for a validation process. After the validation is done successfully, the sender's funds which are in fiat currency will be converted into a stablecoin. After the conversion, these stablecoins are moved through the blockchain to their destination and all these information will be stored in the blockchain database. This database is secured using cryptographic algorithms which ensure integrity of the ledger. When this transaction reaches the recipient, they use their private key to decrypt the transaction and access their funds in their wallet.

For instance, let us assume that Alice from India sends 16,00,000 INR to her friend Bob who lives in the USA at 1 pm. The transaction begins in the first phase with Alice sending funds, firstly the validation process happens, then after the process is completed and added to the blockchain ledger, these funds are used to obtain the equivalent amount of stablecoins in her wallet which is equal to 19,576 USDC at 1 pm, after that this 19,576USDC are moved over the blockchain till its destination, and at final phase, these 19,576USDC stablecoins are again used to obtain the equivalent amount of fiat currency that is 19,576USD at 1 pm exchange rate and credited into Bob's UPI wallet.

All these exchange rates and other information would be noted on the blockchain. This debiting and crediting of money in the user's wallet are similar to trading of digital assets. If you (who is from India) buy any digital assets like USDC, you would buy this with your native currency INR, not with USD or any other currencies, and when you sell your USDC to an US citizen, he would buy your USDC with US

Dollar, the money would be credited to your wallet in your native currency only not in US dollar, same is happening in our mechanism described above, the value of asset is not changing, and only currency is changed.

By interoperating this entire mechanism into an UPI-DeFi app, we can make the transaction process simple and user friendly. However, users are charged an additional fee for the validation and smart contracts' running processes, which are charged by the block chain network; in our case, we choose Ethereum, which is quite low compared to actual remittance taxes to users.

Overview of Proposed Solution

Starting from 1 April, 2022, the Indian government has imposed taxes on cryptocurrencies. One is a 30% tax, which will be charged when a person receives any profits on crypto trading, and another tax is a 1% TDS tax, which is imposed on various order types like selling non-INR crypto pairs, etc. [16–19]. Most of these taxes are charged when we trade any digital assets, on capital gains. In our solution, we are converting the fiat currency into stablecoins, which is nothing but buying the digital assets and moving them over the blockchain till their destination. In most of the countries, taxes are only imposed on capital gains on digital assets but not on buying and transferring them between people. Transferring crypto on the blockchain causes gas prices for the sender, gas prices are the additional fees charged by the blockchain which were very less compared to taxes imposed by third parties, and by using some blockchains like stellar block chain, the gas prices would be in cents per transaction.

Some stablecoins are backed by commodities like physical Gold instead of actual fiat currencies. DigixDAO (DGX), USD coin (USDC) Gold, and Paxos Gold (PAXG) are some examples of stablecoins that are backed by commodities. Using this kind of commodities' backed stablecoins for transactions has some advantages. Commodity-backed stablecoins have high intrinsic value, price stability, and global acceptance compared to fiat currency-backed stablecoins. Moreover, we can use any type of stable coin for a transaction, but using these commodity-backed stablecoins would provide better transparency. In this scenario, users can reduce remittance taxes imposed by the third parties and complete transactions within minutes, instead of days.

In present digital age, banking is like fortress guarding treasure of mammoth properties. These properties are prime target for cyber villains. To tackle this issue, government has cracking down and financial institutions are exploring new avenues such as block chain technology. The main advantage of these blockchain transactions over bank transactions is that blockchain transactions will run and operate without a single point of failure. All transactions are peer-to-peer settlements. Unlike banks, block chain is 24/7 operational. In banks multiple intermediaries and points of failure are involved in settling an order. The process of transferring ownership in traditional system is intricate, as each party maintains their own ledger with their own perspective of actuality (refer Fig. 4). The current system is inefficient and imprecise, as each party involved must update and reconcile their records at the end of the day. In addition to providing securities the transaction gets completed within 3 days [20, 21].

On average due to the involvement of intermediaries, it takes up to three days for overseas transactions processed through Visa to settle [22]. As shown in Fig. 4 because of its centralized structure, each middlemen will take their time to update and reconcile their books at end of every day because each party maintains their own ledger, which results in inefficient system, whereas transactions on blockchains like Ethereum would take anywhere between 15 s and 5 min [23]. So, transactions on blockchain are faster than traditional banking system.

Now let us see transaction fees. Figure 3 shows the transactions between Alice and Bob. Alice sends 16,00,000INR to Bob (who lives in the USA).

In bank-to-bank transaction, the amount incurred as transaction fees would be 46,825INR. The calculations are shown below.

In India there are many methods for overseas transactions, but SWIFT method is famous among all, as it has highest number of transactions (over 37 million per day). If Alice sends 16,00,000INR via SWIFT from India, there are some taxes which will be charged commonly known as remittance taxes. Up to 7,00,000INR there are no taxes charged, but beyond there are 5% taxes on each 1,00,000INR and there are some additional taxes imposed by Government of India, they are 1500INR + 18% GST on 100INR, and this 1500INR might vary from bank to bank, but approximately this entire additional tax would be anywhere around 1825INR (1500 + 18%1500).

Let us calculate these all taxes on 16,00,000INR.

Firstly 5% tax would be on 16,00,000 is: $16,00,000 - 7,00,000 = 9,00,000$ INR.

$5\% \text{ of } 9,00,000 = 45,000$ INR is tax and including those additional taxes = 1825INR.

Total transaction cost is $45,000 + 1825 = 46,825$ INR. As of 7 May, 2023, this is equal to 572.92USD.

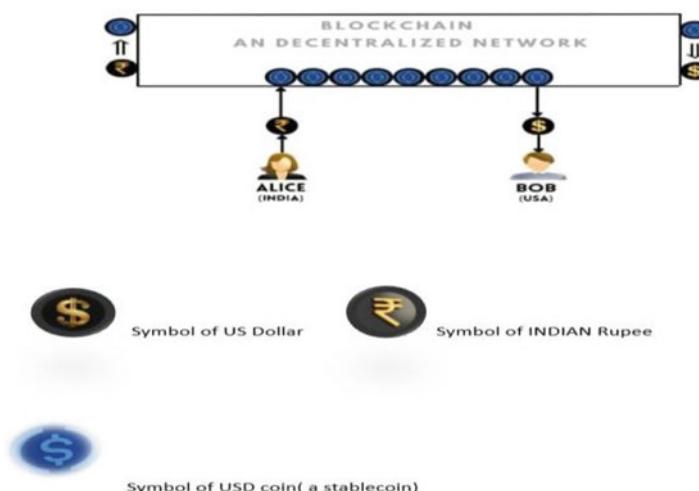


Fig. 3 User flow of proposed decentralized finance application app

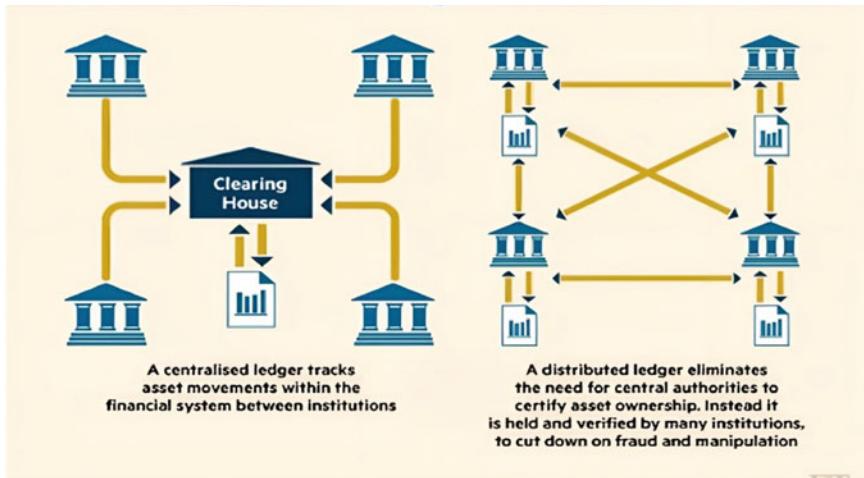


Fig. 4 Centralized and decentralized asset tracking methods

So, to send 16,00,000INR (19,576USD), Alice has to pay 46,825INR (572.92USD) as remittance fees.

Now let us look at the transaction charges if Alice chooses Defi-UPI—the app we proposed, where the transaction process will occur as we described above and that runs Ethereum Blockchain. With the London fork update in Ethereum, changes were made to the transaction fee process. This update is the latest as of 7 May, 2023.

If Alice sends 16,00,000 INR to Bob in this DeFi app, in Ethereum all the transaction fees will be charged and this fee was measured in “Gwei”. In this London update, we need to give some limits for parameters like base fee, tip fee. In Ethereum, the base fee is the minimum fee that the user must pay to have their transactions validated and this price would be calculated based on the current demand for block space on the network. Users can always choose the average base fee to process their transactions. Then, the other is tip, which is a kind of giving reward to miners for processing our transaction, and we can give as much as we want as a reward but this is optional to give. By giving this tip fee to miners, we can attract the miners to validate our transaction even faster.

Let us say the base fee (as of 7 May, 2023) 150 Gwei and the tip is 5 Gwei, additionally, if we use any smart contracts to run in our payment process let us say this additional fee would be 100Gwei. As of 7 May, 2023, all these base prices, tips, and additional smart contract execution fees were taken highest in amount, instead of taking average or lowest in amount to show the maximum amount that Alice has to pay as fees, so depending on network coagulation, final price that Alice has to pay would be less than the total transaction cost, but definitely not more than this transaction cost. As per parameters

$$\text{Basefee} = 150 \text{ Gwei},$$

Tip = 5 Gwei,
 Gaslimit = 21,000 Gwei,
 Additionalfee(forsmartcontract) = 100 Gwei,
 So total Gast price = 255 Gwei(150 + 5 + 10),
 Gascost = gasprice*gaslimit
 = 255*21,000
 = 5,355,000 Gwei
 = 0.005355 ETH.

Therefore, the transaction cost is 0.03381ETH which is equal to 9.46USD (777.16 INR)

So, at last here Alice has to pay less than 9.46USD (777.16INR) instead of 572.92USD (46,825INR). By this, we can observe that Alice was able to save nearly 98.34% of her money by utilizing blockchain transactions, which have significantly lower costs compared to traditional bank transactions [24–27]. We can also check the current base fee, tip fee (or priority fee), and the total cost happened per transaction on Ethereum network [28]. All the current transaction fees happened on 7 May, 2023, are less than our calculated value.

If we consider a blockchain-based UPI works globally, the use of digital assets brings numerous advantages. Some of them are

- A transaction using these methods results in reducing charges for overseas remittance charges, as they eliminate the need for third parties.
- Utilizing a blockchain-based UPI-like app enhances transparency in transactions, allowing for easier tracking and resolution of any potential fraud.
- Transactions would be processed with high speed, with settlements completed within minutes or even seconds depending on block traffic, resulting in efficient transaction processing.

Despite the numerous advantages, here are some drawbacks to consider:

- All taxes cannot be eliminated, and since we are using blockchain for transactions, there are some charges like gas fees, which are fees charged for usage of blockchain. Actually, this fee is very low compared to remittance charges by third parties; by using blockchains like stellar, we can limit the gas fees to cents per transaction.
- Another drawback is that blockchain changes are irreversible. This means let us say that if Alice wants to send her funds to Bob, instead she sends her funds to someone, this is irreversible. This transaction cannot be canceled or reversed. So, while sending funds to someone, users should cross-check whether they are sending to the correct address or not. Well never-the-less this mistaken transaction will be recorded on a blockchain, if needed we can track the transaction destination address, but it will depend on the receiver's decision whether to return the Alice money back or not.

- May not have enough liquidity for digital assets. However, this issue can be addressed by ensuring that the stablecoin is backed by reserves of reliable fiat currency such as USD, held in segmented accounts, and subjected to regular audits by reputable accounting firms.

4 Conclusion

In this unified payment interface-DeFi app, the fiat currency is converted into a stablecoin and moving this stablecoin over the blockchain till its destination would be an efficient solution to reduce any remittance fees on overseas transactions. As the blockchains are actually designed to be decentralized, from sender to receiver, each and every step is recorded on a block, which is immutable, and able to process our transactions within seconds or utmost within a few minutes, with very low transaction charges. We can complete any overseas transactions without depending on any third parties in the most efficient and optimal way possible.

References

1. Pullanoor H (2023) A new 20% tax makes it more expensive to move money out of India, retrieved from Quartz
2. Crail C (2023) Foreign transaction fees, retrieved from forbes Advisor
3. Everything You Need To Know About Tax on International Money Transfer, retrieved from Jupiter (2022)
4. Massimo, Piero D (2017) What is the blockchain, IEEE Xplore
5. Zibin Z, Xie S, Dai H-N, Chen W, Chen X, Weng J, Imran M (2020) An overview on smart contracts: challenges, advances and platforms. Future Generat Comput Syst 105:475–491
6. What is Ethereum, retrieved from <https://ethereum.org/en/what-is-ethereum/>
7. Rosen A (2023) What is cryptocurrency? Retrieved from retrieved from NerdWallet
8. Hayes A (2022) Stablecoins: definition, how they work, and types, retrieved from in- estopedia
9. Luca Fantacci and Lucio Gobbi, Stablecoins, Central Bank Digital Currencies and US Dol- lar Hegemony, retrieved from google schola(2020)
10. Internal wire transaction system via SWIFT, retrieved from AitaNovarica
11. Ideal structure of SWIFT system retrieved from Ideal structure of SWIFT system retrieved from smartencyclopedia.org
12. SEC Charges Ripple and Two Executives with Conducting \$1.3 Billion Unregistered Securities Offering, U.S Securities and Exchange Commissions (2020) retrieved from <https://www.sec.gov/news/press-release/2020-338>
13. Convert USD Coin to US Dollar retrieved from <https://www.coinbase.com/converter/usdc/usd>
14. What is DeFi? Coinbase retrieved from <https://www.coinbase.com/learn/crypto-basics/>
15. Sahu L, Sharma R, Sahu I, Das M, Sahu B, Kumar R (2021) Efficient detection of Parkinson's disease using deep learning techniques over medical data. Expert Syst 12787. <https://doi.org/10.1111/exsy.12787>
16. Sharma R, Kumar R, Sharma DK et al (2021) Water pollution examination through quality analysis of different rivers: a case study in India. Environ Dev Sustain. <https://doi.org/10.1007/s10668-021-01777-3>

17. Ha DH, Nguyen PT, Costache R et al (2021) Quadratic discriminant analysis based ensemble machine learning models for groundwater potential modeling and mapping. *Water Resour Manage.* <https://doi.org/10.1007/s11269-021-02957-6>
18. Dhiman G, Sharma R (2021) SHANN: an IoT and machine-learning-assisted edge cross-layered routing protocol using spotted hyena optimizer. *Complex Intell Syst.* <https://doi.org/10.1007/s40747-021-00578-5>
19. Sharma R, Gupta D, Polkowski Z, Peng S-L (2021) Introduction to the special section on big data analytics and deep learning approaches for 5G and 6G communication networks (VSI-5g6g). *Comput Electri Eng* 95:107507. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2021.107507>
20. Singh PD, Dhiman G, Sharma R (2022) Internet of things for sustaining a smart and secure healthcare system. *Sustain Comput: Inform Syst* 33:100622. ISSN 2210–5379. <https://doi.org/10.1016/j.suscom.2021.100622>
21. Sharma R, Arya R (2021) A secure authentication technique for connecting different IoT devices in the smart city infrastructure. *Cluster Comput.* <https://doi.org/10.1007/s10586-021-03444-8>
22. Sharma R, Arya R (2021) Secure transmission technique for data in IoT edge computing infrastructure. *Complex Intell Syst.* <https://doi.org/10.1007/s40747-021-00576-7>
23. Rai M, Sharma R, Satapathy SC et al (2022) An improved statistical approach for moving object detection in thermal video frames. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11548-x>
24. Rasika V, Rohit S (2022) Dual notched conformal patch fed 3-D printed two-port MIMO DRA for ISM band applications. *Frequenz.* <https://doi.org/10.1515/freq-2021-0242>
25. Sharma N, Sharma R (2022) Real-time monitoring of physicochemical parameters in water using big data and smart IoT sensors. *Environ Dev Sustain.* <https://doi.org/10.1007/s10668-022-02142-8>
26. Anandkumar R, Dinesh K, Obaid AJ, Malik P, Sharma R, Dumka A, Singh R, Khatak S (2022) Securing e-Health application of cloud computing using hyperchaotic image encryption framework. *Comput Electri Eng* 100:107860. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2022.107860>
27. Sharma R, Xin Q, Siarry P, Hong W-C (2022) Guest editorial: deep learning-based intelligent communication systems: using big data analytics. *IET Commun.* <https://doi.org/10.1049/cm2.12374>
28. Sharma R, Arya R (2022) UAV based long range environment monitoring system with Industry 5.0 perspectives for smart city infrastructure. *Comput Indus Eng* 168:108066. ISSN 0360-8352. <https://doi.org/10.1016/j.cie.2022.108066>

A Perspective Review of Generative Adversarial Network in Medical Image Denoising



S. P. Porkodi^{ID} and V. Sarada^{ID}

Abstract This review article discusses using Generative Adversarial Networks (GANs) for image denoising in medical images. GANs have produced optimistic results in enhancing the quality of noisy medical images, which is crucial for accurate diagnosis and treatment. The article provides an overview of the challenges in medical image denoising and the working principle of GANs. The review also summarizes the recent research on using GANs for medical image denoising and compares their performance and significance. Finally, the article discusses the future directions for GAN-based medical image denoising and its potential impact on the healthcare industry.

Keywords Generative adversarial network (GAN) · Structural similarity index (SSIM) · Peak signal-to-noise ratio (PSNR) · Convolutional neural network (CNN)

1 Introduction

Medical image denoising is crucial in improving the quality and accuracy of medical images. Medical images, which include X-rays, Computed Tomography (CT) scans, Magnetic Resonance Imaging (MRI), and ultrasound imaging, are frequently affected by noise, which can decrease the visibility of significant structures and features in the images. Medical image-denoising techniques aim to reduce noise in medical images while preserving important diagnostic information [1]. These techniques can improve the accuracy of medical diagnoses and reduce the need for additional tests

S. P. Porkodi · V. Sarada (✉)

Department of Electronics and Communication Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, 603203 Chengalpattu, Tamil Nadu, India

e-mail: saradav@srmist.edu.in

S. P. Porkodi

e-mail: ps4195@srmist.edu.in

or procedures. Some specific benefits of medical image denoising include improved image quality, more accurate diagnoses, and reduced radiation exposure.

The traditional method deals with noise concerning various filtering techniques. One of the most popular traditional denoising methods is Discrete Wavelet, with fast calculation and simple architecture. Other methods include algorithms based on nonlocal filters [2], and Block-Matching and 3D filtering (BM3D) [3] these methods have do's and dont's. Convolutional Neural Networks (CNNs) [4] have been effectively used for image-denoising problems to overcome this traditional method. CNNs are powerful machine learning models that automatically learn complex image features from the input data, making them ideal for image-denoising applications. CNNs can effectively capture local image features, such as edges, textures, and patterns, which are important for image denoising. They can also handle nonlinear relationships between noisy and clean image pixels, making them more effective than traditional linear filters. One popular CNN architecture for image denoising is the denoising autoencoder (DAE) [5], which consists of a decoder and an encoder. The encoder considers the noisy image input and produces a low-dimensional feature representation, while the decoder restructures the clean data from the low-dimensional feature representation. Other CNN architectures, such as U-Net and Deep Residual Network (ResNet) [6], were also explored to image denoise with favorable results. The drawback is that CNNs try to remove noise from the input image directly, which may lead to the loss of important details and features, they rely on a fixed set of filters that may be unable to identify the intricate relationships between the noisy and clean images, and finally, CNNs require a large dataset to learn the complex features and patterns in the data. So, GANs can overcome some of the limitations of CNNs for medical image denoising by generating synthetic images similar to real images and learning the clean images underlying distribution. First, GANs can generate synthetic images that look identical to real images, which can be utilized as a substitute for noisy images. Secondly, GANs can learn the underlying distribution of clean medical images, which can help to generate more accurate and visually appealing images. Thirdly, GANs can be trained on a relatively small dataset, since it is important for medical image denoising as collecting many clean medical images is challenging and time-consuming.

This paper is structured in the following ways: Sect. 2 describes the related works. Section 3 discusses the various types of denoising methods using GAN. Section 4 discusses different performance metrics used. Section 5 describes the significance of image denoising using GAN. Section 6 discusses conclusion and future scope.

2 Related Works

Image denoising refers to reducing or removing unwanted information from an image although preserving its important details and structures. Noise in an image is typically caused by various factors such as sensor limitations, transmission errors, low-light conditions, or compression artifacts. Image-denoising techniques aim to improve the

image's visual quality by reducing unwanted noise components. Image denoising aims to balance noise reduction and preserve important image features. The process involves analyzing the image data to distinguish between the noise and the underlying image content. Denoising algorithms can reduce or suppress noise components while retaining the essential image details by identifying noise components.

Different image-denoising methods utilize various approaches and algorithms to achieve noise reduction. Here are some common methods used in denoising an image: According to Fan et al. [7], filtering techniques apply specific filters to the image to eradicate the noise components. For example, linear filters such as mean or Gaussian filters smooth out the image by averaging neighborhood pixels. Nonlinear filters including median or bilateral filters effectively preserve edges while reducing noise.

Ben et al. [8] proposed the statistical methods to model the noise and image signals statistically to estimate the clean image from the noisy input. These methods often assume specific probability distributions for the noise and the image. For instance, Gaussian noise is commonly modeled, and algorithms like Wiener filtering or Bayesian estimation estimate the clean image based on the statistical properties.

Donoho et al. [9] discussed the transform-based denoising methods exploit the signal properties in transformed domains such as the Fourier, wavelet, or sparse representations. The image with noise is transformed into the region where the noise is less correlated or concentrated, allowing for better noise and image content separation. Denoising is performed by thresholding or modifying the transformed coefficients and then returning the result to the spatial domain.

Alkinani et al. [10] proposed patch-based denoising techniques that consider local image patches instead of individual pixels. Similar patches are grouped together, and denoising is performed by averaging or adaptively combining the information from these patches. Patch-based methods can effectively preserve image structures and textures while reducing noise.

Tian et al. [11] proposed deep learning approaches that have gained popularity in recent years for image denoising. CNNs are trained on huge dataset to differentiate between noisy and clean images. These models can effectively capture complex noise patterns and learn to reconstruct clean images.

Dey et al. proposed that image denoising using GANs leverages the power of adversarial learning to generate visually appealing and realistic denoised images. Here are some key roles and benefits of using GANs for image denoising [12]: Preserving Image Details, Capturing Complex Noise Patterns, Handling Uncertainty, Enhancing Visual Quality, Handling Non-Stationary Noise, Learning from Large-Scale Data and also Extensibility to Other Image Restoration Tasks. Image denoising using GANs offers advanced noise reduction capabilities, the ability to preserve image details, and the generation of visually appealing and realistic denoised images.

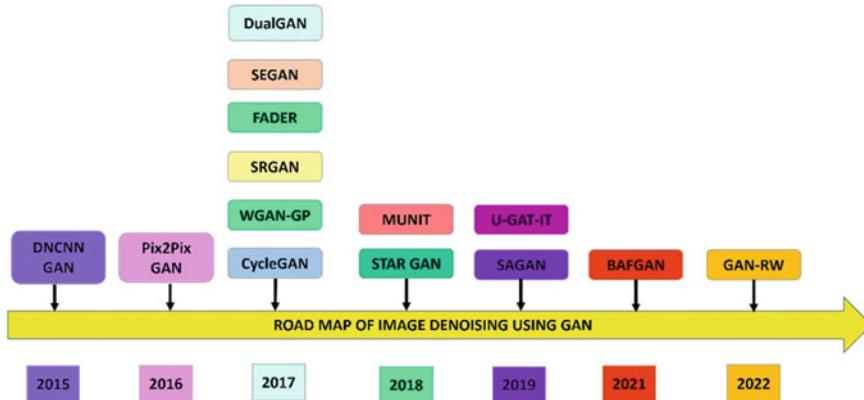


Fig. 1 Road map of image denoising using GAN

3 Various Types of Image-Denoising Methods Using GAN

GANs [13] are a deep learning (DL) method that can denoise an image. Figure 1 shows the road map of image denoising using GAN. The simple concept behind GAN-based denoising is to train a generator network to convert noisy images into noise-free images while instantaneously training a discriminator network to differentiate between generated and authentic clean images. Figure 2 demonstrates the dissimilarity between the traditional and GAN-based image-denoising methods. GAN-based denoising involves training the networks on a noisy and clean image pair dataset to acquire a mapping from noisy to clean images. The noisy images can be artificially generated by incorporating noise into the clean images using a known noise model or obtained from real-world noisy image data. Once trained, the GAN can denoise new noisy images by transmitting them through the network of generators. Let us discuss each denoising GAN. The various types of denoising methods are shown below in Table 1.

4 Performance Metrics

Performance metrics evaluate the characteristics of the denoised images created by a GAN-based denoising model [28]. Some commonly used metrics that include:

Peak Signal-to-Noise Ratio (PSNR): This metric measures the highest probable pixel value ratio to the MSE between the denoised and true clean images. The maximum value of PSNR indicates better denoising performance. The PSNR is given by

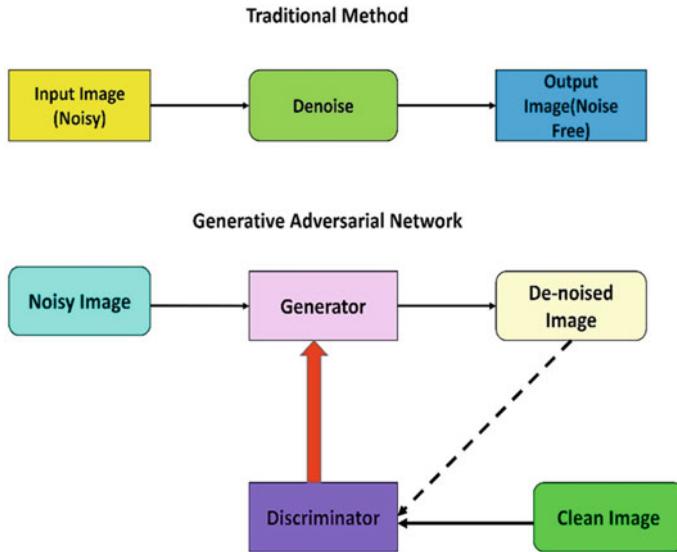


Fig. 2 Difference between traditional and GAN-based image-denoising methods

$$\text{PSNR} = 10 \log_{10} \left(\frac{(L - 1)^2}{\text{MSE}} \right). \quad (1)$$

In this case, L represents the number of the highest possible intensity level (low-intensity value corresponds to zero) in an image.

Structural Similarity Index (SSIM): This metric compares denoised and generated images concerning luminance, structure, and contrast. Higher SSIM values indicate better denoising performance. The SSIM is represented by

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (2)$$

Here, x and y represent a comparison of two images. μ_x and μ_y represent the average luminance values of x and y , respectively. σ_x and σ_y are the standard deviations of x and y . σ_{xy} represents the covariance concerning x and y . The constants are c_1 and c_2 , stabilizing the division when the denominator is close to zero.

Mean Squared Error (MSE): A common metric that estimates the average squared difference between the approximated and true values in a regression problem. The MSE is denoted by:

$$\text{MSE} = \left(\frac{1}{n} \right) * \sum (y_i - \hat{y}_i)^2. \quad (3)$$

Table 1 Various types of denoising methods

Author/ year	GAN type	Method	Advantage	Disadvantages
Zhang et al. [14]	DnCNN-GAN	Hybrid	Improves performance of DnCNN for denoising	May suffer from mode collapse
Isola et al. [15]	Pix2Pix	Conditional	Generates high-quality images with fine details	Requires paired noisy and clean images for training
Zhu et al. [16]	CycleGAN	Unconditional	Can learn mappings between unpaired images	Generates images with less fine details compared to Pix2Pix
Gulrajani et al. [17]	WGAN-GP	Wasserstein	Stable training with strong theoretical guarantees	Requires more training time compared to other GANs
Ledig et al. [18]	SRGAN	Super resolution	Generates high-resolution images from low resolution images	May introduce artifacts in the generated images
Lample et al. [19]	FADER	Disentangled	Learns disentangled representations for attributes and content	Requires paired noisy and clean images for training
Pascual et al. [20]	SEGAN	Speech	Improves speech quality and removes noise	May not perform well on certain types of noise
Yi et al. [21]	DualGAN	Dual	Improves translation accuracy by using two generators and discriminators	Improves translation accuracy by using two generators and discriminators
Choi et al. [22]	StarGAN	Multi-Domain	Can generate images with different styles and attributes	May lead to mode collapse or bias toward certain domains
Huang et al. [23]	MUNIT	Unsupervised	Learns disentangled representations for style and content	Requires large amounts of training data
Zhang et al. [24]	SAGAN	Self-attention	Captures long-range dependencies and improves image quality	Requires more computational resources compared to other GANs
Kim et al. [25]	U-GAT-IT	Unsupervised	Can generate diverse images without paired training data	May generate images with the low visual quality

(continued)

Table 1 (continued)

Author/ year	GAN type	Method	Advantage	Disadvantages
Marcos et al. [26]	BAFGAN	LDCT	Enhance the PSNR of an images	Requires more computational resources compared to other GANs
Zhang et al. [27]	GAN-RW	Ultrasound	Reduce speckle noise	Requires more computational resources compared to other GANs

In the formula, y_i represents the actual or observed data, \hat{y}_i represents the approximated value, and n corresponds to the total no. of data points. The difference between each actual value and its corresponding approximated value is squared, summed up, and then divided by the total no. of data points to calculate the average squared error.

MSE is broadly used in machine learning and statistics as an objective function to estimate the performance of regression models. A lower MSE value indicates better model performance, representing a smaller average squared difference between the expected and actual values.

Mean Absolute Error (MAE): In regression issues, a statistic is employed. To calculate the average absolute dissimilarity between the approximated and true values in a dataset. It is often used to calculate the description of denoising algorithms or models. The MAE is denoted by:

$$\text{MAE} = \left(\frac{1}{n} \right) * \sum |y_i - \hat{y}_i|. \quad (4)$$

In the formula, y_i represents the actual or observed value, \hat{y}_i represents the approximated value, and n corresponds to the total no. of data values. The absolute difference between each actual value and its corresponding approximated value is calculated, summed up, and then divided by the total no. of data to obtain the average absolute error. MAE calculates the average number of errors without taking into account their direction. It is less sensitive to outliers than other error metrics such as Mean Squared Error (MSE) because it does not involve squaring the errors. A lower MAE value represents better model performance, representing a smaller average absolute dissimilarity between the expected and actual values.

Visual Inspection: This is a qualitative evaluation method in which experts visually inspect the denoised images and compare them to the true clean images. This metric can provide additional insights into the excellence of the denoised images that the quantitative metrics may not capture.

5 Significance of Image Denoising Utilizing GAN

Image denoising using GANs has significant practical applications in various fields, such as medical imaging, surveillance, and satellite imagery. Here is some significance of image denoising using GANs:

- GANs have developed a prevalent tool for image denoising for their ability to learn complex distributions and generate high-quality samples.
- Image denoising is an important task in image analysis and computer vision, and it has many practical applications in various fields such as medical diagnosing, surveillance, and satellite imagery.
- GAN-based image denoising has shown great potential in addressing the challenging task of removing complex noise patterns in various types of images, especially for medical images.
- GAN-based image denoising has been demonstrated to produce visually pleasing and perceptually convincing results and has shown to outperform traditional denoising methods.

6 Conclusion

Medical image denoising is a serious task in medical image analysis, as the excellence of the image affects the accuracy of diagnosis and treatment planning. Over the years, various methods have been suggested for medical image denoising, including traditional methods such as wavelet and diffusion-based methods, and more recently, DL-based methods such as CNNs and GANs.

GAN-based image denoising has shown significant promise in removing complex noise patterns in medical images, producing visually pleasing and perceptually convincing results. Furthermore, GAN-based methods have outperformed traditional denoising techniques concerning visual quality and quantitative metrics.

However, despite the significant improvement in this area, several difficulties must be addressed, such as the lack of large-scale annotated datasets, generalization across different imaging modalities, and interpretability of the learned models. Future research in this area should address these challenges further to increase the accuracy and reliability of medical image denoising.

References

1. Sagheer SVM, George SN (2020) A review on medical image denoising algorithms. *Biomed Signal Process Control* 61:102036
2. Mondal, Maitra M (2014) Denoising and compression of medical image in wavelet 2d. *Int J Recent and Innov Trends in Comput Commun* 2(2):1–4

3. Dabov K, Foi A, Katkovnik V, Egiazarian K (2006) Image denoising with block matching and 3d filtering. In: *Image processing: algorithms and systems, neural networks, and machine learning*, vol 6064. SPIE, pp 354–365
4. Thakur RS, Chatterjee S, Yadav RN, Gupta L (2023) Medical image denoising using convolutional neural networks. In: *Digital image enhancement and reconstruction*, Elsevier, pp 115–138
5. Chai Y, Liu H, Xu J, Samtani S, Jiang Y, Liu H (2023) A multi-label classification with an adversarial-based denoising autoencoder for medical image annotation. *ACM Trans Manag Inf Syst* 14(2):1–21
6. Gurrola-Ramos J, Dalmau O, Alarcón TE (2021) A residual dense u-net neural network for image denoising. *IEEE Access* 9:31742–31754
7. Fan L, Zhang F, Fan H, Zhang C (2019) Brief review of image denoising techniques. *Visual Comput Indus Biomed Art* 2(1):1–12
8. Ben Hamza A, Krim H (2001) Image denoising: a nonlinear robust statistical approach. *IEEE Trans Signal Process* 49(12):3045–3054
9. Donoho DL, Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. *J Am Stat Assoc* 90(432):1200–1224
10. Alkinani MH, El-Sakka MR (2017) “Patch-based models and algorithms for image denoising: A comparative review between patch-based images denoising methods for additive noise reduction.” *EURASIP J Image Video Process* 2017(1):1–27
11. Tian C, Fei L, Zheng W, Xu Y, Zuo W, Lin, C.-W (2020) “Deep learning on image denoising: An overview.” *Neural Netw* 131:251–275
12. Dey R, Bhattacharjee D, Nasipuri M (2020) “Image denoising using generative adversarial network.” *Intell Comput: Image Process Based Appl*, 73–90
13. Porkodi SP, Sarada V, Maik V, Gurushankar, K, (2022) “Generic image application using gans (generative adversarial networks): A review.” *Evol Syst*, 1–15
14. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising.” *IEEE Trans Image Process* 26(7):3142–3155
15. Isola P, Zhu J.-Y, Zhou T, Efros AA (2017) “Image-to-image translation with conditional adversarial networks.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134
16. Zhu J-Y, Park T, Isola P, Efros AA (2017) “Unpaired image-to-image trans-lation using cycle-consistent adversarial networks.” In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232
17. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) “Improved training of wasserstein gans.” *Adv Neural Inf Process Syst* 30
18. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z et al (2017) “Photo-realistic single image super-resolution using a generative adversarial network.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681–4690
19. Lample G, Zeghidour N, Usunier N, Bordes A, Denoyer L, Ranzato M (2017) “Fader networks: Manipulating images by sliding attributes.” *Adv Neural Inf Process Syst* 30
20. Pascual S, Bonafonte A, Serra J, (2017) “Segan: Speech enhancement generative adversarial network.” *arXiv preprint arXiv:1703.09452*
21. Yi Z, Zhang H, Tan P, Gong M (2017) “Dualgan: Unsupervised dual learning for image-to-image translation.” In *Proceedings of the IEEE international conference on computer vision*, pp. 2849–2857
22. Choi Y, Choi M, Kim M, Ha J-W, Kim S, Choo J (2018) “Stargan: Unified generative adversarial networks for multi-domain image-to-image translation.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8789–8797
23. Huang X, Liu M-Y, Belongie S, Kautz J (2018) “Multimodal unsupervised image-to-image translation.” In *Proceedings of the European conference on computer vi-sion (ECCV)*, pp. 172–189

24. Zhang H, Goodfellow I, Metaxas D, Odena A (2019) Self-attention generative adversarial networks. In: International conference on machine learning, PMLR, pp 7354–7363
25. Kim J, Kim M, Kang H, Lee K (2019) U-gat-it: unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation. arXiv preprint [arXiv:1907.10830](https://arxiv.org/abs/1907.10830)
26. Marcos L, Alirezaie J, Babyn P (2021) Low dose ct image denoising using boosting attention fusion gan with perceptual loss. In: 2021 43rd annual international conference of the IEEE engineering in medicine and biology society (EMBC), IEEE, pp 3407–3410
27. Zhang L, Zhang J (2022) Ultrasound image denoising using generative adversarial networks with residual dense connectivity and weighted joint loss. PeerJ Comput Sci 8:e873
28. Porkodi SP, Sarada V, Maik V (2023) Dcgan for data augmentation in pneumonia chest x-ray image classification. In: Proceedings of international conference on recent trends in computing: ICRTC 2022, Springer, pp 129–137

Osteoporosis Detection Based on X-Ray Using Deep Convolutional Neural Network



Abulkareem Z. Mohammed and Loay E. George

Abstract In this study, we describe a computer-based technique for identifying osteoporosis by analyzing medical X-ray images utilizing deep convolutional neural networks (DCNNs). During the preprocessing phase, the suggested system prepares the original picture by acquiring the area of interest, enhancing contrast, and reducing noise. Subsequently, the smudging procedure was used to improve the system's accuracy and decrease mistake by creating a nearly identical fragile region throughout the database photos. The next step is using the suggested DCNN model to diagnose the problem. To do this, the dataset was preprocessed, smudged, and then put into the model in two parts: 75% for training and 25% for testing. With Dataset 1 and Dataset 2, the diagnoses' accuracy was 94.7% and 91.5, respectively. It is important to note that two datasets were used: Dataset 1 is the Osteoporosis Knee X-ray Dataset from Kaggle, which has two classes (osteopenia and osteoporosis), and Dataset 2 is from Mendeley, which contains three classes (osteopenia, normal, and osteoporosis).

Keywords Osteoporosis · Deep convolutional neural network · Smudging · X-ray · Computer-based method · Early detection · Diagnosis accuracy

1 Introduction

An higher risk of fractures, particularly to the spine, hip, and wrist, results from the medical disease known as osteoporosis, which causes bones to become brittle and weak [1–3]. It is caused by a gradual loss of bone mass and density that might result in weak bones that are more likely to break [4, 5]. Millions of individuals suffer from osteoporosis, a widespread illness that mostly affects postmenopausal women [6, 7].

A. Z. Mohammed (✉) · L. E. George
Informatics Institute for Postgraduate Studies, Baghdad, Iraq
e-mail: abducspone2@gmail.com

L. E. George
e-mail: phd202010550@iips.icci.edu.iq

A sedentary lifestyle, poor calcium intake, age, and family history are risk factors. Medication and lifestyle modifications may both prevent and treat the illness [8, 9].

In order to diagnose osteoporosis, a patient's medical history and imaging tests—such as DXA, CT, and MRI—are usually used in conjunction with one another. These tests measure bone mineral density and compare it to predetermined threshold values; the results of these tests determine whether a patient has osteoporosis or not. The bone cortex is also examined using X-rays. The hard outer layer of bone that envelops and protects the bony void is referred to as cortical bone. Compact bone, also known as this kind of bone, accounts for around 0.8 of skeletal mass and is vital to body structure and weight bearing due to its high resistance to bending and torsion [13, 14]. Trauma or a bone disease such as osteoporosis may destroy the cortical bone of the legs, arms, and spine. Identifying those who are at risk for fractures and starting therapy to stop future fractures are the two main objectives of osteoporosis detection [15–17].

The primary issue with using X-rays to diagnose osteoporosis is that each person's osteoporosis is different in terms of both its size and location. As a result, it is challenging to discern between the fragile picture and the usual image. The smudging technique was used to solve this issue, and it will be covered in more depth in Sect. 4.2.

The remaining sections of the document are arranged as follows: The relevant papers using the most recent approaches are presented in Sect. 2. A short explanation of the suggested system is provided in Sect. 3. Section 4 outlines the methodology, while Sect. 5 includes the analysis and discussion of the results. Lastly, Sect. 6 discusses the conclusion.

2 Related Works

In order to assess the quantitative usefulness of high-resolution MRI in femoral microstructures, Radominski et al. [18] retrieved the trabecular bone texture features from MRI images and discovered that the majority of the texture parameters were significantly varied. This is only one of several earlier studies on osteoporosis identification. DCNN was suggested by Tomita et al. [19] as a means of identifying osteoporotic VF. The algorithm was able to extract logical characteristics with results that were equivalent to those obtained by practicing radiologists using computed tomography images of the spine as input.

To identify the BMD state of postmenopausal women, Bortone et al. [20] provide a supervised method based on a research-based non-invasive study of static and dynamic bone mineral density (BMD) inspections. The work highlights the need of using machine learning methods to examine the relationship between women's BMD and both static and dynamic baropodometries, such as ANNs and SVMs. In order to identify osteoporosis in teeth, Lee et al. [21] used a convolutional neural network trained using Dental Panoramic X-rays. With no data augmentation, the DCNN's findings are quite good (92.5), surpassing those of oral and maxillofacial radiologists. Pelvic X-rays were used by Liu et al. [22] to diagnose individuals with osteoporosis.

The energy function for the proposed U-net model was produced using the SoftMax method, which utilizes X-rays to evaluate deep aspects of the medullary joint. The study's bone loss and osteoporosis groups' photos showed inadequate diagnosis. Lee et al. [23], combining feature extraction by VGGnet with random forest classification yielded an accuracy of 0.71. This was evaluated using deep learning to identify people with abnormal BMD and alert groups at high risk of osteoporosis.

Yasaka et al. [24] estimated the BMD of the lumbar vertebrae from belly CT pictures. The DXA bone mineral density and the predicted BMD using CNNs were shown to be strongly correlated. Usman et al. [25] using deep CNNs and transfer learning methods, the authors of this study looked for indicators of osteoporosis in knee X-rays. The VGG-16 model was used, both with and without fine-tuning, on X-ray images of patients' knees for accurate labeling. Accuracy, sensitivity, and specificity were used to rank the CNNs' overall effectiveness. The findings imply that the VGG-16 CNN's ability to identify osteoporosis in knee radiographs was greatly enhanced through fine-tuning. With fine-tuning, the VGG-16 was able to attain an overall accuracy of 88%, but without it, it was only able to acquire an accuracy of 80%.

3 Proposed System

Figure 1 depicts the X-ray image-based osteoporosis diagnosis method that is being proposed. In order to improve the accuracy of osteoporosis diagnosis, the process is broken down into three stages: loading images from the dataset, preprocessing to properly prepare the picture by eliminating noise, enhancing contrast, and extricating the area of interest, and smudging stage to obtain agglomerations with close values.

4 Methodology

As was previously said, the suggested system consists of a number of fundamental processes, the first of which is accessing the original picture from the dataset. The second stage entails doing some preliminary processing. Then came the smudging process. Finally, a convolutional neural network is used for categorization. In this article, we will go through each of these steps in great depth.

4.1 Preprocessing

Image preprocessing is the term used to describe the methods used to improve an image's quality, bring out important details, and facilitate processing and analysis. Preprocessing images have the following advantages:

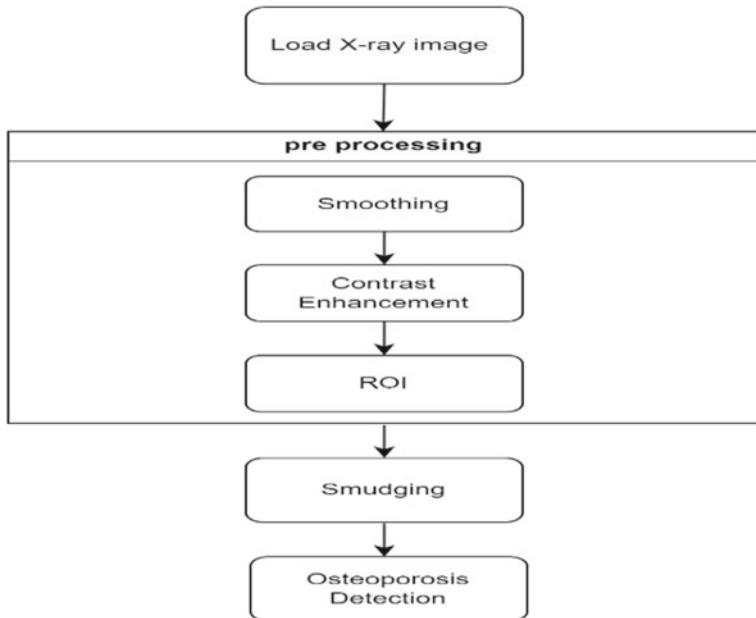


Fig. 1 Workflow of the proposed system

- Smoothing: Image preprocessing can eliminate coherent and random noise that can impact image quality and result in incorrect analysis. A mean filter was employed for noise reduction as well as enhancing features and details of an image, making it easier to extract information. Figure 2 shows the source X-ray image, and Fig. 3 shows the smooth image.
- Enhancing contrast: Contrast enhancement: To uniformly distribute the intensities in the photographs, it is important to increase the contrast in them since the

Fig. 2 Source knee X-ray picture



Fig. 3 Smooth picture**Fig. 4** Enhanced image

images in the database were shot under varying settings. For contrast enhancement, histogram equalization is a frequently used approach. It adjusts the distribution of an image's pixel intensities to improve overall contrast. Histogram equalization's main principle is to dispense pixel intensities over the image's whole dynamic range. Figure 4 displays the improved picture (Figs. 5, 6, and 7).

4.2 Smudging

Due to the diversity of individuals contributing to the dataset, there exist variations in the size and location of injuries, as well as variations in knee size due to differences in body weight (some being obese and others thin). Therefore, it is necessary to develop a method that can align the affected images, enabling easier diagnosis. To achieve this, a smudging process is employed, which is a technique utilized in remote sensing. This process aims to provide a consistent visual appearance to the images, such as gradually assigning a green color to agricultural areas and a cyan color to water bodies, even if they are somewhat dispersed. In this study, the smudging process

Fig. 5 Segmented image by threshold segmentation



Fig. 6 Region of interest coordinates

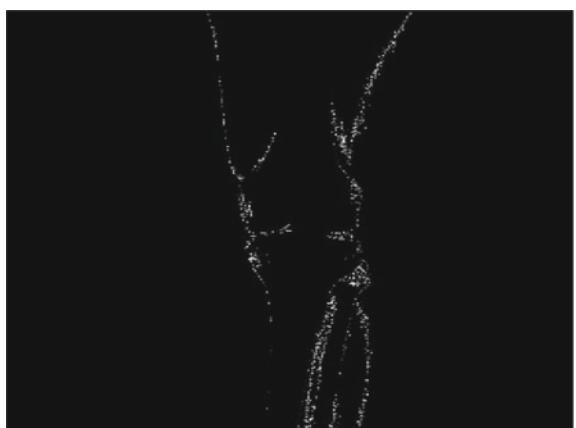


Fig. 7 Region of Interest (ROI)



Fig. 8 Smudge image

combines small, scattered, and convergent areas by employing a mean filter with a large window size. This approach facilitates image interpretation, leading to a more streamlined diagnosis process. Figure 8 illustrates an example of a smudged image.

4.3 Deep Convolutional Neural Network (DCNN)

DCNN was created especially for the analysis and identification of images. This kind of deep learning model is made up of several linked layers of artificial neurons that have been trained to identify patterns in picture data. The convolutional layer, which collects information from the input picture using a mathematical process known as convolution, is the basic component of a deep convolutional neural network (DCNN). Following that, these attributes are processed by a number of artificial neuronal layers, each of which learns ever more intricate representations of the visual input. Usually a completely linked layer, the last layer of the network generates the final classification or prediction. When it comes to picture categorization and object identification, DCNNs have shown to be far more accurate and efficient than standard computer vision algorithms. Figure 9 presents the proposed DCNNs. The recommended system incorporates 55 layers and 40.7 k learnable. All layers and their parameters are shown in Table 1.

ReLU is an abbreviation for a rectified linear unit; BN is for batch normalization; and CONV refers for convolutional layer.

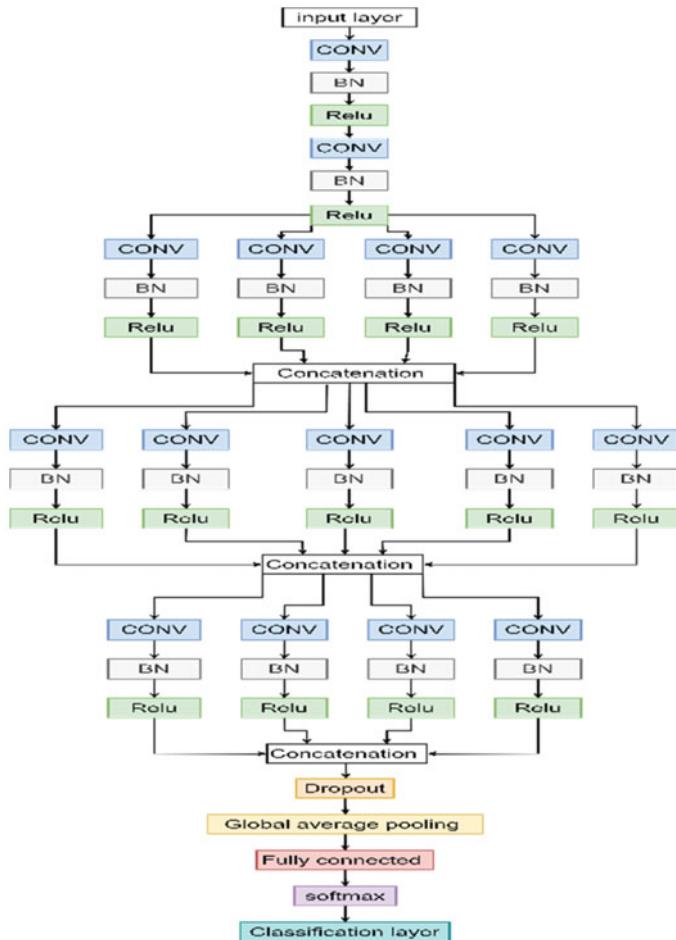


Fig. 9 Proposed DCNN model flow diagram

5 Result Analysis and Discussion

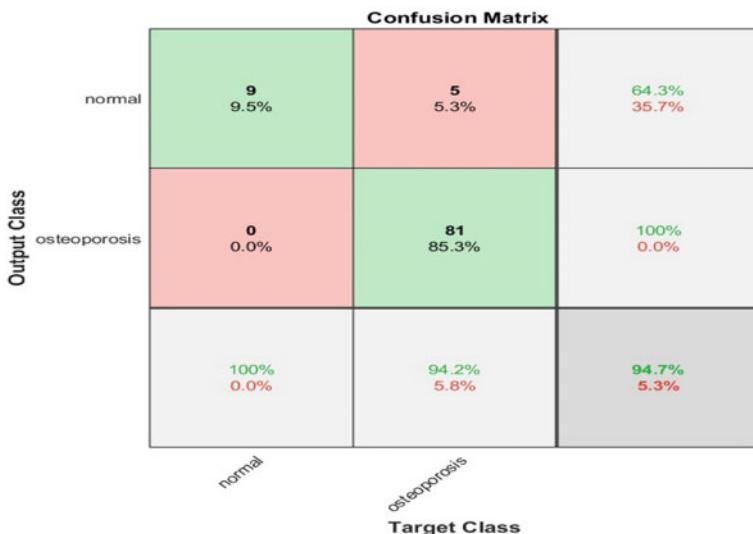
Here, the system adopted two databases: Dataset 1 is Osteoporosis Knee X-ray Dataset from Kaggle, which included two classes (normal and osteoporosis), and Dataset 2 is from Mendeley which included three classes (osteopenia, normal, and osteoporosis). The suggested system prediction using Dataset 1 is shown in Fig. 10. Figure 11 depicts the suggested system prediction based on Dataset 2; it is worth noting that the datasets were divided into 75% for training and 25% for testing.

Notably, the diagnostic accuracy was calculated according to Eq. (1).

$$AC = \frac{TN + TP}{TP + TN + FP + FN}, \quad (1)$$

Table 1 Proposed system's layers with their parameters

Layer	Parameters	Layer	Parameters
Input layer	224*224*3	BN	Default
Conv	3*3	ReLU	Name r23
BN	Default	Block 2	Branch 4
ReLU	Default	Block 2	Branch 5
Block 1	Branch 1	Block 3	Branch 1
Block 1	Branch 2	Block 3	Branch 2
Block 1	Branch 3	Block 3	Branch 3
Block 1	Branch 4	Block 3	Branch 4
Block 2	Branch 1	Block 3	Branch 5
Block 2	Branch 2	Pooling	Default
Block 2	Branch 3	Fully connected	2

**Fig. 10** Dataset 1 confusion matrix

where

An instance of False Positive (FP) occurs when the model predicts something as “Class A” when it is really “Class B.” A False Negative (FN) occurs when the model projected an instance to be “Class B” when it was really “Class A.” True Negative (TN): When an instance of “Class B” was truly “Class B,” the model properly predicted that it was “Class B.” True Positive (TP): An occurrence that was really “Class A” although the model mistakenly projected it to be “Class A.”



Fig. 11 Dataset 2 confusion matrix

The author's findings and those from earlier research on the same datasets are shown in Tables 2 and 3. The author's results were achieved using pre-trained classifiers.

Table 2 Comparing the outcomes of several models using Dataset 1, which can be found at [25]

Model	Accuracy (%)
Pre-trained ResNet50	90.2
Pre-trained VggNet-16	89.1
Pre-trained Xception	90.2
Pre-trained AlexNet	92.3
GoogLeNet [26]	90
VGG-16 [26]	87
ResNet50 [26]	83
Proposed model	94.7

Table 3 Analyzing and contrasting outcomes from various models using Dataset 2, which can be accessible at [27]

Model	Accuracy (%)
VggNet-16 with fine-tuning [28]	88
VggNet-16 without fine-tuning [28]	80
Pre-trained ResNet50	89.8
AlexNet [29]	91
VggNet-16 [29]	86.30
ResNet [29]	86.30
VggNet-19 [29]	84.20
Pre-trained AlexNet	89.8
Pre-trained VggNet-16	88.1
Pre-trained Xception	86.4
Proposed model	91.5

6 Conclusion

The research looked at the efficiency of an X-ray-based approach for identifying osteoporosis. Through a comprehensive review of various sources and results, it was found that the system performed exceptionally well in terms of preprocessing techniques to eliminate noise and identify the region of interest. The postprocessing step, particularly the smudging process, effectively highlighted the area of fragility, thereby facilitating the diagnosis process when utilizing the proposed CNN model.

It is noteworthy that the study focused on two databases of knee X-ray images. When evaluating the performance of the proposed system Dataset 1, an accuracy of 94.7% was achieved. Similarly, when employing the Dataset 2 from Mendeley, the proposed system achieved an accuracy of 91.5%.

These findings suggest that the suggested approach has a high potential for identifying osteoporosis in knee X-ray pictures. Further research and validation using larger datasets and diverse populations would be beneficial to assess the system's generalizability and potential for clinical implementation.

References

1. Derkatch S, Kirby C, Kimelman D, Jozani MJ, Davidson JM, Leslie WD (2019) Identification of vertebral fractures by convolutional neural networks to predict nonvertebral and hip fractures: a registry-based cohort study of dual X-ray absorptiometry. Radiology 293(2):405–411. <https://doi.org/10.1148/radiol.2019190201>
2. Bandirali M et al (2015) A new diagnostic score to detect osteoporosis in patients undergoing lumbar spine MRI. EurRadiol 25:2951–2959. <https://doi.org/10.1007/s00330-015-3699-y>
3. Majumdar SR et al (2008) Multifaceted intervention to improve diagnosis and treatment of osteoporosis in patients with recent wrist fracture: a randomized controlled trial. CMAJ 178(5):569–575. <https://doi.org/10.1503/cmaj.070981>
4. Becker DJ, Kilgore ML, Morrisey MA (2010) The societal burden of osteoporosis. CurrRheumatol Rep 12:186–191. <https://doi.org/10.1007/s11926-010-0097-y>

5. Clynes MA, Harvey NC, Curtis EM, Fuggle NR, Dennison EM, Cooper C (2020) The epidemiology of osteoporosis. *Br Med Bull.* <https://doi.org/10.1093/bmb/ldaa005>
6. Ji M-X, Yu Q (2015) Primary osteoporosis in postmenopausal women. *Chronic Dis Transl Med* 1(01):9–13. [10.1016%2Fj.cdtm.2015.02.006](https://doi.org/10.1016%2Fj.cdtm.2015.02.006)
7. Kanis JA, Cooper C, Rizzoli R, Reginster J-Y, and S. A. B. of the E. S. for C. and E. A. of O. (ESCEO) and the C. of S. A. and N. S. of the I. O. F. (IOF) (2019) European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporosis Int* 30:3–44. [10.1007%2Fs00198-018-4704-5](https://doi.org/10.1007%2Fs00198-018-4704-5)
8. Hans D, Baim S (2017) Quantitative ultrasound (QUS) in the management of osteoporosis and assessment of fracture risk. *J Clin Densitom* 20(3):322–333. https://doi.org/10.1007/978-3-030-91979-5_2
9. Sela EI, Widyaningrum R (2015) Osteoporosis detection using important shape-based features of the porous trabecular bone on the dental X-ray images. *Int J Adv Comput Sci Appl* 6(9):247–250. <https://doi.org/10.14569/IJACSA.2015.060933>
10. Ferizi U et al (2019) Artificial intelligence applied to osteoporosis: a performance comparison of machine learning algorithms in predicting fragility fractures from MRI data. *J Magn Reson Imaging* 49(4):1029–1038. <https://doi.org/10.1002/jmri.26280>
11. Hussain D, Naqvi RA, Loh W-K, Lee J (2021) Deep learning in DXA image segmentation. <https://doi.org/10.32604/cmc.2021.013031>
12. Dimai HP (2017) Use of dual-energy X-ray absorptiometry (DXA) for diagnosis and fracture risk assessment; WHO-criteria, T-and Z-score, and reference databases. *Bone* 104:39–43. <https://doi.org/10.1016/j.bone.2016.12.016>
13. Brett AD, Brown JK (2015) Quantitative computed tomography and opportunistic bone density screening by dual use of computed tomography scans. *J OrthopTranslat* 3(4):178–184. <https://doi.org/10.1016/j.jot.2015.08.006>
14. Fang Y et al (2021) Opportunistic osteoporosis screening in multi-detector CT images using deep convolutional neural networks. *EurRadiol* 31:1831–1842. <https://doi.org/10.1007/s00330-020-07312-8>
15. He QF et al (2018) Radiographic predictors for bone mineral loss: cortical thickness and index of the distal femur. *Bone Joint Res* 7(7):468–475. <https://doi.org/10.1302/2046-3758.77.bjr-2017-0332.r1>
16. Jiang H, Yates CJ, Gorelik A, Kale A, Song Q, Wark JD (2018) Peripheral quantitative computed tomography (pQCT) measures contribute to the understanding of bone fragility in older patients with low-trauma fracture. *J Clin Densitom* 21(1):140–147. <https://doi.org/10.1016/j.jocd.2017.02.003>
17. Shayganfar A, Khodayi M, Ebrahimian S, Tabrizi Z (2019) Quantitative diagnosis of osteoporosis using lumbar spine signal intensity in magnetic resonance imaging. *Br J Radiol* 92(1097):20180774. <https://doi.org/10.1259/bjr.20180774>
18. Radominski SC et al (2017) Brazilian guidelines for the diagnosis and treatment of post-menopausal osteoporosis. *Rev Bras Reumatol* 57:s452–s466. <https://doi.org/10.1016/j.rbre.2017.07.001>
19. Tomita N, Cheung YY, Hassanpour S (2018) Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. *Comput Biol Med* 98:8–15. <https://doi.org/10.1016/j.combiomed.2018.05.011>
20. Bortone I et al. (2018) A supervised approach to classify the status of bone mineral density in postmenopausal women through static and dynamic baropodometry. In: 2018 international joint conference on neural networks (IJCNN), IEEE, 2018, pp 1–7. <https://doi.org/10.1109/IJCNN.2018.8489205>
21. Lee J-S, Adhikari S, Liu L, Jeong H-G, Kim H, Yoon S-J (2019) Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. *Dentomaxillofacial Radiol* 48(1):20170344. <https://doi.org/10.1259/dmfr.20170344>
22. Liu J, Wang J, Ruan W, Lin C, Chen D (2020) Diagnostic and gradation model of osteoporosis based on improved deep U-Net network. *J Med Syst* 44:1–7. <https://doi.org/10.1007/s10916-019-1502-3>

23. Lee S, Choe EK, Kang HY, Yoon JW, Kim HS (2020) The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population. *Skeletal Radiol* 49:613–618. <https://doi.org/10.1007/s00256-019-03342-6>
24. Yasaka K, Akai H, Kunitatsu A, Kiryu S, Abe O (2020) Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. *EurRadiol* 30:3549–3557. <https://doi.org/10.1007/s00330-020-06677-0>
25. <https://www.kaggle.com/datasets/stevepython/osteoporosis-knee-xray-dataset?select=osteoporosis>
26. Abubakar UB, Boukar MM, Adeshina S (2022) Comparison of transfer learning model accuracy for osteoporosis classification on knee radiograph. In: 2022 2nd international conference on computing and machine intelligence (ICMI), IEEE, pp 1–5. <https://doi.org/10.1109/ICMI55296.2022.9873731>
27. <https://data.mendeley.com/datasets/fxjm8fb6mw/1>
28. Abubakar UB, Boukar MM, Adeshina S (2022) Evaluation of parameter fine-tuning with transfer learning for osteoporosis classification in knee radiograph. *Int J Adv Comput Sci Appl* 13(8)
29. Wani IM, Arora S (2022) Osteoporosis diagnosis in knee X-rays by transfer learning based on convolution neural network. *Multimed Tools Appl* 1–25. <https://doi.org/10.1007/s11042-022-13911-y>

Fault Prediction and Diagnosis of Bearing Assembly



Chirag Agarwal, Aman Agarwal, Anmol Tyagi, Dev Tyagi, Mohini Preetam Singh, and Rahul Singh

Abstract Vibration is one of the main causes responsible for the failure of any machinery, which makes it a prominent cause of failure in the detection and prediction of a machine's working baseline. Prediction and maintenance at the right time play a very important role in curbing hazardous situations. It also provides the required data to construct a roadmap to proceed in case of a fault. It also helps in selecting a favorable maintenance procedure, approximation of time required, and the price of maintenance or substitution of the required components. The focus of this paper is to investigate the vibration signals in a “bearing rotation mechanism” for hardware fault prediction and condition-based maintenance. This paper is based on the study and simulation of vibration signals and the primary data is taken from a test setup rig of a rotating bearing mechanism. The data is acquired through the combination of a vibration sensor and a NI DAQ (data acquisition) card to make a suitable dataset and also simulate that data with the help of MATLAB software. This study covers various parameters to study the behavior of the bearing such as Kurtosis, RMS factor, and Skewness. After considering the lack of existing datasets for the parameters necessary to make predictions for the bearing status, a new dataset DBCM or “Dataset for Bearing Condition Monitoring” was designed, containing data in terms of voltage within an interval of 0.04 s. The simulation results verified the readings obtained by hardware setup.

Keywords Vibration · Bearing · Kurtosis · RMS · NI DAQ card · MATLAB

C. Agarwal · A. Agarwal · A. Tyagi · D. Tyagi · M. P. Singh (✉)

Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

e-mail: mohinisingh2008@gmail.com

A. Tyagi

e-mail: anmol.tyagi.01.ask@gmail.com

R. Singh

Product Technology and Development Centre, Larsen & Toubro Ltd, Mumbai, India

1 Introduction

Because of the direct bond between motor structure and bearing assembly performance, it is tough to visualize the advancement of present-day rotating enginery without taking into account the vast implementation of bearings [1], the fault emerging in the motor also depends on bearing fault, and the main origin of bearing collapses includes pitting, spalling, electro erosion, and wear. Facet damage on the raceways or the rotating elements of bearings are the most typical cause of defects. Surface flaws can be classified as either localized faults or distributed faults [2]. Bearing failures will result in serious downtime, increased restoration costs, and even a possible drop in production. In the era of the Internet of Things and Industrial 4.0, enormous actual-time data is collected from bearing health surveillance systems [3]. In general, the literature reports on two regularly used methodologies for bearing problem diagnostics. One way is a prototype-based approach. This method creates a practical representation of the bearings and then categorizes the bearing's faults by keeping a track of the remaining signals between the modeled and collected signals [4] in order to increase plant production and decrease repair costs of these systems, as well as analyze the fitness of running equipment for that valid situation-based surveillance and detection is frequently expected [5]. In this paper, vibration-based signals are used to study the variation of vibration profile which occurs due to different physical conditions of the bearing. After doing a long literary survey on the work done so far in this field, several existing datasets were found, but none of them was useful and relevant. After considering the lack of existing datasets for the parameters necessary to make predictions for the bearing status, a new dataset was designed for this study and this dataset was inspired by a few of the datasets from older studies like "Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification" [6] and "Condition Based Maintenance Fault Database for Testing of Diagnostic and Prognostics Algorithms" [7]. To get authentic and perfect results with the newly designed dataset, a new hardware structure was constructed which is described in the II parts of the paper in the hardware designing section.

The traditional techniques for fault measuring or analyzing are done by professionals, which are well proficient in comprehending sound samples generated by equipment, but human sense-based research may fail or mislead sometimes in terms of the same behavior of the signal. The level of functional bearing can be examined in different ways by using audiovisual and tremors to sense the status of the bearing [8]. Acoustic is suitable for the condition where the immediate translation of vibration is unimpressed which means that the bearing (machine) is in a moving position or submerged in fluid. The way of vibration is suitable for a very noisy environment, where the multiplication and addition of signal vectors may affect the genuine signal. Hence, an appropriate signal-examining technique must be selected to get actual signals from noise. So, the data acquisition and analysis are influential in remote health surveillance of equipment. The method suitable to examine this signal

includes anticipation-based analysis, time-domain based analysis, and frequency-based analysis. Techniques suitable for signal processing are Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT), Wavelet Transform, Hilbert Huang Transform (HHT) [9], and Empirical Mode Decomposition (EMD). Most vibration-analyzing setup uses FFT because it is only suitable for the proposed hardware design which is stationary and also deals with the type of vibration signal acquired from the physical setup which is nexus and consists of dappled signals [10]. Datasets from the sensor are acquired with a NI DAQ card which operates a vibration sensor and also processes data automatically and provides data in terms of voltage for further comparison between different conditions of the bearing.

2 Hardware Designing

Hardware setup consists of different components, devices, and structures. The whole setup is established by joining all materials with the help of bolts and welding different components of hardware.

These components are: AC motor, three bearings, iron shaft, T-shape coupling, pulley, belt, electronic weight machine, manual load controller, belt holding structure, base, loop powered vibration sensor IRD591, sensor holding pad, stud and cable set, bottom base structure, form pad, water, NI DAQ card, connecting cable, Laptop/Desktop, two red and blue 10A wires for power supply. These all are the hardware components that are used to establish the whole setup shown in Fig. 1a, b. The motor works on AC supply which is of 220 V, 2HP, 1.5 kW, single phase, 1500RPM.

Three bearing types of different health states are used (only one bearing is used at a time). The differentiation of those bearings is done as a bearing with negligible defects or a new bearing is called a first-stage bearing a bearing with a small amount of defect, or a used bearing for some time which will also be able for further use is called as a second-stage bearing and a bearing with an extreme number of defects which was used for a very long time and would be difficult to further work with them is called a third-stage bearing. These bearings are attached to a shaft in which one end

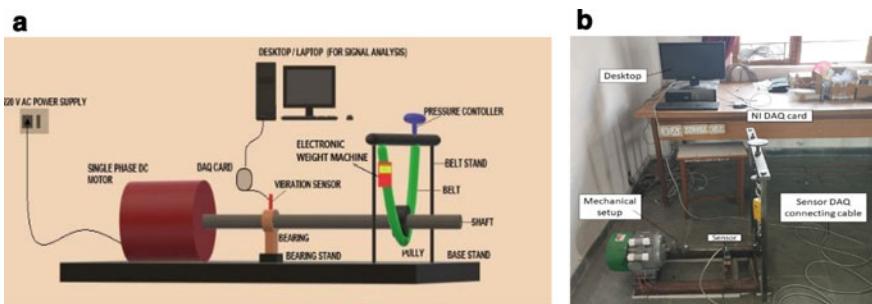


Fig. 1 **a** Mechanical setup hardware design. **b** Original proposed mechanical setup

is connected to a pulley and another end is connected to the motor with the help of a T-shape coupling that provides an interlocking mechanism between them. The whole setup is held by the base structure which is heavyweight, and of a cast iron beam, a small block of white form is also used in between a base structure for minimizing the ground vibrations. A sensor mount on a bearing is connected with a NI DAQ card with the help of a stud and cable which are available by the sensor-providing brand. NI DAQ card is used for data acquisition of the data which is provided by the vibration sensor and to process the signal by the software called NiExpress. A DAQ is connected with Laptop/Desktop by the help of a USB Cable 2.0 A/B which are used as both data transfer and power supply to DAQ card.

Details of the experimental setup and its components are given below.

2.1 AC Motor

Several studies have been undertaken that show that tierce of the population steers below 30% of their graded load [11], i.e., in application motors that are suitable for fulfilling requirements therefore, take A.C. induction motors. The motor is estimated based on the distribution of their applications and required speed, power, torque, and size; therefore, the specification of the motor used in the experimental setup operates at a 220 V voltage, 10A current, and 50 Hz frequency; they provide 1.5KW and 2H.P. power and 1440 RPM. Here, the “U” frame is preferred because it usually provides the prolonged life and the greatest overall results will give an excessive level of efficiency without interfering other functions [12]. Motor convolution padding mechanism, generally are categorized parameters according to their thermal ability. Moderate voltage systems (600 V or less) used for drive utilizations are typically Class F (155C) or Class H (180C).

The samples (motors) are subjected to a number of thermal stresses, mechanical stress, and moisture cycles throughout the qualification testing.

However, IEEE Standards limit electrical stress to 60 Hz and 600 V rms ac [13]. Usage of the full voltage starting method, also known as across-the-line beginning, to start motors since it is the simplest, has the economical equipment costs, and is the highly dependable. A control is used in this manner to seal a contactor and put under full line voltage to the motor input and output terminals. This approach allows the engine to produce the most beginning torque while also providing the quickest acceleration times [14].

Motor losses are aggregated into the following categories [15]:

- Iron core losses—magnetic losses in laminations, inductance, and eddy-current losses.
- Stator resistance—current losses in the windings.
- Rotor resistance—current losses in the rotor bars and end ring.
- Windage and friction—mechanical drag in bearings and cooling fans.

- Stray load losses—magnetic transfer loss in the air gap between the stator and rotor.

2.2 T-Coupling

The primary factors for machine vibrations are imbalance and non-alignment. So, first attention must be given to a rotating shaft for reducing the vibration effect for increased setup lifetime and maintenance cost where the bearing setup is placed. Unlike other works related to the subject, it consists of the coupling parameters as variables [16] (Fig. 2a).

Both couplings are connected by an interlocking between them and a star-like hardcore rubber section is placed as shown in Fig. 2b and c because it helps in perfect interlock between both the shafts and also reduces the vibration between couplings which can occur due to collision in a jaw of coupling due to a small gap between them. A hardcore rubber section also helps in reducing this gap.

The material used in manufacturing is cast iron. It is selected for each of the flanges and the natural rubber is used in the bush. No nut and bolt are used to clamp both the flanges [17].

In its basic form, the SPIDER (JAW) coupling in Fig. 2c comprises a single-piece crucifix-shaped encapsulant element that is hold on to the “jaws” of two mating hubs. Big couplings usually utilize different elements. An upper hand of this structure is the capacity to alter the constitution and hardness of the elastomer used and hence alter the torque potential and torsional rigidity. However, this structure can indulge comparatively less amounts of misalignment [18].

The spider is used between flingers of coupling because jaw coupling insertions should never let the jaws of one hub to touch the opposing hub, which would result in a noisy, “grinding” reaction. That is why, spiders and load paddings are generally equipped with special dots made just to protect the required gap between the metal components [19].

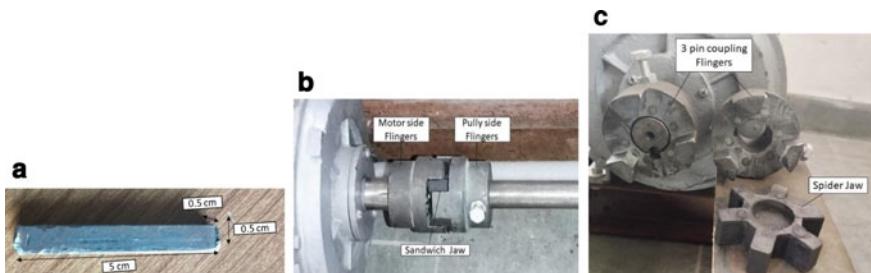


Fig. 2 a Metal key. b Interlocked coupling. c Three-pin coupling flinger and sandwich jaw

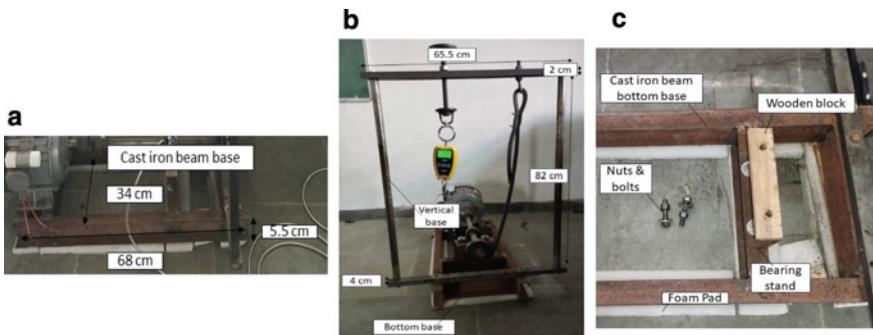


Fig. 3 **a** First base structure. **b** Second base structure. **c** Third base structure

2.3 Setup Holding Base

Base designs always keep the intent of the amount of load mounted on them and condition whether it operates, and material is chosen accordingly. In the proposed experimental setup, the base is formed with the gray type of cast iron and its joints are attached through the help of the electric welding. A base is formed in three structures which are joined with each other with the help of the nut and bolts. The base sections are as:

- The first structure is the main base that is directly in contact with the ground and also holds all weight and the components of the experimental setup. It is used to hold the motor on one end and second structure at opposite end, third structure at between motor and second structure, shaft, jaw coupling, pulley, bearing, sensor shown in Fig. 3a.
- A second structure which is placed at the one end of the first structure through nuts and bolts is used to hold the load control mechanism, belt, electronic weight machine, and hooks shown in Fig. 3b.
- A third structure is placed above the first structure in between a motor and second structure which was used to set up a vibration sensor through two nuts and bolts and placed a small size wooden block to minimize the internal vibration shown in Fig. 3c.

The small size of white synthetic form blocks is placed between the first structure which is near the ground to minimize the ground and frictional vibration which occurs during the run time of the experimental setup.

2.4 Ball Bearing

Bearing acts as a part of the mechanical structure or any structure where any type of rotation is performed because the bearing assists object rotation. Different causes

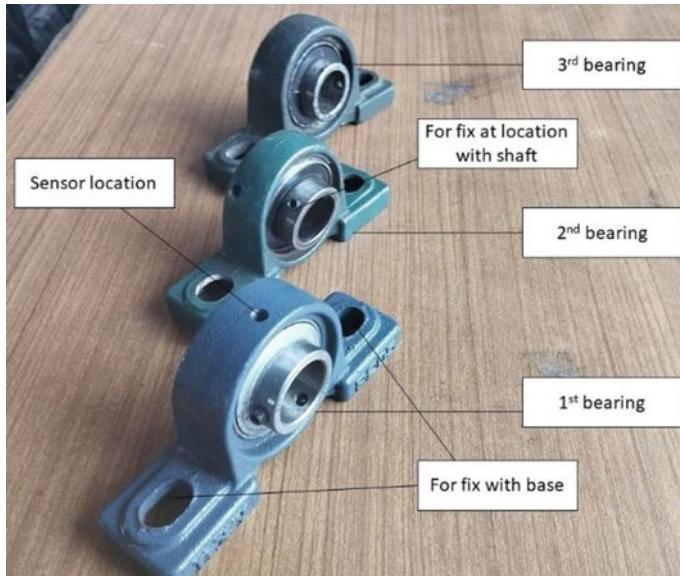


Fig. 4 Ball bearings used in experimental setup

result in bearing collapses which includes excessive-heating, overaxial and radial loads, and electrical stress such as the existence of bearing currents [21]. In the proposed experimental setup, the 2 mm mount bearing is used. A mount bearing is preferred because it is easy to install or mount only with the help of nuts and bolts, and they also have separate points to lubricate them and tiny size screws that are used to hold the bearing in position in the shaft. In the whole experimental duration, a total of three bearings are used in the same type but with different conditions. Different bearing conditions are shown in Fig. 4

- The first bearing conditions are considered as a starting stage. In this stage, the bearing is in perfect working and physical condition, producing very less frictional sound and vibration in running condition; they are almost new bearing which is used directly from the hardware shop.
- The second bearing condition is considered as an intermediate stage. In this stage, the bearings have a few deformities on their surface and also rust is present, producing frictional sound and vibration in running conditions. This bearing is from the junkyard that was used for 3 years in automotive applications.
- The third bearing condition is considered the last stage. In this stage, the bearings have serious deformities on their surface and also a lot of rust is present, producing a loud frictional sound and vibration in running conditions. This bearing is also from the junkyard that is used for 7 years in automotive applications.

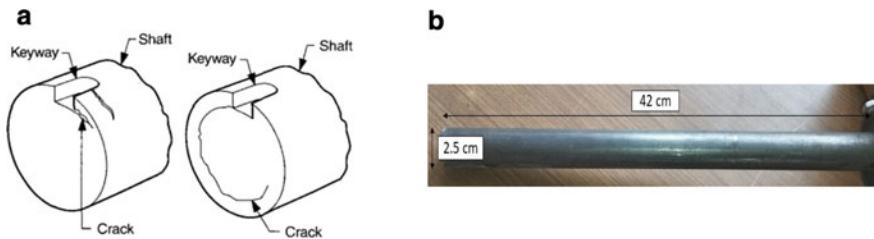


Fig. 5 **a** Peeling-type cracks in shafts usually originate at the keyway [25]. **b** Iron cylindrical shaft

Vibration accounts for the largest portion of bearing noise. The noise in the bearing is dependent on two prime parameters: the origin and traversal of vibration and succeeding radiation [22].

2.5 Shaft

Shaft is a long, cylindrical shaped, whole mass-body object used to transfer a rotary motion to the entire mechanism. The chosen shaft material is steel because it is the most generally used material in uses like construction, petrochemical, and manufacturing industry due to its practical properties. Depending on its features such as strength, ductility, hardness, and cost, there are different classifications of steel types [23] mixed with stainless steel which is generally meant to be used in a rusty and corrosive environment due to its anti-corrosive properties [24]. In the proposed structure used shaft shown in Fig. 5b is of a meter length 42 cm, diameter 2.5 cm, the weight of 5 kg, and also have flinger on both ends to hold pulley and flexible jaw coupling. The area where the fingers are formed which has the highest level of failure due to stress and crack is possible on the shaft surface shown in Fig. 5a.

2.6 Pulley and Belt

A pulley is a type of wheel that is used to drive a belt or transfer a load. Pulleys are selected according to the required application because pulleys are made up of different materials, sizes, efficiency, temperature operation, load handle, and speed. In the proposed structure, the pulley used is made up of cast iron and the diameter is 9.5 cm.

The belt used is a V-like structure in its front surface that is easily fixed in the pulley curved surface which is used to hold the belt. The belt is made up of leather and hard rubber.

2.7 Load Controller

A load controller is used to control the load on the motor-bearing pulley structure through the given load to the pulley with the help of a belt. Loads are in terms of weight, a screw like mechanism is in the upper right corner of the base on which a belt is attached with the help of the hook (freely move) present in the screw. When the screw is rotated in a clockwise direction, then the belt becomes loose which is used for decreasing the load, and when the screw is rotated in an anti-clockwise direction, then the belt becomes tight which is used for increasing the load.

2.8 Electronic Weight Machine

An electronic weight machine is present on the upper left corner of the base with the help of the wire they are winding with the base and a belt is connected with them through the help of the hook. The machine used in the proposed setup is capable of measuring a maximum load of 200 kg. It gives the measured weight in terms of a kilogram, which is a battery-driven device that accepts two AAA batteries that are connected in series.

2.9 NI DAQ Card

National Instruments Data Acquisition card is a card used to convert data from analog to digital and vice versa and also helps to process mechanical or electrical components like voltage, current, pressure, or sound. It consists of sensors, processing and measuring hardware and computers with programming software. LABVIEW has data acquisition, analysis, and presentation tools in a single program for getting data or using a DAQ card [26]. In the proposed setup, NI DAQ card and NIExpress programming software are used. They are connected with the help of a USB 2.0 cable to the computer and get power also from it, and they operate at the 5 V voltage. Here use an analog side to get data from vibration sensors. The sensor has three wires where the green wire connects through the ground of DAQ, red wire transmits data to the in port of DAQ, and a black wire connected to DAQ for powering up of the sensor. (Fig. 6a).

2.10 Vibration Sensor

A vibration sensor is a sensor used to measure the vibration occurring near them. The sensor is installed based on the near-approach condition, and hence, the sensor

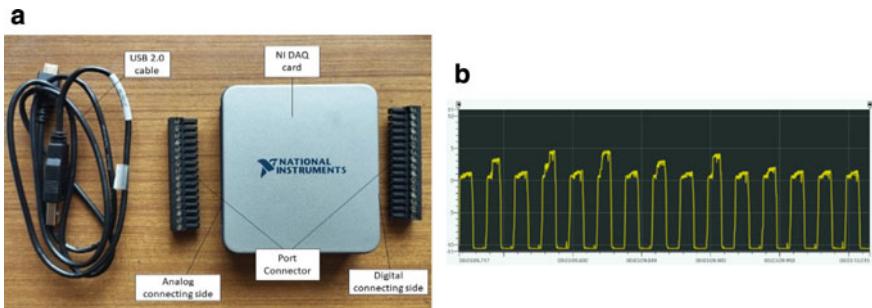


Fig. 6 **a** NI DAQ card. **b** NI DAQ card provided data in analog form

is directly installed on the bearings through a small double-ring screw. The sensor used to analyze vibration is an IRD591 Accelerometer 4–20 mA velocity output via 2 Pin MS Connector. Sensor operates in a 15–20 V DC power supply, settling time is of 2 s, and output impedance is of $600\text{-}\Omega$ max. at 24 V. The case material used to manufacture a sensor is stainless steel, the sensing element is PZT, the weight is 150 gm, the temperature operating range is $-25\text{ }^{\circ}\text{C}$ – $90\text{ }^{\circ}\text{C}$, and the sealing used is IP68 (Fig. 7a–b).

Connector used to connect the sensor with the DAQ card is HS-AA004—non-booted, HS-AA053 or HS-0054—booted. The Vibration sensor should be lightly attached to a flat surface.

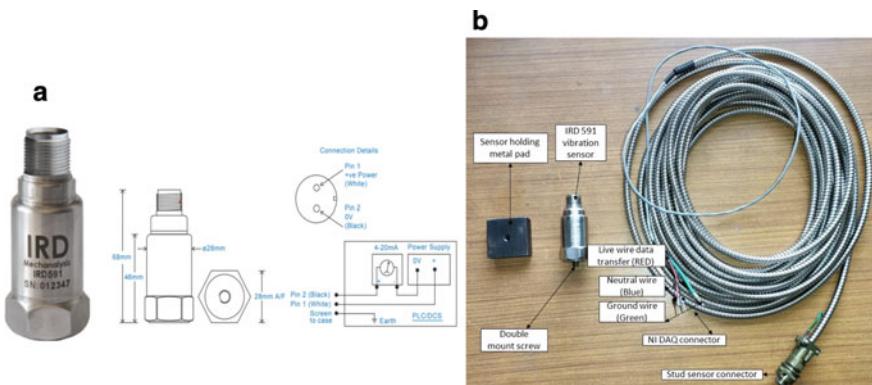


Fig. 7 **a** IRD591 vibration sensor [28]. **b** Sensor connecting equipment

S.No	UNIT	BEARING DATA															
		1st Stage (new bearing)					2nd Stage (some amount of damaged bearing)				3rd stage (extreme amount of damaged bearing)						
		0k	1k	2k	3k	4k	0k	1k	2k	3k	4k	0k	1k	2k	3k	4k	
1	0	-0.0003865	0.000312	8.26E-06	0.001426	0.00091	-1.41722	-0.98888	-10.5582	-6.79987	-2.62613	10.5582	-0.92308	-1.12564	-0.25735	-0.55151	
2	0.024	-0.0003865	0.000312	-0.00051	0.001426	0.00091	-0.63279	-1.88169	0.369681	-7.26046	-1.48947	0.750285	-0.82245	-0.69085	-0.4612	-0.45862	
3	0.048	-0.0003865	0.000312	8.26E-06	0.001426	0.00091	-0.001426	-0.21993	-1.36304	1.197979	-5.62967	-1.25466	0.979607	-0.25864	-0.75663	-0.79406	-0.8335
4	0.072	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.68827	-1.16693	-1.06887	-3.12543	1.85976	-10.0873	-1.19273	-0.47281	-0.9992	-0.85341	
5	0.096	-0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-1.24692	-1.10758	-10.5582	-2.04038	-2.20682	-10.5582	-1.62494	-0.39669	-0.40443	0.127127	
6	0.12	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-1.45335	-0.97985	-10.5582	-6.58957	-2.63387	-10.5582	-1.14886	-1.03404	-0.18768	0.103909	
7	0.144	-0.0003865	0.000312	-0.00051	0.001426	0.00091	-1.00953	-1.75138	-0.2226	-7.32497	-1.33465	0.536115	-0.67924	-0.89212	-0.51668	-0.73472	
8	0.168	-0.0003865	-0.0002	-0.00051	0.001426	0.00091	-0.71665	-1.45722	1.043157	-5.6	-1.20369	1.154113	-0.80052	-0.65085	-0.85083	-0.62247	
9	0.192	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.85857	-1.25853	0.004559	-3.1422	-1.50238	-10.5582	-0.64569	-0.51409	-0.68311	-0.38766	
10	0.216	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-1.20563	-0.51668	-10.5582	-2.89577	-2.32422	-10.5582	-0.83793	-0.64182	-0.31541	-0.19155	
11	0.24	0.0003865	-0.0002	8.26E-06	0.00091	0.001426	-1.13467	-1.0779	-10.5582	-5.81933	-2.53452	-10.5582	-0.91405	-1.02884	0.087133	0.058747	
12	0.264	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.76568	-1.07145	0.199377	-7.29917	-1.38239	0.577401	-0.61215	-1.16693	-0.37213	-0.6973	
13	0.288	0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.65601	-1.47012	1.039286	-6.41926	0.30639	1.240555	-1.06758	-0.5399	-0.8947	-0.42278	
14	0.312	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.76955	-0.59538	-0.69501	-3.21187	-1.71655	-10.5582	-1.23144	0.169703	-0.50248	-0.44055	
15	0.336	0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.87277	-0.9605	-10.5582	-2.17456	-2.70096	-10.5582	-1.68042	-0.80568	-0.23424	-0.01737	
16	0.36	-0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-0.75149	-1.29853	-10.5582	-6.39217	-2.83901	-9.70149	-0.77987	-1.1308	-0.06382	-0.42636	
17	0.384	-0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-0.55667	-1.06371	0.520633	-8.34938	-2.00039	0.248404	-0.93598	-0.87793	-0.05608	-0.59022	
18	0.408	-0.0003865	-0.0002	0.000524	0.001426	0.00091	-0.80181	-1.49463	1.465047	-5.62451	-0.3954	0.99413	-0.92953	-0.36056	-0.80181	-0.57344	

Fig. 8 Overall dataset of randomized data used [27]

3 Experimental Procedure

The working procedure is the same for every bearing: All the bearings are driven for a total of five types of load conditions that load conditions are 0 k (no load), 1 k (1 kg load), 2 k (2 kg load), 3 k (3 kg load), 4 k (4 kg load).

All these load conditions are performed for all three types (first bearing condition, second bearing condition, and third bearing condition) shown in Fig. 4 of bearings, each of them is driving for approx. 5 min in each bearing at least for each load condition, and therefore, each bearing is run for 25 min.

The sensor records data for 5 min and provides data which is of very precise manner because it gives data in the interval of 0.0001 s. The overall dataset used is shown in Fig. 8 and in a procedure called a dataset with the name of BCMD [27].

The sensor's data is accessed by the operating system with the help of an NI DAQ card which converts the sensor's analog generated into digital data which is processed by NI DAQ Express software from the DAQ card which shows data from the sensor, in both digital and graphical form and also calculates various parameters accordingly, like FFT graph and curve fitting (Figs. 9 and 10).

It also provides a recording facility to analyze data and generate CSV files of the sensor data to analyze the complete information or parameters taken from the NI DAQ Express software shown (Fig. 11).

4 Simulation for Fault Diagnosis

To simulate the dataset for fault diagnosis, it was necessary to generate data that mimics the behavior of the real bearing under various fault conditions. The outline of the process involved is as follows:

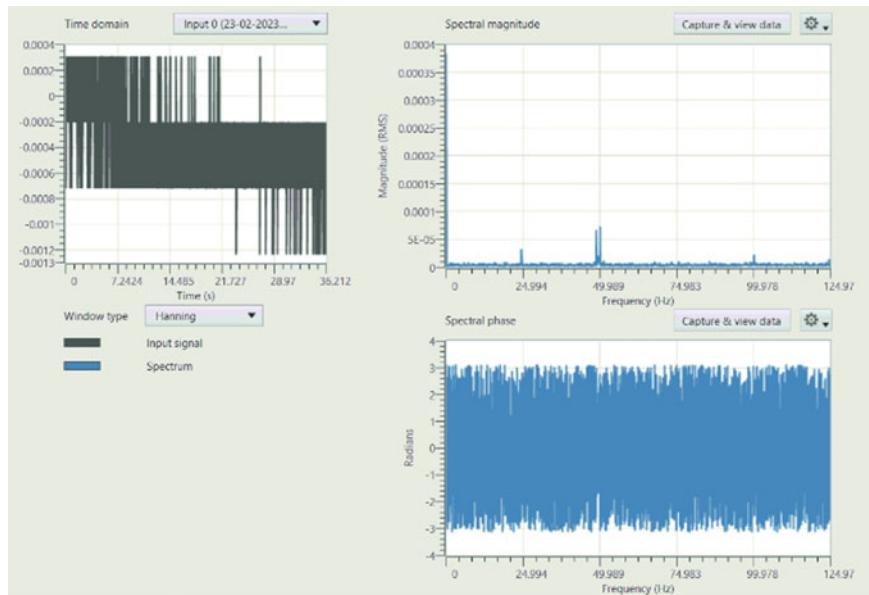


Fig. 9 Above data FFT

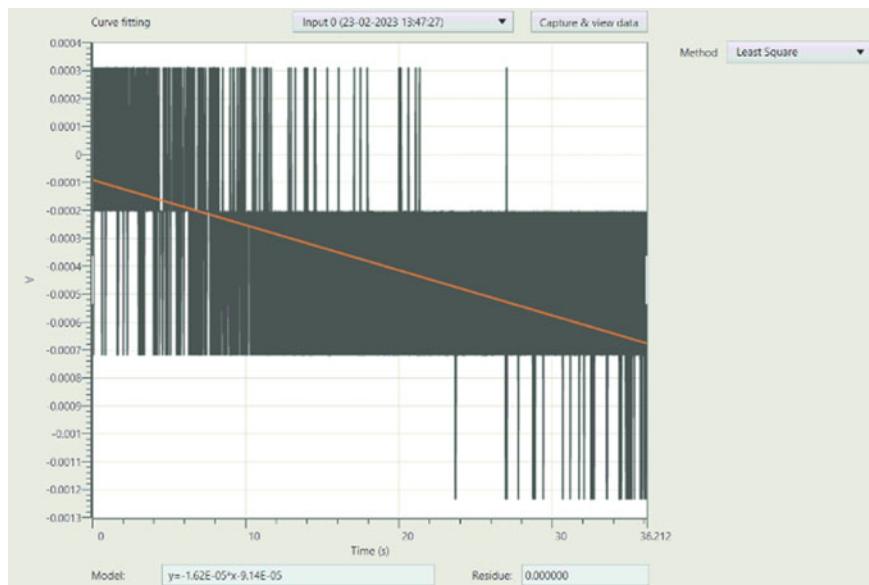


Fig. 10 Curve fitting of data in Fig. 9

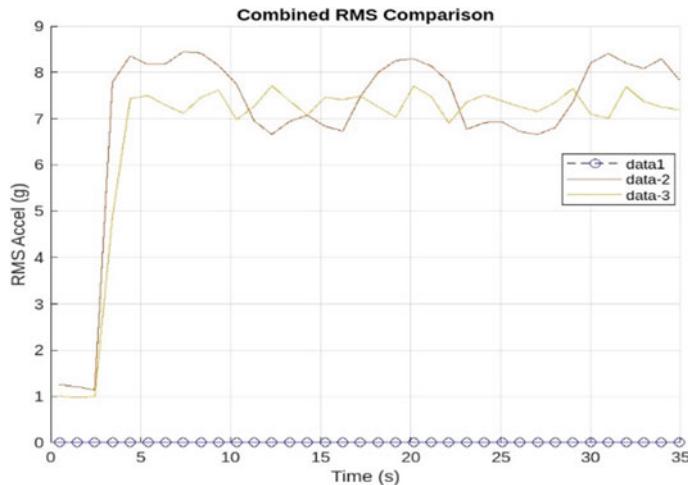


Fig. 11 RMS plot for 0 K load with all three stages

4.1 Parameter and Operating Conditions

The simulation parameters were carefully selected after months of literature survey from older research work in different fields related to the same cause and were considered for producing simulation models that were based on vibration signal analysis. This was further classified into different stages based on the amount of fault that was induced in the bearings. The main parameters which were simulated using MATLAB were as follows:

- *RMS (root mean square)* value is considered as the work done part of the vibration profile signature and thus shows the devastating abilities of the vibration signals.
- *Skewness* shows whether the signal is positively or negatively skewed. A vibration signal with a general distribution, i.e., a general bearing signal will have a skewness value of zero.
- *Kurtosis* is a parameter that is used to classify a signal. It also shows the estimate of the “peaked Ness” of a vibration signal. Signals that have more kurtosis value have larger peaks. Peaks that are more than three times the RMS value of the signal.

The faults in the ball bearing are indicated using different stages which represent a different level of faults from no fault to a moderate fault to a fully faulty bearing. The above-mentioned stages are as follows:

- *Stage One*—no faults, perfect normal bearing as mentioned above.
- *Stage Two*—some faults, moderately used bearing as mentioned above.
- *Stage Three*—fully faulty, worn-out bearing as mentioned above.

Loads were also induced at the same time with the above parameters/operating conditions and stages. The test was conducted for different loads at the above-mentioned fault stages, to obtain the above-mentioned parameters. The loads were 0 K, 1 K, 2 K, 3 K, and 4 K, respectively. For all three stages and five load conditions, a total of 15 datasets were obtained from the physical test rig. The data was refined and tuned as per requirement before simulating it in MATLAB.

5 Result

Below is the first statistical table for the RMS parameter for Stage 1, Stage 2, and Stage 3 with loads 0 K, 1 K, 2 K, 3 K, and 4 K, respectively, followed by RMS plots for all five loads with all three stages in a comparative manner.

From the performed simulation tests and the recorded data, it is visible that the RMS values increased significantly in stages where the bearing is faulty which means that in all these cases, the vibration signals were very high which can lead to wear and tear of the shaft and other components. It also shows that the load applied on the bearings somewhat balanced those high vibration spikes and the difference between the two faulty stages gets reversed which proves that in all cases the vibration signature of moderately faulty or faulty bearing is much higher than the normal or good bearing which proves that the simulation model was able to identify the good bearing and the faulty bearing based on its vibration profile from its RMS values (Table 1).

5.1 Plots for Different Loads

The given curves are plotted between RMS acceleration and time for different loads for all three stages. Here,

- *Data 1*—good bearing or Stage 1.
- *Data 2*—less faulty bearing or Stage 2.
- *Data 3*—total faulty bearing or Stage 3.

Table 1 RMS of all stages at different loads

Load (K)	Stage 1	Stage 2	Stage 3
0 K	0.0014	7.1906	7.3555
1 K	0.00048	7.4156	7.2287
2 K	0.000329	5.7490	5.0194
3 K	0.0013	4.4023	1.9458
4 K	0.000867	1.8008	0.9282

- **For 0K Load**

It can be observed that the data for Stage 1 is stabilized at zero which means there are very low vibrations in the case of a good bearing or Stage 1, whereas, in the case of faulty bearings, the values are much higher, showing signs of destructive vibration signals. In the case of moderate faulty bearing or Stage 2 the values are less distorted which means that the vibrations were high in some regions only where the bearing for pitted or rusted whereas, in the case of total faulty bearing or Stage 3, the values are more distorted with continuous peaks which means that the vibrations were high in all regions because the whole bearing was pitted and rusted. The results suggest that the vibrations are high and consistent in the case of a faulty bearing or Stage 3 which the simulation model easily identifies (Fig. 11).

- **For 1 K Load**

The values are still showing signs of destructive vibration signals, but the values are less spiked now than earlier. In the case of moderately faulty bearing or Stage 2, the values are less peaked and distorted which means that the vibrations were high in some regions only where the bearing for pitted or rusted but the load somewhat countered it whereas, in the case of total faulty bearing or Stage 3, the values are more distorted with continuous peaks which means that the vibrations were high in all regions because the whole bearing was pitted and rusted and the load has somewhat countered it in this case too (Fig. 12).

- **For 2K Load**

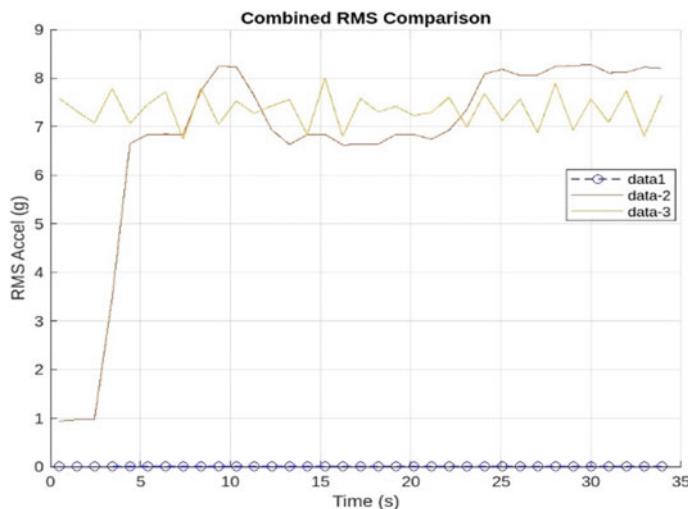


Fig. 12 RMS plot for 1 K load with all three stages

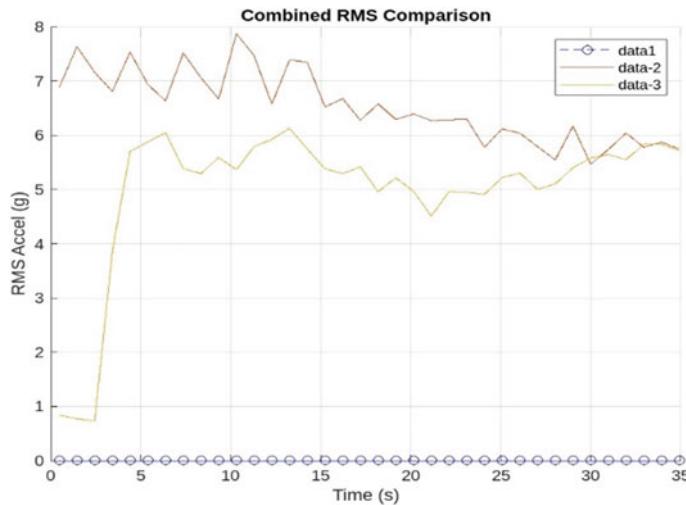


Fig. 13 RMS plot for 2 K load with all three stages

The vibration signatures of Stage 2 and Stage 3 have stopped overlapping due to the increase in load and their places have started getting reversed. The values are still showing signs of destructive vibration signals, but the values are less spiked now than earlier. In the case of moderately faulty bearing or Stage 2, the values are less peaked and distorted which means that the vibrations were high in some regions only where the bearing was pitted or rusted but the load somewhat countered it whereas, in the case of total faulty bearing or Stage 3, the values are more distorted with continuous peaks which means that the vibrations were high in all regions because the whole bearing was pitted and rusted and the load has countered it (Fig. 13).

• For 3K Load

It can be observed that the data for Stage 1 is stabilized at zero in the case of a 3 K load which means that there are very low vibrations in the case of a good bearing or Stage 1, whereas, in the case of faulty bearings, the effect of the load has started to change the peak values and the vibration signature of the bearing. The gap between the vibration signatures of Stage 2 and Stage 3 has increased, representing a much bigger difference between the values due to the increase in the load. The values are still showing signs of destructive vibration signals, but the values are less spiked now than earlier. In the case of moderately faulty bearing or Stage 2, the values are less peaked and distorted which means that the vibrations were high in some regions only where the bearing was pitted or rusted but the load somewhat countered it whereas, in the case of total faulty bearing or Stage 3, the values are more distorted with continuous peaks which means that the vibrations were high in all regions because the whole bearing was pitted and rusted and the load has somewhat countered it in

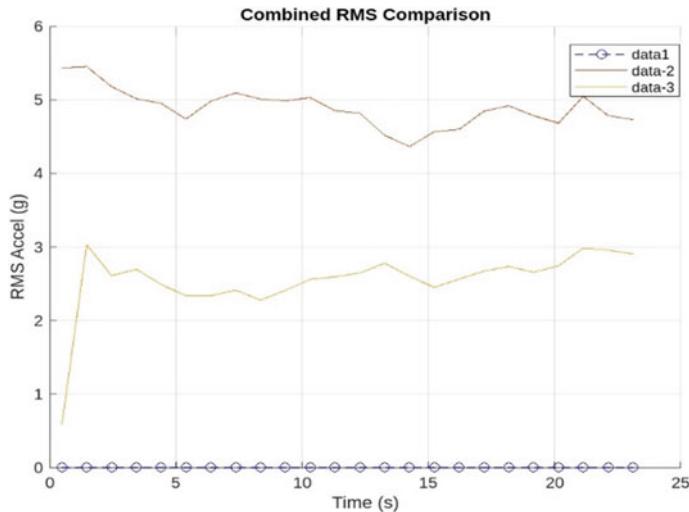


Fig. 14 RMS plot for 3 K load with all three stages

this case too but the results still suggest that vibrations are high and consistent in the case of totally faulty bearing which the simulation model easily identifies (Fig. 14).

• For 4K Load

It can be observed that the data for Stage 1 is stabilized at zero in the case of 4 K load which means that there are very low vibrations in the case of a good bearing or Stage 1, whereas, in the case of faulty bearings, the effect of the load has started to change the peak values and the vibration signature of the bearing. The gap between the vibration signatures of Stage 2 and Stage 3 has increased way too much, representing the biggest difference between the values due to the increase in the load (Fig. 15).

The values are still showing signs of destructive vibration signals, but the values are less spiked now than earlier. In the case of moderately faulty bearing or Stage 2, the values are less peaked and distorted which means that the vibrations were high in some regions only where the bearing was pitted or rusted but the load somewhat countered it whereas, in the case of total faulty bearing or Stage 3, the values are more distorted with continuous peaks which means that the vibrations were high in all regions because the whole bearing was pitted and rusted and the load has somewhat countered it in this case too but the results still suggest that vibrations are high and consistent in the case of totally faulty bearing which the simulation model easily identifies. All these plots are used to compare all the stages for different loads for our first parameter, i.e., RMS values. Now, moving on to the other parameters which are Skewness & Kurtosis in the form of a statistical table.

From the performed simulation tests and the recorded data, it is visible that the Skewness values which are almost stable around zero significantly in stages where the bearing is normal or in good condition without any load. As the load increases,

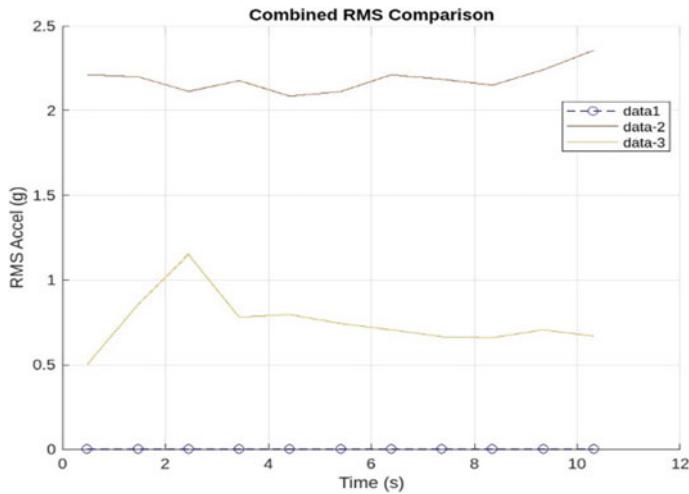


Fig. 15 RMS plot for 4 K load with all three stages

the values get positively and negatively skewed. It also shows that in the case of the two faulty stages, the skewness values are less stable to unstable around zero. After inducing loads in case of faulty bearings, the values are still unstable and are negatively and positively skewed. In some cases, the values were far away from the benchmark which is zero. This proves that the simulation model was able to identify the good bearing and the faulty bearing based on its vibration profile from its Skewness values (Table 2).

From the performed simulation tests and the recorded data, it is visible that the Kurtosis values of Stage 1 have peaked to almost twice the original values as the load increases after taking a dip at around 2 k load. It also shows that in the case of the two faulty stages, the kurtosis peaked at total faulty stages when the applied load was more than 2 K. This shows that after a certain amount which in this case is 2 K, the kurtosis values start changing significantly. Also, in the case of normal or good bearing, the kurtosis values act differently than that of faulty bearing kurtosis values (Table 3).

Table 2 Skewness of all stages at different loads

Load (K)	Stage 1	Stage 2	Stage 3
0 K	0.051	-0.0242	0.0228
1 K	0.1167	0.0738	-0.0334
2 K	-0.2729	-0.0592	0.0862
3 K	-0.2991	-0.0975	-0.7860
4 K	-0.6284	-0.4129	-6.4708

Table 3 Kurtosis of all stages at different loads

Load (K)	Stage 1	Stage 2	Stage 3
0 K	2.1378	1.1257	1.0700
1 K	2.3660	1.2151	1.0852
2 K	1.3971	2.3431	2.4529
3 K	4.4218	2.0184	3.7399
4 K	4.3355	2.6684	57.9135

6 Conclusion

The approach is used in detecting bearing faults with the help of MATLAB for simulating a virtual model based on the required parameters and operating conditions which are reciprocated from a physical hardware test rig's dataset. The procedure used incorporates the most appropriate feature selection according to the required conditions, parameters, and faults which results in a customized yet basic simulation model. The three stages are shown simultaneously for a comparative approach which makes it easier to study in both a graphical and statistical manner. Kurtosis, Skewness, and RMS were the parameters that were taken into consideration for this study and in all the physical and simulation tests.

7 Summary

This study presents an approach toward developing a mechanical structure for finding deformities and the nature of physical faults which occur in a ball bearing when it runs in a continuous rotation manner. This is done by analyzing the voltage signals which were generated from the vibration sensor in running conditions which can help in predicting and developing a condition-based maintenance system for the bearing by analyzing different parameters through sensor-generated graphs and processed datasets. The data recorded for each stage and condition of the bearing indicates that heavy vibrations occur in the case of bearings with irregular inner race surfaces and balls with rust which could lead to destructive wear and tear of the hardware and can have serious consequences. The effect of combined defect affects the vibration of the bearing mechanism too. It is also seen that the categorized accuracy of this MATLAB-based simulation model is as good as any other traditional machine learning or mathematical modeling-based techniques. The results obtained can vary with a bigger dataset. A small training dataset is taken for this research as in physical situations less researched data is present. The outcomes indicate the probable implementation of simulation for creating a virtual-based model which can become beneficial in the detection of flaws for suing in situation-based repairs to prevent any disastrous negligence and to decrease any usage cost.

8 Future Scope

The concerned project's work can be extended further in various aspects like oil and lubrication analysis of the Rolling Bearings. Bearing is a mechanical structure, and like any other moving or rotating structure, it also requires proper lubrication in order to minimize friction and increase the overall efficiency. Hence, it is necessary to analyze the oil or lubricant and to locate as well as predict the possible faults that might occur because of impurities present in the lubricant. The work with revolutions per minute (RPM) and its variations with load can be done.

Acknowledgements This work was supported by TIH-IoT CHANAKYA UG/PG Internship/Fellowship Program 2022 with Larsen & Toubro Ltd as Industrial mentors. The Authors would like to acknowledge Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India for giving the required resources and infrastructure for this project.

References

- Li B, Chow M-Y, Tipsuwan Y, Hung JC (2000) Neural-network-based motor rolling bearing fault diagnosis. *IEEE Trans Indus Electron* 7(5)
- Dolenc B, Boškoski P, Juričić Đ (2015) Distributed bearing fault diagnosis based on vibration analysis. Elsevier, *Mechanical Systems and Signal Processing*, 11 June 2015
- He M, He D (2017) A deep learning based approach for bearing fault diagnosis. *IEEE Trans Indus Appl*. <https://doi.org/10.1109/TIA.2017.2661250>
- He D, Li R, Zhu J (2013) Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Trans Indus Electron* 60(8)
- Esfahani ET, Wang S, Sundararajan V (2013) Multisensor wireless system for eccentricity and bearing fault detection in induction motors. *IEEE/ASME Trans on Mechatron* 1083–4435
- Mohanty S, Gupta KK, Raju KS, Singh A, Snigdha S (2013) Vibro acoustic signal analysis in fault finding of bearing using empirical mode decomposition. In: IEEE international conference on advance electronic systems, 21–23 Sept. 2013, pp 29–33
- Lessmeier C, Kimotho JK, Zimmer D, Sextro W (2016) Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification. In: European conference of the prognostics and health management society 2016
- Bechhoefer E (2013) Condition based maintenance fault database for testing of diagnostic and prognostics algorithms. Data Assembled and Prepared on behalf of MFPT
- Rai VK, Mohanty AR (2006) Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert Huang transform. 21:2607–2615
- Mohanty S, Gupta KK, Raju KS (2014) Comparative study between VMD and EMD in bearing fault diagnosis. Central Electronics Engineering Research Institute (CEERI)/Council of Scientific & Industrial Research (CSIR)
- Bonnett AH (1992) Understanding the changing requirements and opportunities for improvement of operating efficiency of A.C. motors. *IEEE PCIC-CH3142—July 92*
- Bonnett AH (1994) An update on AC induction motor efficiency. *IEEE Trans Indus Appl* 30(5)
- IEEE Standard Test Procedure for Evaluation of Systems of Insulating Materials for Random-Wound AC Electric Machinery, IEEE Standard 117, 1974/ANSI Standard C50.32 (1976)
- Larrabee J, Pellegrino B, Flick B (2005) Induction motor starting methods and issues. *IEEE PCIC-2005-24*

15. Malinowski J, McCormick J, Dunn K (2004) Advances in construction techniques of AC induction motors: preparation for super-premium efficiency levels. *IEEE Trans Indus Appl* 40(6)
16. Verucchi C, Bossio J, Bossio G, Acosta G (2016) Misalignment detection in induction motors with flexible coupling by means of estimated torque analysis and MCSA, 0888–3270/& 2016 Elsevier Ltd
17. Kumar S, Sehgal R, Kumar R, Bhandari S (2012) Vibrations analysis of 4 jaw flexible coupling considering unbalancing in two planes. *Int J Sci Technol* 1(11)
18. Johnson CM (1996) An introduction to flexible couplings. World Pumps
19. McCullough (2000) Understanding spiders and their role in jaw couplings. World Pumps
20. de Campos MF, Rolim Lopes LC, Magina P, Lee Tavares FC, Kuniyoshi CT, Goldenstein H (2005) Texture and microtexture studies in different types of cast irons. *Mater Sci Eng A* 398:164–170
21. Muetze A, Strangas E (2016) The useful life of inverter-based drive bearings: methods and research directions from localized maintenance to prognosis. *IEEE Ind Appl Mag* 22(4):63–73
22. Bunt WT (1941) Ball bearing noise and vibration. *The J Acoust Soc America*
23. Dowson P, Bauer D, Laney S (2009) Selection of materials and material related processes for Centrifugal compressors and steam turbines in the oil and petrochemical industry. In: Proceedings of the thirty-seventh turbomachinery symposium, pp 189–209
24. Elango P (2014) A review paper on methods of improvement of wear, corrosion and hardness properties of austenitic stainless steel 316L. *Int J Eng Res Rev* 2(4):18–23
25. Das A (1996) Metallurgy of failure analysis. McGraw-Hill
26. Bogdan M (2009) Measurement experiment, using Ni Usb-6008 data acquisition. *J Electri Electron Eng*
27. Agarwal C (2020) Bearing_Condition_monitoring_data.csv, chirag1236/bearing-condition-monitoring-data.csv
28. IRD Mechanalysis Limited, “IRD591 Accelerometer 4–20mA velocity output via 2 Pin MS Connector (datasheet)”, Sales Department, IRD Mechanalysis Limited, E8–14, Bhumi World Industrial Park, Thane—421311, Maharashtra, India

Bearing Fault Diagnosis Using Machine Learning Models



Shagun Chandrvanshi, Shivam Sharma, Mohini Preetam Singh, and Rahul Singh

Abstract The bearing serves as a crucial element of any machinery with a gearbox. It is essential to diagnose bearing faults effectively to ensure the machinery's safety and normal operation. Therefore, the identification and assessment of mechanical faults in bearings are extremely significant for ensuring reliable machinery operation. This comparative study shows the performance of fault diagnosis of bearings by utilizing various machine learning methodologies, including SVM, KNN, Linear Regression, Ridge Regression, XGB Regressor, AdaBoost Regressor, and Cat Boosting Regressor. Bearings are like the unsung heroes of the mechanical world, immensely supporting and guiding the smooth motion in everything, from your car's wheel to the propeller in a ship. However, like other mechanical components, over the course of time, the constant use of bearings can lead to wear and tear, which may ultimately result in a fault. Bearing faults can manifest in several ways, including vibration, noise, heat, and changes in lubrication that reduce the efficiency of a machine. Therefore, it is essential to regularly monitor the bearings and inspect them to detect any issues early on. The aim of this present work is to use the various ML methodology, and their application on the bearing's data to watch the condition of the machine's bearing. The present work is carried out in four phases. In the first phase, the data of various loads is collected. In the second phase, the data undergoes an Exploratory Data Analysis (EDA). During the third phase, the data undergoes both training and testing processes to evaluate its effectiveness. In the fourth and final

S. Chandrvanshi · M. P. Singh (✉)

Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

e-mail: mohinisingh2008@gmail.com

S. Chandrvanshi

e-mail: shagunrajputmrt@gmail.com

S. Sharma

Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India

R. Singh

Product Technology and Development Centre, Larsen & Toubro Ltd, Mumbai, India

phase, the model that gives the highest accuracy among all is chosen. The present approach is based on the various machine learning algorithms and their application.

Keywords Bearing · Gearbox · Fault diagnosis · Exploratory data analysis (EDA)

1 Introduction

Condition-based monitoring (CBM) is a crucial practice for maintaining the reliability of rotating machinery by identifying potential issues at an early stage. This approach helps prevent system failure by allowing for preventative measures to be taken in a timely manner. Bearing fault diagnosis is particularly important as bearings undergo distinct stages when subjected to different loads, including the healthy stage, damage stage, and extra damage stage. Early detection of faults is crucial in preventing total system failure. In this study, various machine learning models will be employed to conduct diagnosis of faults in bearing. These algorithms will analyze the behavior of bearings under varying loads and identify any deviations from normal functioning, using factors such as sound emissions, vibrations, and variations in temperature. The findings of this search can reveal the enhancement of predictive maintenance strategies to enhance the reliability and performance of rotating machinery.

Bearing faults are a common cause of machinery breakdown and can lead to equipment damage and high maintenance costs. To prevent such issues, early diagnosis and detection of faults are crucial. In the past few years, several methods have been introduced to detect bearing faults by analyzing vibration signals. One such approach is the utilization of statistical characteristics obtained from the vibratory signals. In this regard, an innovative way of diagnosing issues in bearings by utilizing statistical characteristics extracted from vibratory signals was submitted in [1]. The authors demonstrated the effectiveness of their approach in identifying healthy and faulty bearings through experimental data and compared their approach with existing methods, showing promising results. Their research suggests that their approach can be employed in condition monitoring of industrial machinery to prevent equipment failure and reduce maintenance costs.

Ball bearings serve as a fundamental component of machinery that revolves, and their faults can lead to equipment breakdown and high maintenance costs. Machine learning techniques have recently been suggested for the accurate diagnosis of ball bearing faults. Kankar et al. [2] investigated for ball bearing fault diagnosis by employing vibration signals and statistical characteristics obtained from the bearings. They utilized different types of techniques for machine learning, which include Artificial Neural Networks, Decision Trees, and Support Vector Machines, to categorize healthy and faulty bearings based on vibration signals and statistical features.

Their research showed that the SVM algorithm outperformed the other methods, demonstrating high accuracy in fault diagnosis. The study concludes that machine

learning methods can be effective in diagnosing ball bearing problems and are appropriate in condition monitoring of industrial machinery. The suggested technique was assessed on a set of vibration signals obtained from bearings having various types of defects and juxtaposed with conventional machine learning methods [3].

It presents a novel technique for locating bearing faults by fusing CNNs and SVM. It highlights the importance of identifying bearing issues early to prevent machinery failures as well as reduce maintenance costs. The authors collected signals generated by a bearing test bench while extracting features using CNN. These features were then fed into an SVM for fault diagnosis. The combined CNN-SVM approach for fault diagnosis outperformed other existing methods for accuracy and computational efficiency, according to the results of the proposed approach's evaluation on three bearing fault datasets. The study concludes that the suggested approach is capable of detecting bearing faults, enhancing machinery performance, and increasing reliability [4]. Deep learning algorithms have also been tested for the similar studies on bearing faults [5]. The paper introduces fundamental concepts of SVMs, such as kernel methods, and the geometry of linear classifiers, to provide a basic understanding of SVMs [6, 7].

In 2008, a comparative study to assess the effectiveness of Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to identify flaws in rolling element bearings using wavelet transform preprocessing. This study utilized vibration signals obtained from the bearings and extracted statistical features for classification. The features extracted were utilized to train and test SVM and ANN algorithms for the purpose of distinguishing healthy and faulty bearings. Results of the study demonstrated that SVM outperformed ANN for both accuracy and speed of finding errors. As a result, the study concludes that SVM is a viable option for detecting bearing faults through preprocessing with wavelet transforms [8]. They are using an SVM-based classifier to assess the performance of the feature ranking techniques [9]. Compare the performance of their method with conventional techniques and show its higher accuracy and robustness [4].

An innovative strategy for detecting issues in the bearings by applying the weighted KNN algorithm for fault diagnosis is also done. The Authors extracted analytical features from vibration data gathered from bearings and employed the K-nearest neighbor algorithm to classify healthy and faulty bearings. The added feature of assigning weights to the neighbors based on their distances improved both the precision and the dependability of the proposed approach. The authors tested their approach on experimental data and compared it with existing methods, demonstrating its effectiveness in bearing fault diagnosis. The study concludes that the suggested method is suitable for condition monitoring of industrial machinery, leading to improved maintenance and reliability. The research highlights the potential benefit of the weighted K-nearest neighbor for bearing fault diagnosis [10].

Heng and Nor [11] presented a statistical analysis approach for monitoring the state of bearings with rolling elements using sound and signaling from vibrations. The study concludes that statistically analyzing vibration and sound signals can be a reliable way to monitor bearing conditions as well as a tool for predicting faults in rotating machinery. This research highlights the importance of early identification

and detection of bearing problems to avoid system failure and enhance the reliability of industrial machinery.

To monitor equipment and diagnose faults, [12] performed a thorough review of the Support Vector Machine. The authors explained the fundamental concepts of SVM, its strengths compared to other machine learning techniques, and its various applications in predictive maintenance. They also conducted an extensive review of existing studies that used SVM for fault diagnosis tasks such as bearing, gear, and pump fault diagnosis. The review concluded that SVM is a promising tool for fault diagnosis in machinery due to its ability to handle complex and nonlinear data and its potential to achieve high accuracy and generalization performance. The review of literature emphasizes their significance for SVM monitoring the condition of machines and highlights its potential to enhance the reliability and safety of industrial machinery.

2 Methodology

2.1 SVM

A category of supervised machine learning approach is Support Vector Machine (SVM), which operates on the principles of statistical learning theory. The present research highlights that Support Vector Machine (SVM) is a potent approach to perform classification and regression tasks, particularly in scenarios where the data available is inadequate, as in the context of fault diagnosis. This research proposes the utilization of Support Vector Machines (SVMs) for several reasons, such as their capability to provide high precision, overcome overfitting, and work well with limited data for fault identification. Additionally, SVMs possess the potential to perform efficiently in less time and are suitable for handling high-dimensional data without overfitting due to their automatic regularization mechanism. Moreover, SVMs demonstrate a promising approach for various applications as they can effectively operate with small datasets.

To enhance the accuracy of prediction, our model utilizes multiple SVM kernels, including linear, sigmoid, fine Gaussian, medium Gaussian, and coarse Gaussian kernels.

2.2 SVM Kernels

The kernel function is a powerful mathematical tool that enables Support Vector Machines (SVM) to perform complex, nonlinear classifications of one-dimensional data in higher-dimensional spaces. Through the utilization of a kernel function on the input dataset, the Support Vector maps the original data into a higher-dimensional

space where the data can be more effectively separated. The kernel function can be chosen based on the nature of the data and the desired separation of the classes.

- **Linear Kernel Function**—The SVM utilizes the linear kernel function to transform the input data to a higher-dimensional space, which is a straightforward yet effective approach, making it easier to separate and classify. Its effectiveness lies in its ability to capture linear relationships between variables.
- **Polynomial**—It is also used to convert data into a space with a higher number of dimensions, its uniqueness lies in its capacity to capture intricate connections between data points, even when they are not linearly separable.
- **Radial**—It is used for classification and clustering. It measures the similarity between two points based on their distance between each other.
- **Quadratic**—Quadratic kernel is a type of radial basis function used for nonlinear classification. It maps data points into a higher-dimensional space to make them separable.

2.3 KNN

KNN, also known as K-nearest neighbor, is a machine learning algorithm that falls under the category of supervised learning techniques. This algorithm is used for making highly accurate predictions. However, KNN is used for classification and regression problems.

Classification problem:—When the targeted column or dependent variable has a fixed number of categories, then it comes under classification. We use two approaches to solve the classification problem, i.e., Euclidean distance and Manhattan distance.

Regression problem:—Whenever the targeted column or dependent variable is continuous in nature, it becomes a regression problem. In this, we take the average distance of the data points by choosing the k value, where k is a hyperparameter and its value is not fixed.

2.4 Decision Tree

A popular machine learning method for its ability to provide transparency, flexibility, and feature selection capabilities is the decision tree. Unlike other machine learning methods that are often seen as “black boxes,” decision trees generate a tree-like structure that is intuitive and easy to interpret. This characteristic is especially useful in industries like healthcare, law, and finance where interpretability is essential.

One of the key strengths of the decision tree is its versatility in handling both categorical and numerical data. This makes them highly adaptable to a wide range of problems, from predicting customer churn to detecting fraudulent transactions. Moreover, decision trees are highly robust to missing data and can handle noisy

data, making them ideal for real-world applications where data quality is not always perfect.

Another major advantage of the decision tree is its feature selection capabilities. By identifying the most relevant feature in the dataset, the decision tree can help researchers reduce the dimensionality of the data, leading to improved model accuracy and faster training times. This feature is particularly useful in domains where the cost of data acquisition or storage is high.

Overall, the decision tree is a powerful and effective machine learning technique that has found broad application across a variety of research domains. Their transparency, flexibility, and feature selection capabilities make them a popular choice for researchers who require an intuitive and effective algorithm for their work.

2.5 *Random Forest*

Another supervised class of machine learning algorithm is random forest. This algorithm is an advanced version of the decision tree algorithm. It is used to tackle the problem of overfitting. Overfitting is the limitation of the decision tree which can arise when a model performs well with training data but it fails to perform well with test data. Random forest has low bias and low variance.

Random forest is a bagging technique which does not impact by outliers as it gives equal opportunities to all the features by selecting randomly. It can handle both categorical and continuous data, as it is highly versatile in nature.

2.6 *Regression*

Regression is a statistical technique used to examine how one or more independent variables and a dependent variable are related to one another, by finding a mathematical equation that best describes the pattern of the data. Based on past observations, it can be used to predict future outcomes. Financial, economic, and social science fields use it extensively. It has various types, i.e., linear, ridge, logistic, polynomial, and lasso regression.

- **Linear Regression**—Linear regression is a mathematical method utilized to develop a model for the association between an independent variable and one or several dependent variables. It presupposes a straight-line association between the variables and strives to identify the optimal fit line that decreases the total sum of the squared discrepancies. It comprises two variations: positive linear regression and negative linear regression.
- **Ridge Regression**—It is a version of linear regression, in which a small bias(weight) is introduced for better prediction. Ridge considers square of weights and reduces complexity. It includes a penalty term in the loss function to address

multicollinearity and improve the model's generalization performance. It works by shrinking the coefficients toward zero, resulting in a more stable and less complex model.

3 Methodology of Machine Learning Model

Data Preparation—The dataset is loaded using the Pandas library. The train_test_split function from the scikit-learn module is used to divide the dataset into training and testing sets after it has been loaded. The target variable “Outcome” is removed from the feature set X , and assigned to the target variable y .

Initialization—Initialize different classifiers and regressors.

Training—Train each classifier and regressors on the training set using the fit method.

Prediction—Make predictions on the testing set for each classifier and regressor using the predict method of the ML models.

Accuracy Calculation—The accuracy of each classifier and the regressor is computed by utilizing the accuracy_score function from the scikit-learn library.

4 Techniques for Extracting and Selecting Features from Data

When using a machine learning approach, there are several statistical features that can be obtained from the data. These features work wonders in streamlining data analysis by efficiently filling in missing data and eliminating redundant information.

In order to acquire a thorough collection of characteristics, numerous statistical characteristics are derived. Further elaboration of them is provided as follows:

Mean—The statistical measure that characterizes the central tendency of a dataset is known as the mean value, which is often referred to as the average value of a signal. To determine the mean of a dataset, one must meticulously sum up every value within the collection, and subsequently divide the resulting summation by the total count of data points included within the set.

Variance—Variance is a statistical measure that describes the spread or distribution of data points within a dataset. It is calculated by determining the average squared distance of each data point from the mean of the entire dataset. It tells how much the data deviates from the mean value. By analyzing the variance of a dataset, we can gain insight into the level of variability within the data and make more informed decisions based on the trends and patterns that emerge.

Standard deviation—The standard deviation is a widely used statistical measure, it should be noted that it is not a direct indicator of the energy content present in a vibration signal; rather, it is a statistical parameter that quantifies the degree of variation or dispersion of the data points from the mean value. However, it can still

be used as a useful tool in analyzing vibration signals and detecting potential faults in mechanical systems.

Skewness—A statistical metric called skewness shows how asymmetrically the distribution of data is distributed. The skewed distributions which are positive have a longer tail on the right side, while the distributions which are negative have a longer tail on the left. The skewness is determined by a distribution's third central moment, which is influenced by the mean, standard deviation, and number of data points. The skewness equation is used to quantify the degree of asymmetry in the dataset, allowing any outliers or biases to be identified.

Kurtosis—Kurtosis is a statistical indicator that assesses how much the pointedness or levelness of a distribution deviates from a normal distribution.

Root Mean Square (RMS)—RMS also known as root mean square, is a mathematical measure that captures the degree of deviation from the mean by computing the square root of the average of the squared differences of each value from the mean. This measure is particularly useful in the analysis of signals and waveforms, as it provides a measure of the overall magnitude or energy content of the signal.

5 Relationship Between the Statistical Features

- The relationship between these features can be mathematically explained as follows:
- The square of the standard deviation represents variance.

$$\text{Standard Deviation} = \sqrt{\text{Variance}}$$

- Skewness, kurtosis, mean, variance, and standard deviation are all measures of the central tendency and spread of the data distribution.
- RMS is a measure of the overall deviation of the data distribution from the mean.
- Mean, variance, and standard deviation reveal information about the central tendency and spread, whereas skewness and kurtosis reveal information about the shape of the data distribution.
- A high skewness value indicates that the distribution is asymmetrical, while a low value indicates symmetry.
- A distribution that has a high value for kurtosis has very peaks, whereas a distribution that has a low value is flatter.
- The mean and variance can be used to describe a normal distribution, while skewness and kurtosis can be used to identify non-normal distributions.
- RMS can be used to quantify the overall error in a prediction model.

6 Data Description

This study involves the utilization of a dataset comprising three stages of bearing faults, with the inclusion of a healthy bearing in the first stage for comparative analysis. The dataset comprises data at different loads, ranging from 0 to 4 k, for every stage of bearing faults. Additionally, the distribution of data within each stage is not uniform. Three phases make up the dataset, where the first stage represents a healthy bearing, the second stage includes data of bearings with certain degrees of damage, and the third stage comprises data of bearings that have undergone more severe damage (see Fig. 1).

- **First Stage**

The initial stage of the dataset corresponds to the healthy bearing stage, with the inclusion of readings at various loads of 0, 1, 2, 3, and 4k.

- **Second Stage**

The second stage of the dataset encompasses data of bearings that have undergone some degree of damage, with recordings available at different loads of 0k, 1k, 2k, 3k, and 4k, depicting the damage to the bearings.

- **Third stage**

The third stage of the dataset includes data of bearings that have undergone an additional level of damage beyond the second stage. Recordings of these bearings are available at various loads of 0, 1, 2, 3, and 4k, providing insights into the nature and extent of the extra damage sustained by the bearings.

S.No	UNIT	BEARING DATA														
		1st Stage (new bearing)				2nd Stage (some amount of damaged bearing)				3rd stage (extreme amount of damaged bearing)						
		0k	1k	2k	3k	4k	0k	1k	2k	3k	4k	0k	1k	2k	3k	4k
1	0	-0.0003865	0.000312	8.26E-06	0.001426	0.00091	-1.41722	-0.98888	-10.582	-6.79987	-2.62613	-10.582	-9.92308	-1.12564	-0.25735	-0.55151
2	0.024	-0.0003865	0.000312	-0.00051	0.001426	0.00091	-0.63279	-1.88169	0.369681	-7.26046	-1.48947	0.750285	-0.82245	-0.69085	-0.4612	-0.45862
3	0.048	-0.0003865	-0.0002	8.26E-06	0.001426	0.000426	-0.21933	-1.36304	1.197979	-5.62967	-1.25466	0.976067	-0.25864	0.75665	-0.79406	-0.83355
4	0.072	0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.68827	-1.16693	-1.06887	-3.12543	-1.85976	-10.0873	-1.19273	-0.47281	-0.9992	-0.85341
5	0.096	-0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-1.24692	-1.10758	-10.582	-2.04038	-2.20682	-10.582	-1.62494	-0.39669	-0.40443	0.127127
6	0.12	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-1.45335	-0.97985	-10.582	-6.58957	-2.63387	-10.582	-1.14886	-1.03404	-0.18768	0.103903
7	0.144	-0.0003865	0.000312	-0.00051	0.001426	0.00091	-1.00953	-1.75138	0.2226	-7.32497	-1.33467	0.536113	-0.67924	-0.89212	-0.51668	-0.73472
8	0.168	-0.0003865	-0.0002	-0.00051	0.001426	0.00091	-0.71665	-1.45722	1.043157	-5.6	-1.30369	1.154113	-0.80052	-0.65085	-0.85083	-0.62247
9	0.192	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.85857	-1.25853	0.004559	-3.1422	-1.50238	-10.582	-0.64569	-0.51409	-0.68311	-0.38766
10	0.216	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-1.20563	-0.51668	-10.582	-2.89577	-2.32422	-10.582	-0.83793	-0.64162	-0.31541	-0.19155
11	0.24	-0.0003865	-0.0002	8.26E-06	0.00091	0.000426	-1.13467	-1.0779	-10.582	-5.81933	-2.53452	-10.582	-0.91405	-1.02888	0.087131	0.058747
12	0.264	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.76568	-1.07146	0.199377	-7.29917	-1.38239	0.577401	-0.61215	-1.16693	-0.32315	-0.6973
13	0.288	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.65601	-1.47012	1.039286	-6.41926	0.30631	1.240553	-1.06758	-0.5399	-0.8947	-0.42378
14	0.312	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.76955	-0.59538	-0.69601	-3.21187	-1.71655	-10.582	-1.23144	0.169703	-0.50249	-0.44055
15	0.336	-0.0003865	-0.0002	8.26E-06	0.001426	0.00091	-0.87277	-0.9605	-10.582	-2.17456	-2.70096	-10.582	-1.68042	-0.80568	-0.23542	-0.01737
16	0.36	-0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-0.75149	-1.29853	-10.582	-6.39217	-2.83901	-9.70149	-0.77987	-1.1308	-0.06382	-0.42636
17	0.384	0.0003865	-0.0002	8.26E-06	0.00091	0.00091	-0.55667	-1.06371	0.520633	-8.4938	-2.00039	0.248404	-0.93598	-0.87793	-0.05608	-0.59022
18	0.408	-0.0003865	-0.0002	0.000524	0.001426	0.00091	-0.80181	-1.49463	1.465047	-5.62451	-0.3954	0.954113	-0.92593	-0.36056	-0.80181	-0.57344

Fig. 1 Overall dataset utilized [29]

7 Comparative Study of Statistical Features

- **First Stage**

Table 1 shows the statistical features of a bearing dataset from the first stage, which is the healthy stage. The mean values for each set of 1000 data points are close to zero, indicating that there is no significant bias in the data. Since the variance is quite low, it is likely that the data are closely grouped around the mean. The root mean square values also show that the data is well-concentrated around the mean.

The skewness values are close to zero for the first two sets of data, indicating a symmetric distribution. However, the skewness becomes negative for the remaining sets of data, demonstrating the distribution's left-handed skewness. Finally, the kurtosis values are mostly above the normal range of -3 to $+3$, indicating that the signals have high kurtosis.

The first stage bearing dataset's statistical characteristics generally indicate that the data is consistently distributed, firmly clustered around the mean, with some minor skewness and heavier tails. Overall, the statistical features of the first stage bearing dataset suggest that the data is regularly distributed and tightly clustered around the mean, with some slight skewness and heavier tails.

- **Second Stage**

Table 2 presents the statistical features of a bearing dataset from the second stage, which is considered the damage stage, at different loads. The mean values of the datasets are considerably different from zero and show variations among the different sets, indicating the presence of a significant bias in the data. The variance values are relatively large, indicating that the data is more spread out and not as tightly clustered around the mean as in the first stage.

The root mean square values also indicate that the data is more dispersed compared to the first stage. The skewness values are close to zero for the first and third sets, indicating a symmetric distribution. However, the skewness turns negative for the

Table 1 Statistical features of first stage at different load

Statistical feature	1st stage (0 k)	1st stage (1 k)	1st stage (2 k)	1st stage (3 k)	1st stage (4 k)
Mean	-0.001285	-0.000393	-0.000203	0.001239	0.00081
Variance	3.155e-07	8.365e-08	6.723e-08	7.874e-08	6.109e-08
Root mean square	0.001403	0.000488	0.000329	0.001270	0.000867
Skewness	0.0502	0.0006	-0.2728	-0.2990	-0.6284
Standard deviation	0.000562	0.000289	0.000259	0.000281	0.000247
Kurtosis	2.1385	2.3659	1.39714	4.4217	4.3355

Table 2 Statistical features of second stage at different load

Statistical feature	2nd stage (0 k)	2nd stage (1 k)	2nd stage (2 k)	2nd stage (3 k)	2nd stage (4 k)
Mean	-4.661462	-4.525684	-4.848079	-4.072137	-1.634369
Variance	29.975	34.509	9.547	2.798	0.571
Root mean square	7.190577	7.415622	5.749048	4.402333	1.800769
Skewness	-0.02424	0.07384	-0.05923	-0.0974	-0.4129
Standard deviation	5.475	5.875	3.090	1.672	0.756
Kurtosis	1.1257	1.2150	2.3431	2.0183	2.6683

remaining sets, indicating a leftward skewed distribution. The standard deviation values increase with the load, indicating a widening spread of the data.

In summary, the statistical features of the second stage bearing dataset suggest that the data is not normally distributed and exhibits a considerable bias. In comparison to the previous stage, the data is more dispersed and has a higher variance and standard deviation. The presence of negative skewness and heavier tails in some sets indicates that the distribution may be skewed to the left, further emphasizing the damage to the bearing and degradation of the system.

- **Third Stage**

The statistical features of the bearing dataset from all three stages show that as the damage becomes more severe, the data becomes less symmetric, more spread out, and less dispersed. In Table 3 the variance is relatively large for the first two stages but decreases significantly for the third stage. The root mean square values also decrease with increasing damage severity. The skewness is close to zero for the first and third stages, but becomes negative for the later stages, indicating severe left-skewness. The standard deviation values decrease as the damage becomes more severe. Finally, the kurtosis values increase significantly for the later stages, indicating more peaked and heavier tails, with an extremely high kurtosis value for the fifth set of data. These statistical properties shed light on the bearing's deterioration process and may be applied to the creation of efficient monitoring and preventative maintenance plans.

8 Result

Table 4 represents the precision of diverse machine learning techniques on the bearing dataset across all three stages of degradation. The models were trained to categorize the condition of the bearing based on the given data. The SVM model with a linear kernel function achieved the highest accuracy of 98.5%, followed closely by the models with quadratic and ridge kernels, achieving 98.3% and 98.1%, respectively.

Table 3 Statistical features of third stage at different load

Statistical feature	3rd stage (0 k)	3rd stage (1 k)	3rd stage (2 k)	3rd stage (3 k)	3rd stage (4 k)
Mean	-4.872336	-4.718315	-4.569394	-1.771385	-0.737018
Variance	30.364	29.991	4.315	0.648	0.318
Root mean square	7.355545	7.228699	5.019388	1.945778	0.928224
Skewness	0.228	-0.03344	0.8618	-0.78598	-6.47081
Standard deviation	5.511	5.477	2.077	0.805	0.564
Kurtosis	1.0700	1.0852	2.4529	3.7399	57.9134

The other models like KNN, Random Forest, Linear Regression, XGB Regressor, Adaboost Regressor, and Random Forest also achieved good accuracy ranging from 97.2 to 98.2%. However, the Catboosting Regressor model achieved the lowest accuracy of 94.6%.

Overall, the high accuracy rates of these models demonstrate their effectiveness in predicting the health status of the bearing across all three stages of degradation (see Fig. 2).

The models can be used for developing predictive maintenance strategies to detect and mitigate any potential damage in the early stages.

Table 4 Accuracy with distinct approaches

S. No.	Model name	Accuracy (%)
1	Linear kernel function [SVM]	98.5
2	Polynomial kernel [SVM]	98.2
3	Radial kernel [SVM]	98.2
4	Quadratic kernel [SVM]	98.3
5	KNN	98.2
6	Decision tree	95.9
7	Random forest	97
8	Linear regression	97.3
9	Ridge regression	98.1
10	XGB regressor	97.2
11	Adaboost regressor	97.5
12	Catboosting regressor	94.6

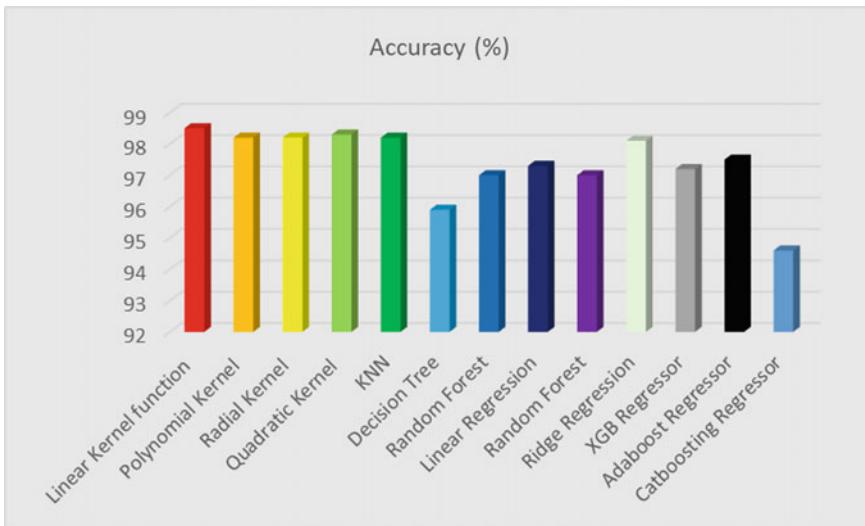


Fig. 2 Accuracy of different models

9 Conclusion

In conclusion, the analysis of the bearing dataset reveals significant changes in data characteristics as the bearing undergoes degradation. Statistical features indicate tighter clustering in the healthy stage and increased spread and skewness in damaged stages, with severe damage showing even greater dispersion. SVM models exhibit high accuracy for classification, particularly with complex datasets, while different kernel functions are effective for linear and nonlinear separation. KNN is computationally expensive but performs well, while decision trees have lower accuracy but random forest improves with ensemble learning. Linear regression is suitable for regression tasks. SVM with a linear kernel achieves the highest accuracy overall. These findings can inform monitoring and predictive maintenance strategies for bearings, saving costs by identifying potential failures in advance.

10 Summary

This study compares multiple machine learning methods for diagnosing faults in bearings of machinery. The research focuses on monitoring bearing conditions and detecting faults early on. The study involves data collection, exploratory data analysis, training and testing, and selecting the most accurate model. Machine learning techniques such as SVM, KNN, decision trees, random forest, and regression are employed for fault diagnosis and classification tasks. These approaches contribute

to improved maintenance and reliability of industrial machinery through early fault detection. Feature extraction and selection play a vital role in analyzing data, and statistical measures provide valuable insights for informed decision-making. The present research highlights the effectiveness of Support Vector Machine (SVM) as a potent approach for classification and regression tasks, especially in scenarios where data is limited, such as fault diagnosis.

11 Future Scope

Incorporating deep learning methods like Convolutional Neural Networks (CNN) with the goals of improving accuracy, real-time monitoring, and multimodal data integration would broaden the paper's potential application areas for bearing problem diagnostics. The paper can go into real-time monitoring of bearing conditions, which enables preventative maintenance and prompt issue diagnosis. This could entail creating an online monitoring system that continuously examines sensor data in real-time and sends out prompt alerts or cautions when abnormal behavior is found. Research can investigate methods to enhance computing effectiveness, reduce latency and ensuring the viability.

Acknowledgements The TIH-IoT CHANAKYA UG/PG Internship/Fellowship Program 2022, with Larsen & Toubro Ltd as Industrial mentors, provided invaluable support for this work. The Authors extend their sincere gratitude to the Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India, for their generous provision of the necessary resources and infrastructure that enabled the successful completion of this project.

References

1. Altaf M, Akram T, Khan MA, Iqbal M, Ch MMI, Hsu CH (2022) A new statistical features based approach for bearing fault diagnosis using vibration signals. Sensors 22(5):2012
2. Kankar PK, Sharma SC, Harsha SP (2011) Fault diagnosis of ball bearings using machine learning methods. Expert Syst Appl 38(3):1876–1886
3. He M, He D (2017) Deep learning based approach for bearing fault diagnosis. IEEE Trans Ind Appl 53(3):3057–3065
4. Han T, Zhang L, Yin Z, Tan AC (2021) Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine. Measurement 177:109022
5. Zhang S, Zhang S, Wang B, Habetler TG (2020) Deep learning algorithms for bearing fault diagnostics—a comprehensive review. IEEE Access 8:29857–29881
6. Cristianini N, Shawe-Taylor NJ (2000) An introduction to support vector machines. Cambridge University Press, Cambridge
7. Widodo A, Yang B-S (2007) Review on support vector machine in machine condition monitoring and fault diagnosis. Mech Syst Signal Process 21:2560–2574
8. Tyagi CS (2008) A comparative study of SVM classifiers and artificial neural networks application for rolling element bearing fault diagnosis using wavelet transform preprocessing. Int J Mech Mechatron Eng 2(7):904–912

9. Vakharia V, Gupta VK, Kankar PK (2016) A comparison of feature ranking techniques for fault diagnosis of ball bearing. *Soft Comput* 20:1601–1619
10. Sharma A, Jigyasu R, Mathew L, Chatterji S (2018) Bearing fault diagnosis using weighted K-nearest neighbor. In: 2018 2nd international conference on trends in electronics and informatics (ICOEI), May, IEEE, pp 1132–1137
11. Heng RBW, Nor MJM (1998) Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Appl Acoust* 53(1–3):211–226
12. Widodo A, Yang BS (2007) Support vector machine in machine condition monitoring and fault diagnosis. *Mech Syst Signal Process* 21(6):2560–2574
13. Wang B, Zhang X, Xing S, Sun C, Chen X (2021) Sparse representation theory for support vector machine kernel function selection and its application in high-speed bearing fault diagnosis. *ISA Trans* 118:207–218
14. Patle A, Chouhan DS (2013) SVM kernel functions for classification. In: 2013 international conference on advances in technology and engineering (ICATE), January, IEEE, pp 1–9
15. Feizizadeh B, Roodposhti MS, Blaschke T, Aryal J (2017) Comparing GIS-based support vector machine kernel functions for landslide susceptibility mapping. *Arab J Geosci* 10:1–13
16. Brenning A (2023) Interpreting machine-learning models in transformed feature space with an application to remote-sensing classification. *Machine Learn* 1–17
17. Xu G, Liu M, Jiang Z, Söffker D, Shen W (2019) Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. *Sensors* 19(5):1088
18. Vakharia V, Gupta VK, Kankar PK (2017) Efficient fault diagnosis of ball bearing using ReliefF and random forest classifier. *J Braz Soc Mech Sci Eng* 39(8):2969–2982
19. Kamat P, Marni P, Cardoz L, Irani A, Gajula A, Saha A, Kumar S, Sugandhi R (2021) Bearing fault detection using comparative analysis of random forest, ANN, and autoencoder methods. In: Communication and intelligent systems: proceedings of ICCIS 2020, Springer, Singapore, pp 157–171
20. Li C, Sanchez RV, Zurita G, Cerrada M, Cabrera D, Vásquez RE (2016) Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mech Syst Signal Process* 76:283–293
21. Guo X, Chen L, Shen C (2016) Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement* 93:490–502
22. Soualhi A, Medjaher K, Zerhouni N (2014) Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression. *IEEE Trans Instrum Meas* 64(1):52–62
23. Zhang R, Tao H, Wu L, Guan Y (2017) Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. *IEEE Access* 5:14347–14357
24. Lei Y, He Z, Zi Y (2009) Application of an intelligent classification method to mechanical fault diagnosis. *Expert Syst Appl* 36(6):9941–9948
25. Lu W, Li Y, Cheng Y, Meng D, Liang B, Zhou P (2018) Early fault detection approach with deep architectures. *IEEE Trans Instrum Meas* 67(7):1679–1689
26. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. *IEEE Trans Neural Networks* 13(2):415–425
27. Esakimuthu Pandarakone S, Mizuno Y, Nakamura H (2019) A comparative study between machine learning algorithm and artificial intelligence neural network in detecting minor bearing fault of induction motors. *Energies* 12(11):2105
28. Abboud D, Elbadaoui M, Smith WA, Randall RB (2019) Advanced bearing diagnostics: a comparative study of two powerful approaches. *Mech Syst Signal Process* 114:604–627
29. Agarwal C (2023) Bearing_Condition_monitoring_data.csv. chirag1236/bearing-condition-monitoring-data.csv

A High-Payload Image Steganography Based on Shamir's Secret Sharing Scheme



Sanjive Tyagi, Maysara Mazin Alsaad, and Sharvan Kumar Garg

Abstract This paper introduces an image steganography system that hides a high capacity of confidential data by utilizing a distributed computing scheme that offers the embedding of smaller, equally segmented confidential images across multiple cover images. The proposed strategy developed a strong security mechanism associated with self-synchronized methods by utilizing Shamir's Secret Sharing (SSS) Scheme. To ensure extraordinary security for hidden confidential data within the cover image, double-layer security has been designed. In the primary layer, the Secret Distributing Scheme (SDS) divides a single larger secret image into identically smaller sub-images, and the SSS scheme then creates an encoded structure of smaller sub-images for assigning and restoring the distributed decomposed confidential images within the proposed shareholders. The SSS scheme will reorder the distributed confidential decomposed images into a predefined order by the authorized recipient at the moment of revealing secrets by interpolating polynomials. The next layer is steganography, where the image is segmented into 2×2 squares of pixels and navigated in a crisscross way; the pixel-value difference is computed among non-overlapping Red, Green, Blue (RGB) pixels in a diagonal direction within a focused square of pixel area. Secret bits are hidden within the RGB pixels of the cover image by using the presented innovative pixel-value differencing (PVD) strategy. The results of the research indicate that the presented system provides an inventive mechanism in the form of a double layer of security and vastly enhances embedding capability.

Keywords Secret distributing · Image steganography · Pixel-value differencing · Embedding capacity · Authentication · Steganalysis

S. Tyagi (✉) · S. K. Garg

Department of CSE, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India

e-mail: tosanjive@gmail.com

M. M. Alsaad

Department of Computer Science, Gujarat University, Ahmedabad, India

S. K. Garg

Department of CSE, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India

1 Introduction

Security is important everywhere to protect data from unauthorized access. Thus, internet privacy is important. As a result, techniques for protecting information have been developed, including passwords, fingerprinting, eye-lock, and secret keys. Copyright act violations, hacking, and other forms of data exploitation are on the rise in today's internet-based society. Furthermore, a rival who wishes to obstruct the transmission or reception of information may intentionally alter the data carried by an open network channel. Because of this, it is crucial to utilize digital security methods to encrypt private data or disguise it as a different type of file or model. That is why cutting-edge encryption and steganography are anticipated to play a significant role in the framework used to safeguard digital data. One of the cryptography's key purposes is to safeguard secret data by making it unintelligible. However, steganography is preferred since it conceals the secret data within another digital file, reducing the possibility of the hidden data being discovered. Encrypted information has the potential downside of raising suspicions and being susceptible to influence from non-authorized parties. Steganography is an emerging field of study with the ultimate aim of offering cutting-edge improvements to the infrastructure used to safeguard sensitive data [1].

To conceal confidential information behind a grayed cover image with a large payload, Wu and Tsai [2] invented a PVD technique. The presented method is based on the premise that the same amount of secret bits need not be installed in every pixel. The study found that the differences between groups of closed pixels determined the total number of embedded bits. There was also evidence to imply that the PVD technique might offer a significantly larger embedding capacity without sacrificing the quality of the stego digital picture. While smooth regions have lower differences between neighboring pixels, the opposite is true for areas near edges. The more bits can be covertly implanted, the higher the PVD.

In [3], the author proposes a steganography technique wherein an authorized sender, aided by a collection of n members, implants the hidden picture. In order to calculate the amount of embedded bits using three-pixel-value differencing, [4] proposes a multi-pixel differencing approach. In [5], the author proposed using Least Significant Bit (LSB) replacement for multi-pixel-value differencing steganography. Two improved PVD techniques were proposed in [6] for use with the block-based concealment strategy. The author of [7] proposed an LSB exchange between adjacent blocks of four pixels. The author of [8] described a steganographic method that uses the differential pixel values of multimedia files. PVD-based steganography paired with an LSB method was provided by the author [9] to increase the concealment bound. An improvement is achieved by Five Pixel Pair Differencing [10].

In the proposed scheme, color images are used as the cover medium to increase the embedding capacity of secret data, and a novel PVD-based algorithm is being designed.

2 Related Work

In [11], the authors offer a method for distributing a single secret image among several different shadow images. According to [12], using block-Discrete Cosine Transformation (DCT) to conceal information in many stock photos is one method of a distributed image steganographic strategy. Multiple confidential images can be concealed within a single visible image using the DWT-based steganography method described by the author in [13]. The stego images provided in [14] are more resistant to multiple steganalysis assaults because of the dispersed technique used to generate them. Intra-block variation was used in a steganographic construction [15]. For the purpose of image steganography, the pixel-connecting approach was introduced [16]. In [17], the author proposes a method of covertly sharing information among a set of n individuals. PVD steganography is introduced in [18], whereby confidential bits are encoded in a set of two adjacent pixels (pixel and pixel + 1). The approach described in [19] uses visual cryptography to enable the secure and anonymous transmission of patient information across a network. The author in [20] uses Shamir's covert sharing technique to provide a method for transferring medical images without compromising their integrity. In [21], a method for communicating confidential images based on a polynomial is paired with the visual cryptography method. In [22], two techniques, PVD-based steganography and LSB steganography, are introduced together. The author provided a method of hashing using quadratic probing in [23], which is useful for encoding secret byte sequences. The author of [24] presented a secure image steganography with a large capacity for embedding.

This study provides a high-payload image steganography method that spreads out secret information over several carrier files using Shamir's Secret Sharing (SSS) Scheme without degrading image quality.

3 Proposed Steganography Technique

The presented PVD-based distributed steganography secured with Shamir's Secret Sharing methodology offers the embedding of smaller, equally split confidential images across several cover images. The risk of a steganography attack is increased with conventional methods of hiding a single confidential digital image by embedding it secretly in one digital cover file. To uncover such concerns, a Secret Distributing Scheme (SDS) with a novel PVD strategy is presented along with a double-layer security system. Smaller shares increase efficiency and transmission; hence, the concept of image steganography can be enhanced to create invisible communication of secret data between the sender and receiver.

3.1 Secret Distributing Scheme (SDS)

In this subsection, a mechanism of distributing confidential images across k number of members has been discussed. Considering the objective is to improve the capacity of secret bits to be hidden, input data to the algorithm is organized as secret data D , which is decomposed into k number of equal partitions as $\{D_1, D_2, D_3, \dots, D_k\}$, and communicated to k number of recipients as $\{R_1, R_2, R_3, \dots, R_k\}$ and to be embedded inside k number of cover images denoted by $\{C_1, C_2, C_3, \dots, C_k\}$, and after the embedding process, the obtained stego images are as given $\{SI_1, SI_2, SI_3, \dots, SI_k\}$.

3.2 Shamir's Secret Sharing (SSS) Scheme for Protecting Hidden Secret Information

An effective secret sharing method, Shamir's secret sharing divides up sensitive data among a group of people in such a way that the secret cannot be revealed until the majority of the group work together to combine their knowledge. Our proposed approach uses Shamir's Secret Sharing (SSS) Scheme to create an encoded structure of smaller sub-images for assigning and restoring the distributed, decomposed confidential images within the proposed shareholders. Further, the SSS Scheme verifies stego image files on the receiving end prior to initiating the extraction procedure during the disclosure phase.

- Only valid stego images are utilized in the extraction process, and if one is not received, an attempt is submitted to be resent.
- The proposed SSS technique mathematically split the secret key, S into k number of secret sub-keys, for sharing among K number of shareholders $\{R_1, R_2, R_3, \dots, R_k\}$, represented by S_k as $\{X_k, Y_k\}$, where X_k denotes shareholder number and Y_k denotes split secret sub-key attached with each stego image. Toward recipient side secret key, S can be revealed by reassembling S_k at the time of extraction by valid recipient R_k ($1 \leq k \leq N$).

3.2.1 Model to Demonstrate the Presented Scheme

Number of shareholders or recipients $\{R = R_1, R_2, R_3, R_4, \dots, R_N\}$ in proposed example, N is taken as 4, and N and m are also taken equal.

Steps of developed algorithm for authentication are shown as given below:

- Let us consider secret key S is 11, taken prime number arbitrarily and create polynomial as given

$$f(x) = 11.x^0 + 7.x^1 + 3.x^2 + 1.x^3$$

Here a_1 is 7, a_2 is 3, a_3 is 1, taken arbitrarily as prime number and secret key S denoting a_0 is 11, taken arbitrarily as prime number

Compute decomposed Y_k for each recipient R_k , $1 \leq k \leq 4$ by utilizing the (Lagrange Polynomial) Eq. 1.

$$f(x) = \sum_{j=0}^N y_j \prod_{m=0, m \neq j}^N \frac{x - x_m}{x_j - x_m} \quad (1)$$

- Authorized recipient gather secret sub-key Y_k attached with each stego image from m shareholders, where m and N are taken equal to 4 as $\{x_k(\text{public_key}), Y_k(\text{secret sub key})\}$ and computed values for each recipient are given.

R_1 is $\{x_1(1), Y_1(22)\}$,

R_2 is $\{x_2(2), Y_2(45)\}$,

R_3 is $\{x_3(3), Y_3(86)\}$,

R_4 is $\{x_4(4), Y_4(151)\}$

- Disseminate m stego images to m receivers via m unreliable nodes.
- Now secret key S is computed toward recipients' side using Eq. 1 by interpolating polynomial.

$$x^3 + 3 \cdot x^2 + 7 \cdot x^1 + 11 \quad (2)$$

Here secret key is obtained from free coefficient shown in Eq. 2.

- The shared stego images are accepted for extraction if the computed secret key S is valid and equal to 11 in this example; otherwise, a request is made to resent the stego images.

3.3 Proposed PVD-Based Steganography Method

In this section, we have developed the second layer of security, viz., the pixel-value differencing (PVD)-based steganography algorithm. The presented approach considers that the cover image is made up of pixels, which are organized in the form of a matrix of dimension $L \times B$. Every cell of an image is made up of 24 bits (RGB).

In the presented approach, the concept of choosing the embedding locations is as follows: In case 1, if intensity variation is less than 9, then this smooth region is kept away from the embedding process; in case 2, if the area with the next higher-intensity variation embeds 1 bit in each RGB element of pixel; in case 3, if the area with the next higher-intensity variation embeds 2 bits in each RGB element of pixel; and in

case 4, if the area with the next higher-intensity variation embeds 3 bits in each RGB element of pixel.

3.3.1 PVD-Based Embedding Process

Here proposed PVD-Based Embedding Process is being described.

Input: Cover image having length and breadth ($L \times B$) and a secret image with an appropriate size of ($u \times v$) that can be embedded in the cover image.

Output: Stego image having length and breadth ($L \times B$) obtained after embedding secret image having length and breadth ($u \times v$) into cover image with length and breadth ($L \times B$).

Algorithm for Embedding Process

- 1: For pixel = $\text{pixel}_{1,1}$ to $\text{pixel}_{L,B}$, where $L \times B$ is the cover image's dimensions and L and B are taken equal.
- 2: Choose a non-overlapping pair of pixels from the diagonal of a 2×2 grid.
- 3: Calculate PVD diff_c between two $\text{pixel}_{i,j}$ and $\text{pixel}_{i+1,j+1}$ for Red, Green, Blue (RGB) component, where c denotes (RGB).
- 4: Using the LSB approach, hidden bits are encoded in the image based on the differences in pixel values between the three primary colours (RGB). The embedding bit count is split into four distinct categories.
 - i. If PVD diff_c lies in range 0 and 8, then no secret bit is embedded.
 - ii. If PVD diff_c lies in range 9 and 16, then one secret bit is embedded in red, green, blue component of that pixel.
 - iii. If PVD diff_c lies in range 17 and 24, then two secret bits are embedded in red, green, blue component of that pixel.
 - iv. If PVD diff_c lies in range 25 and 255, then three secret bits are embedded in red, green, blue component of that pixel.
- 5: Pick secret bits into variable Bites-to-Embed as per case of Step 4.
- 6: Embed the secret bit(s) by picking the pair of pixels from one of four categories determined in Step 4. This is done separately for the R, G, and B channels of that pixel. Target embedding pixel is determined out of $\text{pixel}_{i,j}$ and $\text{pixel}_{i+1,j+1}$ on the basis of their decimal value, so compute $d1 = \text{decimal value of } \text{pixel}_{i,j}$ and $d2 = \text{decimal value of } \text{pixel}_{i+1,j+1}$.
 - i. If case i of Step 4 is true then secret bits are not hidden.
 - ii. If case ii of Step 4 is true then, if $d1 < d2$ then embed three bits by replacing one bit using 1 LSB into $\text{pixel}_{i,j}$ (Red component), $\text{pixel}_{i,j}$ (Green component), and $\text{pixel}_{i,j}$ (Blue component) respectively, else embed three bits into $\text{pixel}_{i+1,j+1}$ by replacing one bit using 1 LSB into $\text{pixel}_{i+1,j+1}$ (Red component), $\text{pixel}_{i+1,j+1}$ (Green component), and $\text{pixel}_{i+1,j+1}$ (Blue component) respectively.

- iii. If case iii of Step 4 is true then, if $d_1 < = d_2$ then embed six bits by replacing two bits using 2 LSB into $\text{pixel}_{i,j}$ (Red component), $\text{pixel}_{i,j}$ (Green component), and $\text{pixel}_{i,j}$ (Blue component) respectively, else embed six bits into $\text{pixel}_{i+1,j+1}$ by replacing two bits using 2 LSB into $\text{pixel}_{i+1,j+1}$ (Red component), $\text{pixel}_{i+1,j+1}$ (Green component), and $\text{pixel}_{i+1,j+1}$ (Blue component) respectively.
 - iv. If case iv of Step 4 is true then, if $d_1 < = d_2$ then embed nine bits by replacing three bits using 3 LSB into $\text{pixel}_{i,j}$ (Red component), $\text{pixel}_{i,j}$ (Green component), and $\text{pixel}_{i,j}$ (Blue component) respectively, else embed nine bits into $\text{pixel}_{i+1,j+1}$ by replacing three bits using 3 LSB into $\text{pixel}_{i+1,j+1}$ (Red component), $\text{pixel}_{i+1,j+1}$ (Green component), and $\text{pixel}_{i+1,j+1}$ (Blue component) respectively.
- 7: Calculate PVD E_{diff_c} for a chosen pair of embedded pixels.
- 8: If the PVD E_{diff_c} of a given pair of embedded pixels does not fall within the same range as the four-case scenario established in Step 4, the pair of pixels taken is not used for embedding.
- 9: Steps 1–9 should be repeated until every possible pair of non-overlapping blocks along the diagonal has been chosen.
- 10: end

PVD-Based Extracting Process can be achieved by reverse processing the above PVD-Based Embedding Process, and a secret image with dimension $(u \times v)$ is obtained as output.

The above procedures are intended to accomplish a novel steganography model designed according to PVD-based steganography without influencing the aesthetic quality of the stego image.

4 Results and Experimental Findings

Here, we examine the effectiveness of the presented PVD steganography technique using a wide range of PNG and BMP files. Software is being developed in NetBeans IDE 8.2 using Java, and it is having real-world ramifications. We tested out the presented PVD-based algorithm on a variety of example cover images [25], some of which are presented in Fig. 1. Tables 1 and 2 display the computation outcomes of an experiential study based on the inspection and validation of a proposed PVD-based method. Anaconda and Python-3 were used to calculate PSNR values.

4.1 Robustness and Varying Embedding Capacity

It is observed from the computations in Table 1 that an average embedding bit rate of 3.38 bits per pixel block (BPPB) is obtained based on selecting consecutive 2

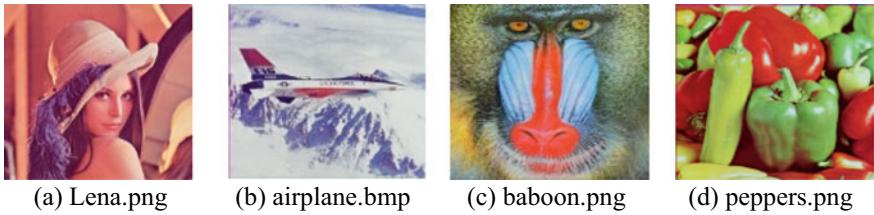


Fig. 1 Sample cover images [25] used to hide secret image

$\times 2$ pixel blocks; an embedding bit rate of 3.00 BPPB is acquired by selecting odd or even locations of consecutive blocks; and by choosing each of the cover image's subsequent blocks, a greater embedding bit rate of 6.00 BPPB is obtained, as shown in Fig. 2 graphically. Henceforth, our presented approach has multiple, varying embedding capacities.

It can be seen that changeable embedding rates of 1.50 bits per pixel (BPP), 0.75 BPP, 0.75 BPP, and 0.375 BPP can be managed by selecting a different position for the target (2×2) pixel square of the cover image, as calculated in Table 1 and graphically shown in Fig. 3. Thus, keeping in mind the trade-off between imperceptibility and embedding capacity, anyone from any of the four cases can be taken as per requirement. Hence, the presented approach overcomes the trade-off; convincingly, our introduced technique is robust.

4.2 Comparative Analysis

A comparison was made between well-known image steganographic schemes and the proposed scheme in terms of their embedding capacity and their PSNR values. Observations from computations are given as follows.

4.2.1 Improved Visual Quality

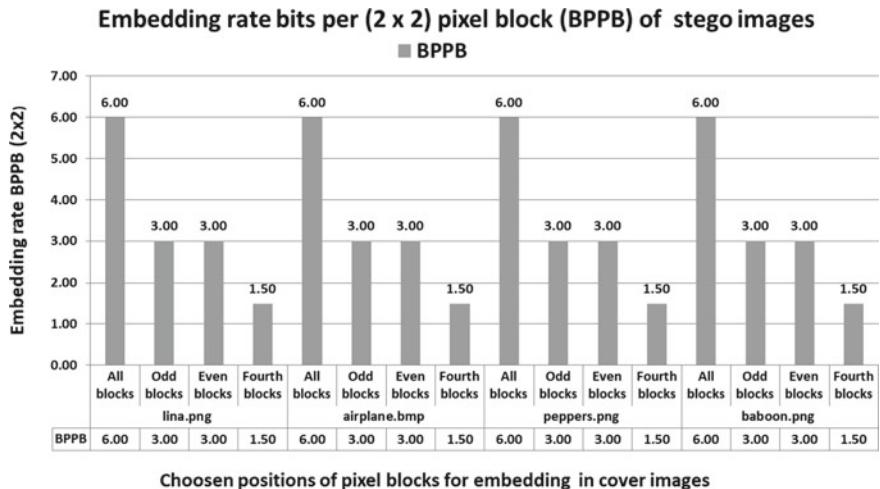
On comparing PSNR values of the proposed approach with PSNR values of existing PVD image steganographic schemes (Chang et al. [26]; Ker [27]; Zhang and Wang [28]), it is observed that PSNR values (Table 2) of the proposed approach in the range 55.4826–57.5493 dB and embedding capacity, i.e., 0.844 BPP–0.848 BPP (Table 2), perform better than existing image steganography in terms of high embedding capacity and visual reliability of steganographic digital images, as graphically shown in Figs. 4 and 5; hence, the presented approach is robust and maintains imperceptibility even with sufficiently high embedding capacity.

Table 1 Computed PSNR values and embedding rate organized by choosing various places of focused (2×2) pixel square of cover image

S. No.	Cover image	Focused pixel square (2×2) position utilized	Embedding capacity in bits	PSNR	ER: BPPB	ER: BPP
1	lina.png	All successive blocks	2,457,600	52.2206	6.00	1.500
		Odd successive blocks	1,228,800	55.4226	3.00	0.750
		Even successive blocks	1,228,800	55.4226	3.00	0.750
		Fourth successive blocks	614,400	58.8644	1.50	0.375
		Average	1,382,400	55.4826	3.38	0.844
2	airplane.bmp	All successive blocks	2,457,600	52.4278	6.00	1.500
		Odd successive blocks	1,228,800	55.3384	3.00	0.750
		Even successive blocks	1,228,800	55.3384	3.00	0.750
		Fourth successive blocks	614,400	59.4046	1.50	0.375
		Average	1,382,400	55.6273	3.38	0.844
3	peppers.png	All successive blocks	2,457,600	53.4432	6.00	1.500
		Odd successive blocks	1,228,800	58.6434	3.00	0.750
		Even successive blocks	1,228,800	58.6434	3.00	0.750
		Fourth successive blocks	614,400	59.4672	1.50	0.375
		Average	1,382,400	57.5493	3.38	0.844
4	baboon.png	All successive blocks	2,457,600	54.4458	6.00	1.500
		Odd successive blocks	1,228,800	56.2436	3.00	0.750
		Even successive blocks	1,228,800	56.2436	3.00	0.750
		Fourth successive blocks	614,400	59.8456	1.50	0.375
		Average	1,382,400	56.69465	3.38	0.844

Table 2 Comparative analysis of proposed image steganography scheme

Cover image	Zhang and Wang [26]		Ker [27]		Chang et al. [28]		Proposed method	
(640 × 640)	PSNR (db)	ER (BPP)	PSNR (db)	ER (BPP)	PSNR (db)	ER (BPP)	PSNR (db)	ER (BPP)
lina.png	47.75	0.660	49.23	0.760	49.61	0.777	55.4826	0.844
airplane.bmp	42.31	0.683	46.17	0.770	48.23	0.787	55.6273	0.846
peppers.png	45.71	0.685	49.10	0.780	49.43	0.792	57.5493	0.848
baboon.png	46.18	0.690	47.34	0.796	47.93	0.799	56.6946	0.845

**Fig. 2** Embedding rate (BPPB) managed by positions of targeted pixel blocks

4.2.2 Anti-steganalysis System

When comparing the security of the stego image with existing image steganographic schemes, it is clear that the security of the hidden data is not given extra attention. However, our proposed approach adds an extra layer of security by using the SSS scheme with a higher embedding capacity. The SSS scheme provides a structure for the distribution of sub-image packets and reassembles them at destination accurately by the authorized recipient. Therefore, the presented security methods provide unbreakable security and work well as anti-steganalysis systems.

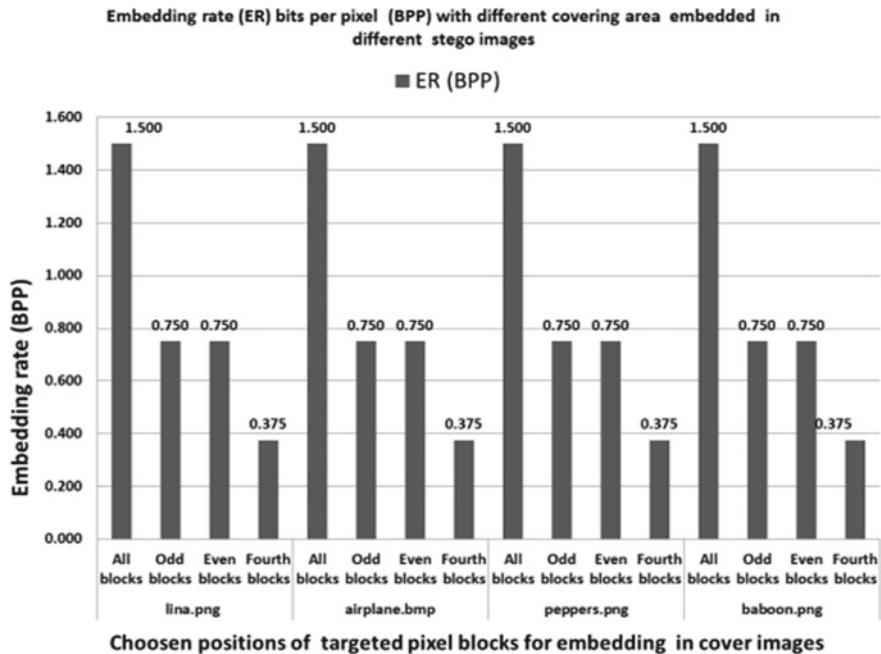


Fig. 3 Embedding rate (BPP) managed by positions of targeted pixel blocks

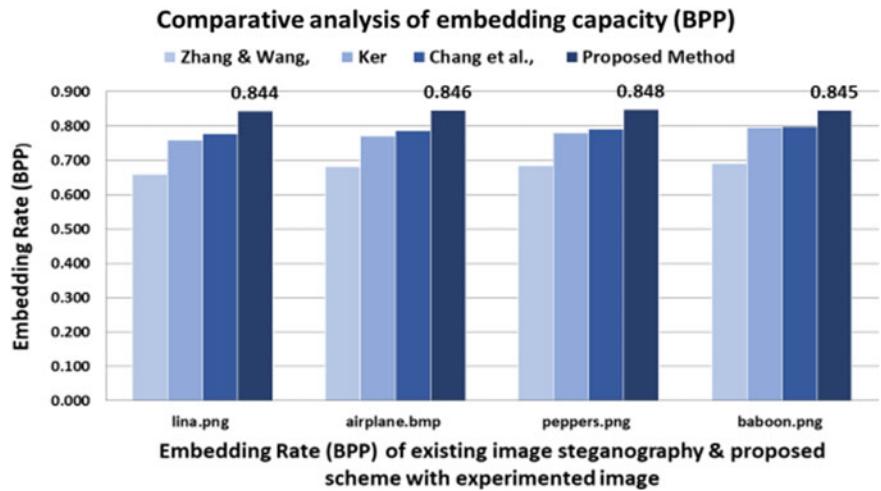


Fig. 4 Comparative analysis based on embedding capacity (BPP)

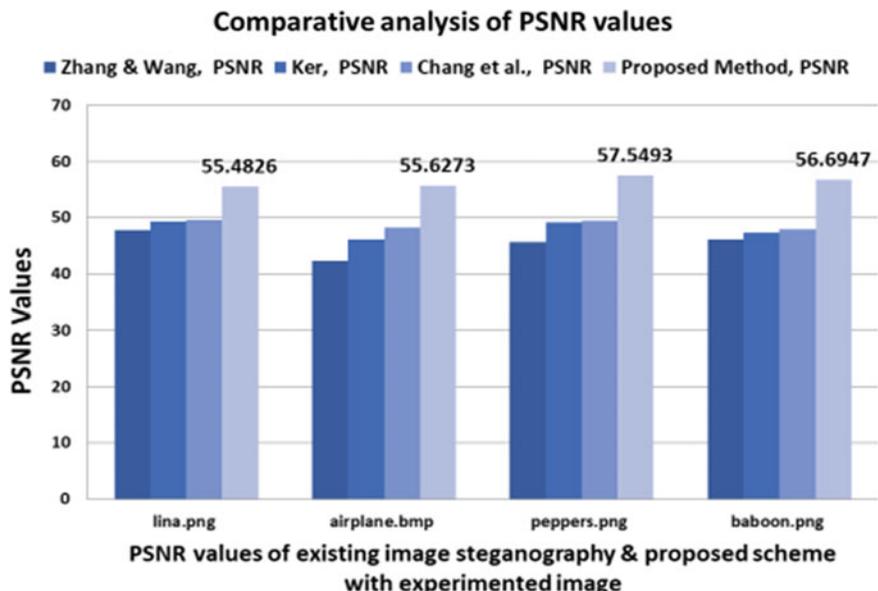


Fig. 5 Comparative analysis based on visual quality (PSNR)

5 Conclusion

The presented steganography strategy developed a strong security system associated with the self-synchronized method by utilizing Shamir's Secret Sharing Scheme. Double layers of security have been developed to ensure the highest level of safety for secret data hidden within a cover image. In the SDS layer, a single secret image of larger dimensions is divided into a number of secret images of progressively smaller dimensions, and the SSS produces a structure of self-synchronization generated for allotting and re-establishing disseminated partitioned sub-images among intended recipients. On comparing the presented scheme with prominent existing PVD image steganographic schemes, it is observed that the embedding rate of 0.844 BPP (average case of all chosen cases) is better than the various mentioned existing approaches. It is also observed that (case 1) 1.50 BPP, (case 2) 0.75 BPP, and (case 3) 0.375 BPP, varying embedding rates are achieved by choosing changeable focused locations of the (2×2) pixel square. Keeping in view the trade-off between imperceptibility, embedding capacity, and security, any case may be selected as per requirement out of the mentioned cases; hence, robust image steganography is achieved. Successful steganalysis examination of the presented steganography system using PSNR demonstrates that the introduced approach is strongly against steganalysis attack. Additionally, the introduced approach is combined with distributed computing, which improves embedding capacity, and its high security is developed by implementing our proposed Shamir's Secret Sharing Scheme. The presented steganography system

gives immensely protected reproduction of distributed confidential hidden images. Evaluation, testing, and steganalysis assessment of the proposed steganographic system all showed that it is extraordinarily secure, both in terms of high payload and sufficient high visual quality of the stego image. The proposed inventive methodology is realistic for all practical purposes.

References

1. Tyagi S, Dwivedi RK, Saxena AK (2020) A novel PDF steganography optimized using segmentation technique. *Int J Inf Tecnol* 12:1227–1235. <https://doi.org/10.1007/s41870-019-00309-7>
2. Wu D-C, Tsai W-H (2003) A steganographic method for images by pixel-value differencing. *Pattern Recogn Lett* 24(9–10):1613–1626
3. Chang C-C, Hwang R-J (2004) A new scheme to protect confidential images. *J Interconnect Netw* 5(3):221–232
4. Yang CH, Weng CY (2006) A steganographic method for digital images by multi-pixel differencing. In: Proceedings of international computer symposium, Taipei, Taiwan, pp 831–836
5. Jung K-H, Ha K-J, Yoo K-Y (2008) Image data hiding method based on multi-pixel differencing and LSB substitution methods. In: Proceedings of international conference on convergence and hybrid information technology (ICHIT 08), pp 355–358
6. Liu J-C, Shih M-H (2008) Generalizations of pixel-value differencing steganography for data hiding in images. *Fundament-Informatiae* 83(3):319–335
7. Liao X, Wen Q-Y, Zhang J (2011) A steganographic method for digital images with four-pixel differencing and modified LSB substitution. *J Vis Commun Image Represent* 22(1):1–8
8. Yang C-H, Weng C-Y, Tso H-K, Wang S-J (2011) A data hiding scheme using the varieties of pixel-value differencing in multimedia images. *J Syst Softw* 84(4):669–678
9. Liao X, Wen QY, Shi S (2011) Distributed steganography. In: Proceedings of seventh international conference on intelligent information hiding and multimedia signal processing, pp 153–156
10. Priyadarshini I, Sharma R, Bhatt D et al. (2022) Human activity recognition in cyber-physical systems using optimized machine learning techniques. *Cluster Comput*. <https://doi.org/10.1007/s10586-022-03662-8>
11. Priyadarshini I, Alkhayyat A, Obaid AJ, Sharma R (2022) Water pollution reduction for sustainable urban development using machine learning techniques. *Cities* 130:103970. ISSN 0264–2751. <https://doi.org/10.1016/j.cities.2022.103970>
12. Pandya S, Gadekallu TR, Maddikunta PKR, Sharma R (2022) A study of the impacts of air pollution on the agricultural community and yield crops (Indian Context). *Sustainability* 14:13098. <https://doi.org/10.3390-su142013098>
13. Bhola B, Kumar R, Rani P, Sharma R, Mohammed MA, Yadav K, Alotaibi SD, Alkwai LM (2022) Quality-enabled decentralized dynamic IoT platform with scalable resources integration. *IET Commun* 00:1–10. <https://doi.org/10.1049/cmu2.12514>
14. Deepanshi, Budhiraja I, Garg D, Kumar N, Sharma R (2022) A comprehensive review on variants of SARS-CoVs-2: challenges, solutions and open issues. *Comput Commun*. ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2022.10.013>
15. Ahsan Habib AKM, Hasan MK, Islam S, Sharma R, Hassan R, Nafi N, Yadav K, Alotaibi SD (2022) Energy-efficient system and charge balancing topology for electric vehicle application. *Sustain Energy Technol Assessments* 53(Part B):102516.ISSN 2213-1388. <https://doi.org/10.1016/j.seta.2022.102516>

16. Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. *Comput Electri Eng* 105:108543. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2022.108543>
17. Sharma R, Rawat DB, Nayak A, Peng S-L, Xin Q (2023) Introduction to the special section on survivability analysis of wireless networks with performance evaluation (VSI-networks survivability). *Comput Netw* 220:109498. ISSN 1389-1286. <https://doi.org/10.1016/j.comnet.2022.109498>
18. Ghildiyal Y, Singh R, Alkhayyat A, Gehlot A, Malik P, Sharma R, Vaseem Akram S, Alkwai LM (2023) An imperative role of 6G communication with perspective of industry 4.0: challenges and research directions. *Sustain Energy Technol Assessments* 56:103047. ISSN 2213-1388. <https://doi.org/10.1016/j.seta.2023.103047>
19. Ahasan Habib AKM, Hasan MK, Alkhayyat A, Islam S, Sharma R, Alkwai LM (2023) False data injection attack in smart grid cyber physical system: issues, challenges, and future direction. *Comput Electri Eng* 107:108638. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108638>
20. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Alkwai LM, Kumar S (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electri Eng* 107:108617. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
21. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerg Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>
22. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electri Eng* 108:108715. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
23. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun.* ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
24. Tyagi S, Dwivedi RK, Saxena AK (2019) High capacity steganography pro-tected using Shamir's threshold scheme and permutation framework. *Int J Inn Technol Expl Eng* 8(9S):784–795
25. Petitcolas P (1997–2020) The information hiding homepage. https://www.petitcolas.net/watermarking/image_database/
26. Chang K-C, Huang PS, Tu T-M, Chang C-P (2007) Adaptive image steganographic scheme based on tri-way pixel-value differencing. In: Proceedings of the IEEE interna-tional conference on systems, man, and cybernetics (SMC 07), pp 1165–1170
27. Ker A (2005) Improved detection of LSB steganography in gray-scale image. In: Lecture notes in computer science, pp 97–115
28. Zhang X, Wang S (2004) Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security. *Pattern Recogn Lett* 25(3):331–339

Design and Comparison of Various Parameters of T-Shaped TFET of Variable Gate Lengths and Materials



Jyoti Upadhyay, Tarun Chaudhary, Ramesh Kumar Sunkaria, and Mandeep Singh

Abstract T-Shaped Dual-Gate TFET channel provides larger tunneling junction area which offers more electrons, which enhances I_{ON} current. This paper will provide a description and see different variations by using substrate material like Silicon, Gallium Arsenide, Germanium at different channel lengths (46 nm, 36 nm, 26 nm, 16 nm, 10 nm). For insulating material, we have taken Hafnium oxides (HfO_2) and Silicon-oxide. We have analyzed and compared crucial characteristics such as the I_{ON}/I_{OFF} ratio, the subthreshold swing (SS_{avg}), transconductance, BTBT, electric field (EF), and surface potential. Germanium material which shows better I_{ON}/I_{OFF} ratio and better subthreshold swing, Sensitivity than Silicon and Gallium Arsenide. Electric field and Surface Potential of device getting improved as we decrease the channel width.

Keywords DG-TFET Dual-Gate-T-shaped TFET · BTBT band-to-band tunneling · Hafnium oxides (HfO_2) · Trap-assisted tunneling (TAT) · Gallium Arsenide (GaAs)

1 Introduction

Leading to deep Nano-CMOS technology, the demand for higher packing density, higher velocity, a smaller footprint, and lower power consumption, including scaling of MOS devices, is getting deeper. Low-voltage circuit design methods have grown to be a fascinating topic of study for the circuit design community [6]. As the size

J. Upadhyay (✉) · T. Chaudhary · R. K. Sunkaria · M. Singh
Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, India
e-mail: jyoti1997u@gmail.com

T. Chaudhary
e-mail: chaudharyt@nitj.ac.in

R. K. Sunkaria
e-mail: sunkariark@nitj.ac.in

of MOSFETs shrinks into the nanoscale region, scaling of MOS devices becomes more difficult [2]. It has problem of short-channel effects (SCEs). TFETs, unlike MOSFETs, allow charge to flow in both the reverse and forward gate bias states [2]. The current in MOSFETs is primarily due to the thermionic way of releasing free charge carriers [2].

The limitation of MOSFET is overcome by TFET by manipulation of band-to-band tunneling (BTBT) [1], for which MOSFET carrier transport provides benefits over thermionic injection [1]. A TFET is excellent for use in low-power applications since its subthreshold swing is better [1].

The TFET is excellent for high-frequency applications as well as analogue and RF applications since it is resistant to the short-channel effects [2, 3]. It exhibits a quite low leakage current measured in femto amperes (fA) [1]. Here, we have design a T-Shaped Dual-Gate TFET with different materials and insulating materials at different channel lengths which vary from 46 nm, 36 nm, 26 nm, 16 nm, 10 nm, observing which TFET structure will give an optimized and better I_{ON}/I_{OFF} . Parallelly, we observe subthreshold slope, Electric Potential, band-to-band tunneling, and transconductance of structure. Different structures are proposed to improve performance of TFET. This demonstrates how T-TFETs can be used in both analogue and digital applications [6]. Concurrent Si-TFETs, however, suffer from a number of shortcomings including low I_{ON} , big SS_{avg} , tiny I_{ON}/I_{OFF} , high V_T , the presence of ambipolar currents, and drain-induced current augmentation because of huge Silicon substrate bandgap [2]. To use these in digital applications, the reverse voltage at gate, or the ambipolar current, must be lowered [2]. Because of their difficulty in logic applications due to their ambipolar behavior at increased V_{DS} bias, TFETs require more complex circuit designs than CMOS [5].

Using a weakly doped drain is a natural way to control ambipolar behavior [5]. Because of its better subthreshold swing and bigger tunnel junction area, the usage of T-shaped channels has been shown to improve overall TFET performance when compared to standard shaped channels in experiments. The usage of T-shaped channels has been shown to improve overall TFET performance when compared to standard shaped channels in experiments [2]. Subthreshold design techniques, deep subthreshold design techniques, with other low-voltage circuit design methods have all received a great deal of attention [6].

However, with an exponential sensitivity to voltage and temperature in these areas, there are significant difficulties. Identifying device architectures with steep slopes as a good contender [6]. Some difficulties, such as low on-current (I_{on}), continue to restrict the application of TFETs [5], Trap-aided tunneling (TAT) [5] led to SS deterioration, and the PIN structure produced ambipolar switching behavior [5]. For low-power circuits, a significant alternative to metal–oxide–semiconductor FET technology has arisen in the form of the steep-slope tunneling field-effect transistor (TFET) [6]. TFETs have issues because of their low ON current (I_{ON}), ambipolar behavior, and higher Miller capacitances. [6].

As a result, TFETs have been seen as a possible alternative to MOSFETs since they can achieve minimum 60 mV/decade subthreshold swing (SS_{avg}) at ambient temperature and have a better I_{ON}/I_{OFF} ratio due to their low-voltage operating capability [2].

T-shaped channel allows for more control of gate potential between any two gates, thereby increasing I_{ON} current. BTBT, or band-to-band tunneling, is the method used by TFET [3] because of the TFET's low leakage current, low subthreshold slope, high sensitivity, and feasibility for low-power applications. It has been used in a number of implementations [3]. There have been reports that TFETs' ON current can be increased by incorporating high-k dielectric and hetero-gate dielectric [3]. Ambipolarity in a TFET device is observed to be decreased by the presence of a drain pocket (DP) and a gate-drain underlap area [3]. Device design engineering and bandgap engineering, which involves including a heterojunction across source-channel interface, are said to boost the device's sensitivity by raising the likelihood of tunneling and consequently the rate at which BTBT is generated [3]. Designed structure leads to better functionality, including an increment in the I_{ON}/I_{OFF} current ratio along with reduction in the subthreshold slope. Ambipolarity is reduced by optimizing DP doping and length [3]. This paper includes a thorough investigation of device performance metrics, including I_{ON}/I_{OFF} current ratio, subthreshold slope, threshold voltage, and drain current sensitivity of device [3].

Even leakage current is in almost femto amperes allowing it to be used in high-frequency applications [2]. For applications requiring low power and low voltage, it has been found that T-TFET circuits outperform their FinFET equivalents [6]. Generally, TFETs are known for their low off currents, which contribute to their low-power consumption. However, achieving high on currents has been a challenge in TFET technology. While TFETs have demonstrated promising results in terms of reducing power consumption, further research and development are still required to improve their on-current performance to make them competitive with conventional transistor technologies like MOSFETs.

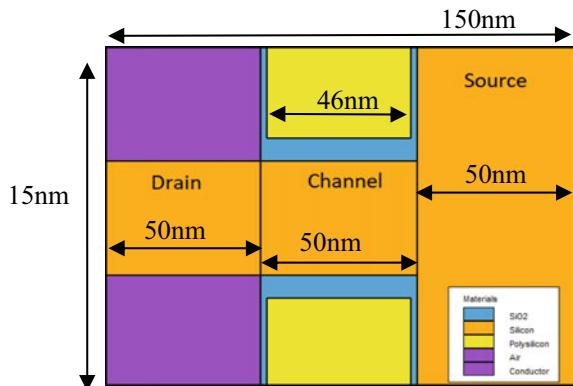
The rest of this essay is structured as follows: The device structure, parameter properties, and usage are described in Sect. 2 [6]. Section 3 compares circuits based on their performance metrics. The results and performance are included in Sect. 4. Finally, Sect. 5 brings this paper to a close.

2 Device Structure and Simulation

The structures are given names based on the materials they are made of: Gallium Arsenide-based T-Shaped—Dual-Gate TFET (GaAs-TC-DG-TFET) and Silicon-based T-Shaped—Dual-Gate TFET (Si-TC-DG-TFET), where Silicon is used to combine the source, channel, and drain regions. We use double gates instead of single-gate TFET as it gives extra control ability over the channel.

Figure 1 is designed structure defined with Silicon substrate with 46 nm channel length. Similar kind, we have designed T-shaped TFET with Silicon, Germanium, Gallium Arsenide with 46 nm, 36 nm, 26 nm, 16 nm, 10 nm. We have designed device structure of TFET in Tshapd with overall structure length 150 nm and wide 15 nm (Table 1).

Fig. 1 Device structure of 46 nm channel length on Silicon substrate



3 Comparative Analysis on Devices and Discussion

3.1 *ON Current and OFF Current*

TFET is a type of transistor that operates based on quantum tunneling principles. It is intended to provide a high I_{on}/I_{off} current ratio with low threshold voltage making it suited for low-power applications. The TFET has an inherent advantage over conventional transistors in terms of reducing power consumption. The current that flows between both the drain and the source of a TFET when it is in the conducting or “on” condition is referred to as the “on current.” The transistor’s “off” condition, or non-conducting condition, is referred to as the “off current” (Table 2).

3.2 *Subthreshold Swing (SS_{avg})*

In the subthreshold region, it is the rate at which the voltage at the gate varies in relation to the drain current. It is an important parameter because it directly impacts the energy efficiency and switching speed of the transistor. TFETs exploit quantum tunneling through a thin barrier, allowing for steeper subthreshold swing values. Various factors, such as the device design, material properties, interface effects, and process variations, can influence the subthreshold swing (Table 3).

3.3 *Transconductance (g_m)*

Transconductance is related to V_{GS} and I_D of a transistor. It indicates how effectively the transistor can control the current flow through it. In the case of a tunnel field-effect

Table 1 Device specification parameters

Device specification parameter	Abbreviation	Silicon (<i>Si</i>)	Germanium (<i>Ge</i>)	Gallium Arsenide (GaAs)
Source (Dopant Type)	N_S	$1 \times 10^{22} \text{ cm}^{-3}$ (P-type)	$1 \times 10^{20} \text{ cm}^{-3}$	$1 \times 10^{20} \text{ cm}^{-3}$
Drain (Dopant Type)	N_D	$1 \times 10^{22} \text{ cm}^{-3}$ (N-type)	$1 \times 10^{18} \text{ cm}^{-3}$	$1 \times 10^{17} \text{ cm}^{-3}$
Channel (Dopant Type)	N_C	$1 \times 10^{16} \text{ cm}^{-3}$ (P-type)	$1 \times 10^{20} \text{ cm}^{-3}$	$1 \times 10^{20} \text{ cm}^{-3}$
Source Length	L_S	50 nm	50 nm	50 nm
Drain length	L_D	50 nm	50 nm	50 nm
Channel length	L_C	46 nm, 36 nm, 26 nm, 16 nm, 10 nm	46 nm, 36 nm, 26 nm, 16 nm, 10 nm	46 nm, 36 nm, 26 nm, 16 nm, 10 nm
Oxide length	L_O	50 nm, 40 nm, 30 nm, 20 nm, 14 nm	50 nm, 40 nm, 30 nm, 20 nm, 14 nm	50 nm, 40 nm, 30 nm, 20 nm, 14 nm
Source Thickness	T_S	15 nm	15 nm	15 nm
Drain Thickness	T_D	5 nm	5 nm	5 nm
Channel Thickness	T_C	5 nm	5 nm	5 nm
Oxide Thickness	t_{ox}	1 nm	1 nm	1 nm
Work Function	W_{FG}	3.8 eV and 4.0 eV	3.8 eV and 4.0 eV	3.8 eV and 4.0 eV
Total Thickness of Device	L_{total}	15 nm	15 nm	15 nm
Total length of Device	T_{total}	150 nm	150 nm	150 nm

transistor (TFET), the transconductance is an important parameter that characterizes its amplification capability (Table 4).

4 Results

Using Silvaco TCAD tool, we designed design and compared the devices on different aspects which are major contributing factor in performance of device which include the transfer characteristics, the transconductance, the band-to-band to tunneling, the electric field, the Electric Potential.

Table 2 Comparison of I_{ON} and I_{OFF} in different materials and channel lengths

Substrate	Silicon			Germanium			Gallium Arsenide		
Channel length	I_{ON}	I_{OFF}	I_{ON}/I_{OFF}	I_{ON}	I_{OFF}	I_{ON}/I_{OFF}	I_{ON}	I_{OFF}	I_{ON}/I_{OFF}
46 nm	8.00E-06	1.51E-18	5.28E+12	2.25E-04	1.63E-15	1.38E+11	2.93E-05	5.59E-18	5.25E+12
36 nm	2.00E-06	2.03E-17	9.86E+10	7.03E-05	1.78E-15	3.94E+10	1.04E-06	1.05E-17	9.91E+10
26 nm	3.82E-06	1.36E-18	2.82E+12	1.20E-04	1.76E-15	6.82E+10	1.09E-06	7.52E-18	1.45E+11
16 nm	2.27E-06	1.82E-15	1.24E+09	7.71E-05	3.89E-15	1.98E+10	1.01E-06	7.60E-18	1.33E+11
10 nm	1.14E-06	1.05E-11	1.00E+05	5.40E-05	1.86E-10	2.89E+05	1.07E-06	2.83E-14	3.80E+07

Table 3 Subthreshold swing on different materials and channel lengths

Channel length of device (nm)	Subthreshold swing (SS_{avg}) (mV/decade)		
	Silicon	Germanium	Gallium Arsenide
46	28.1	33.9	23.05
36	45.6	33.2	35.15
26	28.5	35.3	49.16
16	64.8	50	76.29
10	159.5	140.7	156.5

Table 4 Transconductance on different materials and channel lengths

Channel length of device	Transconductance (g_m) siemens (μS)		
	Silicon	Germanium	Gallium Arsenide
46 nm	26.3	69.65	67.2
36 nm	12.6	84.79	3.2
26 nm	20.2	79.65	3.44
16 nm	18.1	74.26	2.91
10 nm	6.9	72.35	3.29

4.1 Transfer Characteristics of Device

In below layout of transfer characteristics, as we are decreasing the channel length of device, our current strength is getting improved, but on same side the subthreshold swing is getting poor and switching performance decreases as I_{ON}/I_{OFF} decreases. Germanium substrate device shows better current on compared to Silicon and GaAs substrate and better switching performance (Fig. 2).

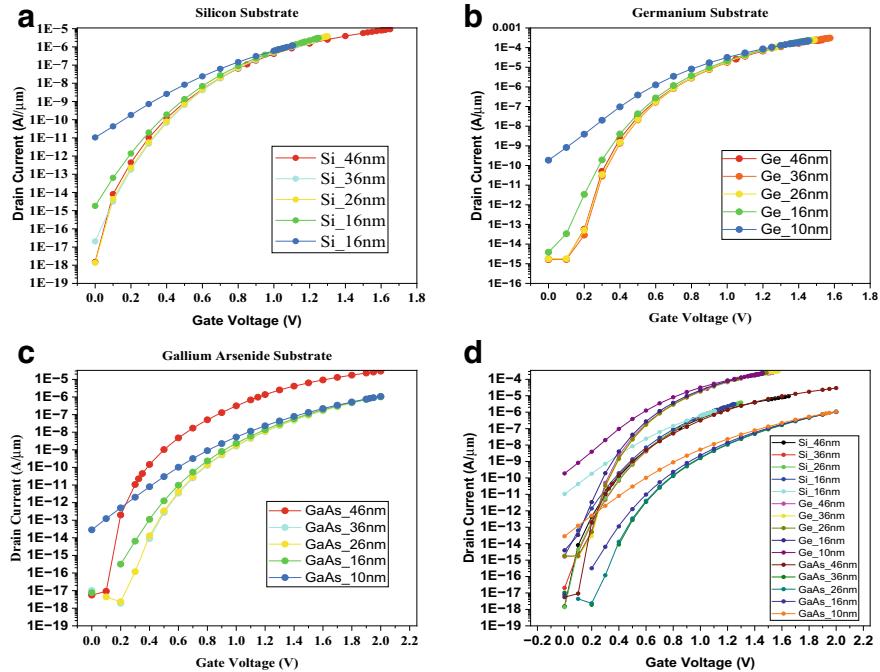


Fig. 2 Transfer characteristics (I_D vs V_{GS}) for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined transfer characteristics for all materials

4.2 Transconductance Analysis (g_m)

With increasing channel of device, the I_{ON}/I_{OFF} ratio of device is decreased, and subthreshold swing decreases. Germanium substrate shows better switching capability than Silicon and GaAs. So, Germanium has better sensitivity in regards with other material. As we decrease channel length, the transconductance of devices performance is degraded (Fig. 3).

4.3 Band-to-Band Tunneling

The phenomenon of BTBT, where charge carriers' tunnel through a thin barrier across source, drain of the transistor. This tunneling process allows the TFET to achieve a significant reduction in the subthreshold swing, enabling it to operate at low voltages. TFETs are designed with a narrow energy bandgap material in the channel region, typically a heavily doped semiconductor or a heterojunction interface. This narrow bandgap enables efficient band-to-band tunneling. The steep subthreshold swing is

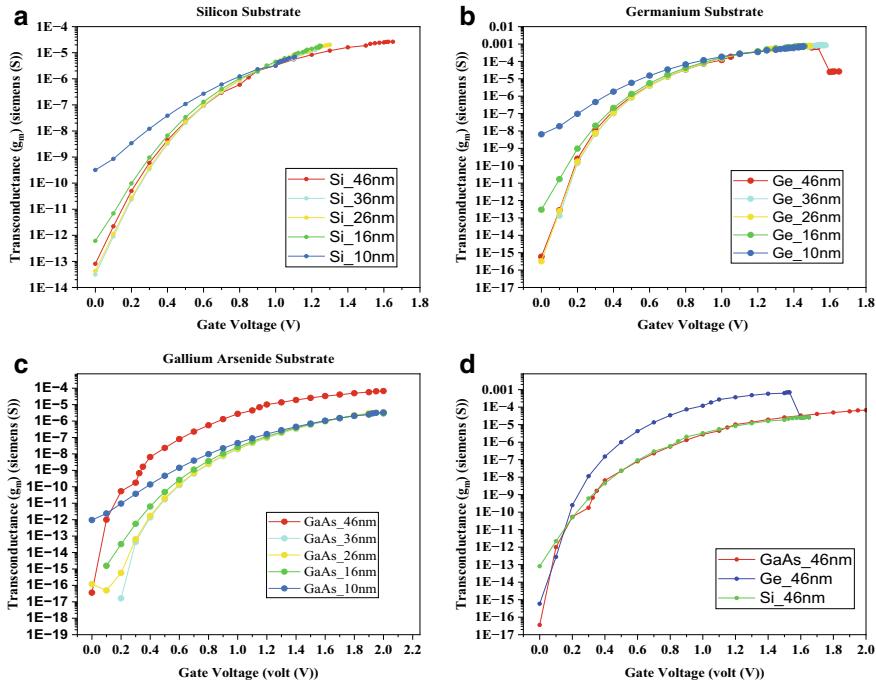


Fig. 3 Transconductance versus gate voltage for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined transconductance all materials at 46 nm

one important benefit. Band-to-band tunneling allows TFETs to maintain a high I_{ON}/I_{OFF} current ratio, which means that the device can effectively turn on and off the current flow, even at low supply voltages. This characteristic is crucial for achieving low-power operation and high-energy efficiency. On comparing BTBT effect, GaAs shows better tunneling compared to all substrates, and while increasing the channel length across all devices, the tunneling is getting improved. With 10 nm GaAs device, tunneling is seen better. Tunneling gets improved as we decrease the device channel length (Fig. 4).

4.4 Electric Field

The electric field is responsible for shaping the energy profile within the TFET device, enabling efficient tunneling across the narrow bandgap region (Fig. 5).

As moving from source toward the drain region, the strength of the electric field increases, which leads to creating a high electric field zone close to the source-channel junction. The electric field acts as a driving force for charge carriers, enabling

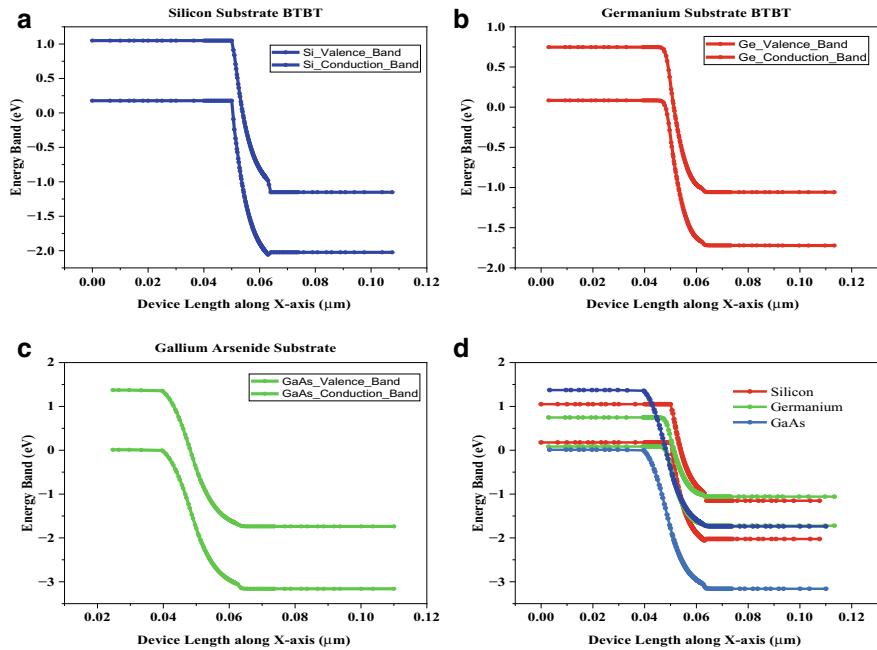


Fig. 4 Band energy across device length for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined band energy at 10 nm all materials

them to overcome the potential barrier and tunnel through the energy bandgap. The high electric field region across the source–channel junction reduces the effective barrier height, facilitating the band-to-band tunneling process. This allows charge carriers, either electrons or holes, to tunnel through the narrow bandgap material and flow from the source toward the drain. The electric field profile in a TFET can be influenced by various factors, including the device structure, doping profile, and applied bias. Optimizing the electric field distribution is important for achieving efficient tunneling and controlling the transistor's performance. GaAs shows better EF than Germanium and Silicon. The strength of EF increases as we increase the device channel length, which increases flow of charge carries and strengthen Current.

4.5 Surface Potential

It refers to the potential energy of charge carriers at the surface of the device. It is essential in controlling the band alignment and carrier transport in TFETs. In TFETs, the surface potential is particularly important in the channel region where band-to-band tunneling occurs.

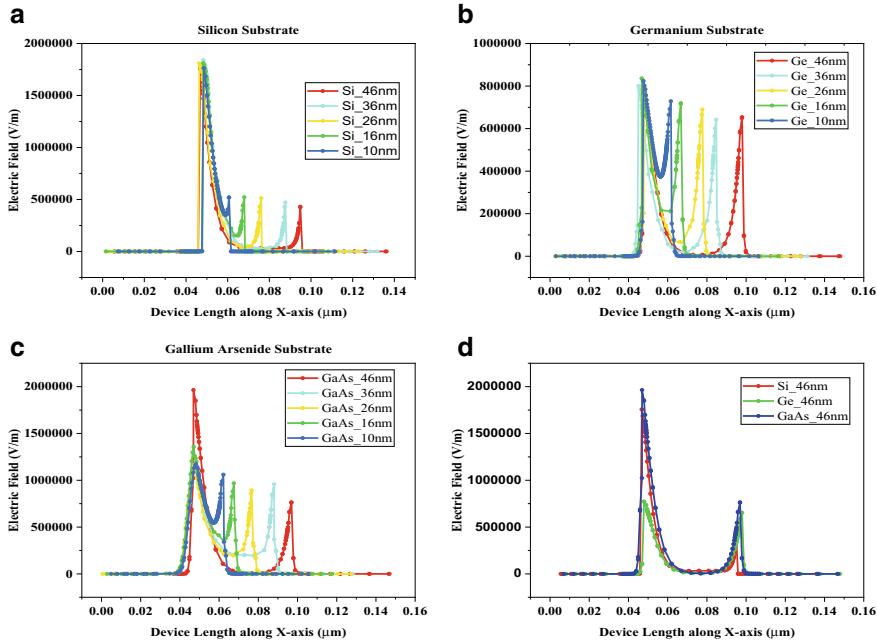


Fig. 5 Electric field for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined EF of 46 nm all materials

The surface potential influences the energy profile within the device, affecting the barrier height and the probability of tunneling through the narrow bandgap material. The voltage at gate controls the surface potential by modulating the charge distribution and the energy bands in the channel region.

Increase in V_{GS} above V_T , the surface potential is lowered, reducing the barrier height. This enables the tunneling of charge carriers from the source to the drain, leading to an increased ON-state current. Optimizing the surface potential in TFETs is critical for achieving desirable performance characteristics. Figure 6 shows surface potential layout, when device is in OFF state, and Fig. 7 shows surface potential layout, when device is in ON state. GaAs substrate shows more surface potential, which make band energy difference small, and provide better tunneling which enhances device performance.

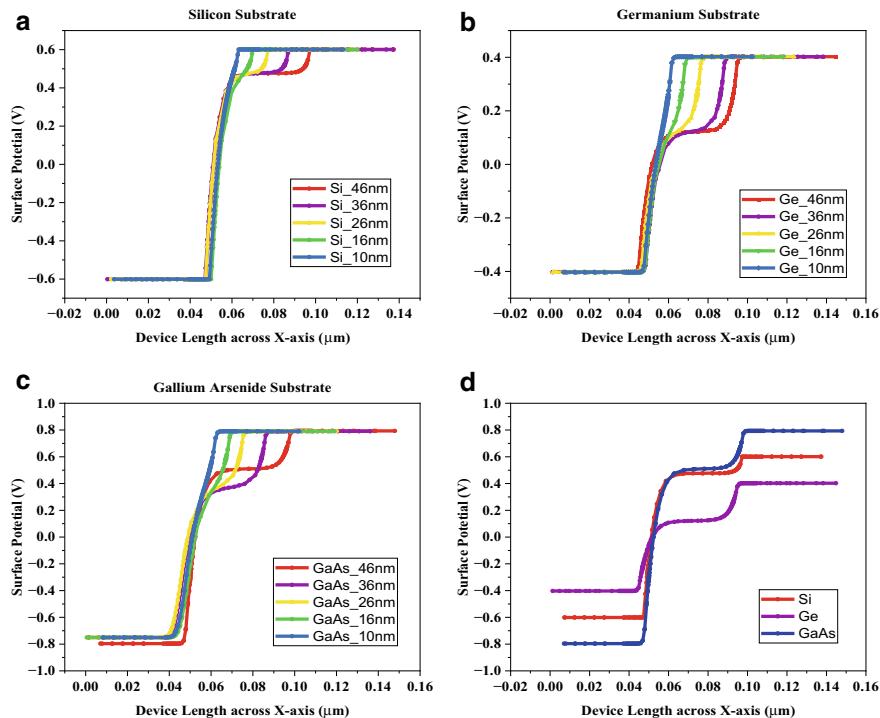


Fig. 6 Surface potential across device length for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm in OFF state, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined surface potential of 46 nm all materials

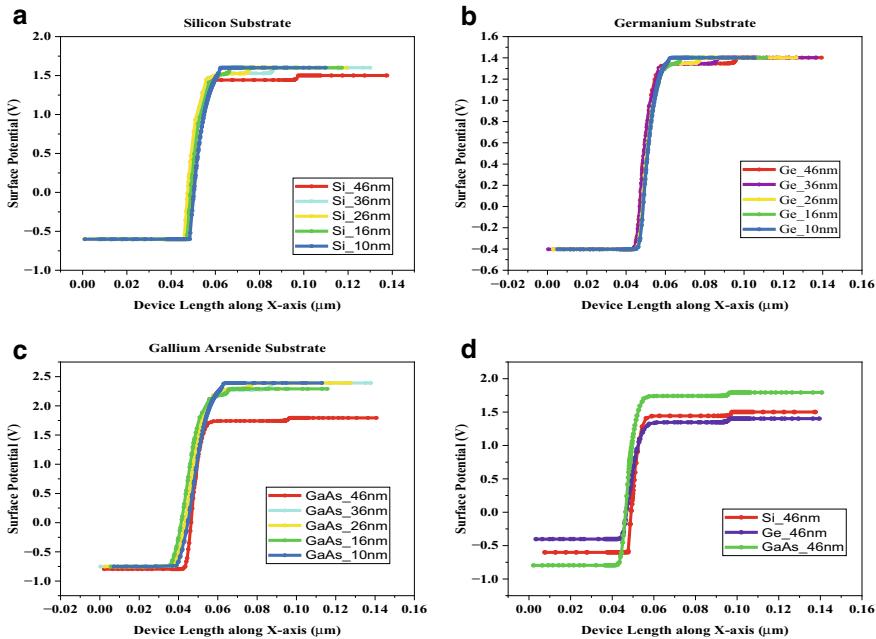


Fig. 7 Surface potential across device length for 46 nm, 36 nm, 26 nm, 16 nm, 10 nm in ON state, **a** Silicon substrate, **b** Germanium substrate, **c** Gallium Arsenide substrate comparison, **d** Combined BTBT of 46 nm of all materials

5 Conclusion

Using different materials and varying channel lengths of T-Shaped-DG-TFET, parameters like I_{ON} and I_{OFF} , transconductance, BTBT of device, electric field, surface potential at ON and OFF state have been investigated. While decreasing the channel length, different materials support and enhance particular parameter. While decreasing channel length, Germanium and Silicon shows best outcome, where in Band to Band is better in GaAs. But as we decrease width more, the performance of device is getting degraded. Subthreshold and switching of device got degraded. To overcome this, for the source and drain, many materials may be utilized of different widths which may optimize the parameter if device has a broad bandgap material in the drain area and a narrow bandgap material near the source. Complexity of the fabrication process to achieve a high-performance TFET device, this strategy may successfully suppress the ambipolar conduction effect. We can subsequently improve it by optimizing the process variation parameter.

References

1. Bijoy G, Bhattacharjee D, Dash DK, Bhattacharya A, Sarkar SK (2018) Demonstration of T-shaped channel tunnel field-effect transistors. In: 2018 2nd international conference on electronics, materials engineering and nano-technology (IEMENTech), IEEE, pp 1–5
2. Yadav R, Singh M, Chaudhary T (2022) Comparative analysis of T-shaped Tfet structures for low power applications. *Int J Electr Electron Data Commun (IJEEDC)* 10(7):37–41
3. Shaw N, Mukhopadhyay B (2022) Modeling and performance analysis of a split-gate T-shape channel DM DPDG-TFET label-free biosensor. *IEEE Sens J* 23(2):1206–1213
4. Ye H, Hu J (2020) A new type of N-type TFET with tri-input terminals using T-shaped structure. In: 2020 IEEE 20th international conference on nanotechnology (IEEE-NANO), IEEE, pp 124–127
5. Liu C, Ren Q, Chen Z, Zhao L, Liu C, Liu Q, Wenjie Y, Liu X, Zhao Q-T (2019) A T-shaped SOI tunneling field-effect transistor with novel operation modes. *IEEE J Electron Devices Soc* 7:1114–1118
6. Dubey PK, Kaushik BK (2020) Evaluation of circuit performance of T-shaped tunnel FET. *IET Circuits Devices Syst* 14(5):667–673
7. Chong C, Liu H, Wang S, Chen S, Xie H (2021) Study on single event effect simulation in T-shaped gate tunneling field-effect transistors. *Micromachines* 12(6):609
8. Mittal M, Khosla M, Chawla T (2022) Design and performance analysis of delta-doped hetro-dielectric GeOI vertical TFET. *SILICON* 14(10):5503–5511
9. Prabhat S, Prakash Samajdar D, Yadav DS (2021) A low power single gate l-shaped tfet for high frequency application. In: 2021 6th international conference for convergence in technology (i2CT), IEEE, pp 1–6
10. Priyadarshini I, Sharma R, Bhatt D et al (2022) Human activity recognition in cyber-physical systems using optimized machine learning techniques. *Cluster Comput.* <https://doi.org/10.1007/s10586-022-03662-8>
11. Priyadarshini I, Alkhayyat A, Obaid AJ, Sharma R (2022) Water pollution reduction for sustainable urban development using machine learning techniques. *Cities* 130:103970. ISSN 0264–2751. <https://doi.org/10.1016/j.cities.2022.103970>
12. Pandya S, Gadekallu TR, Maddikunta PKR, Sharma R (2022) A study of the impacts of air pollution on the agricultural community and yield crops (Indian Context). *Sustainability* 14:13098. <https://doi.org/10.3390/su142013098>
13. Bhola B, Kumar R, Rani P, Sharma R, Mohammed MA, Yadav K, Alotaibi SD, Alkwai LM (2022) Quality-enabled decentralized dynamic IoT platform with scalable resources integration. *IET Commun* 00:1–10. <https://doi.org/10.1049/cmu2.12514>
14. Deepanshi, Budhiraja I, Garg D, Kumar N, Sharma R (2022) A comprehensive review on variants of SARS-CoVs-2: challenges, solutions and open issues. *Comput Commun* ISSN 0140–3664. <https://doi.org/10.1016/j.comcom.2022.10.013>
15. Ahsan Habib AKM, Hasan MK, Islam S, Sharma R, Hassan R, Nafi N, Yadav K, Alotaibi SD (2022) Energy-efficient system and charge balancing topology for electric vehicle application. *Sustain Energy Technol Assessments* 53(Part B):102516. ISSN 2213–1388. <https://doi.org/10.1016/j.seta.2022.102516>
16. Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. *Comput Electri Eng* 105:108543. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2022.108543>
17. Sharma R, Rawat DB, Nayak A, Peng S-L, Xin Q (2023) Introduction to the special section on survivability analysis of wireless networks with performance evaluation (VSI–networks survivability). *Comput Netw* 220:109498. ISSN 1389–1286. <https://doi.org/10.1016/j.comnet.2022.109498>
18. Ghildiyal Y, Singh R, Alkhayyat A, Gehlot A, Malik P, Sharma R, Vaseem Akram S, Alkwai LM (2023) An imperative role of 6G communication with perspective of industry 4.0: challenges and research directions. *Sustain Energy Technol Assessments* 56:103047. ISSN 2213–1388. <https://doi.org/10.1016/j.seta.2023.103047>

19. Ahsan Habib AKM, Hasan MK, Alkhayyat A, Islam S, Sharma R, Alkwai LM (2023) False data injection attack in smart grid cyber physical system: issues, challenges, and future direction. *Comput Electri Eng* 107:108638. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108638>
20. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Alkwai LM, Kumar S (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electri Eng* 107:108617. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
21. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerging Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>
22. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electri Eng* 108:108715. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
23. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun* ISSN 0140–3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
24. Paul R (2022) Performance investigation and optimization of 2-D material based double gate tunneling field-effect transistor (DG-TFET). In: 2022 International conference on advancement in electrical and electronic engineering (ICAEEE), IEEE, pp 1–4
25. Wang H, Han G, Liu Y, Zhang J, Hao Y, Jiang X (2017) The performance improvement in SiGeSn/GeSn p-channel hetero Line Tunneling FET (HL-TFET). In: 2017 IEEE electrical design of advanced packaging and systems symposium (EDAPS), IEEE, pp 1–3
26. Lin H-H, Hu VP-H (2018) Device design of vertical nanowire III-V heterojunction TFETs for performance enhancement. In: 2018 7th international symposium on next generation electronics (ISNE), IEEE, pp 1–4

Experiment to Find Out Suitable Machine Learning Algorithm for Enzyme Subclass Classification



Amitav Saran, Partha Sarathi Ghosh, Umasankar Das,
and Thiagarajan Chenga Kalvinathan

Abstract Proteins play a major role in determining many characteristics and functions of living beings. Prediction of protein classes and subclasses is one of the prominent topics of research in bioinformatics. Machine learning methods are widely used for prediction purposes, also applied for classification and subclassification of proteins. The problem is to classify the proteins to the corresponding subclass they belong to and choose a suitable machine learning method which can be used for better subclass classification. The objective is to compare the performances of three existing machine learning methods: logistic regression, support vector machine (SVM), and random forest, for protein subclassification. For this study the methods are implemented, and their results are compared by varying the number of samples of different subclasses and varying the number of subclasses. Logistic regression and support vector machine are used as a binary classifier for predicting multiple classes with $\log_2(n)$ number of classifiers for n class labels. It is observed that both random forest and support vector machine provide almost same accuracy for smaller data size, but as the data size increases random forest performs better than SVM.

Keywords Protein · Subclasses · Support vector machine (SVM) · Random forest (RF) · Logistic regression (LR) · Binary classes

Supported by organization x.

A. Saran (✉) · P. S. Ghosh · U. Das · T. C. Kalvinathan
LS Discovery, Cognizant, Chennai, India
e-mail: amitav.saran@cognizant.com

P. S. Ghosh
e-mail: parthasarathi2.ghosh2@cognizant.com

U. Das
e-mail: umasankar.das@cognizant.com

T. C. Kalvinathan
e-mail: thiyagarajan.kalvinathan@cognizant.com

1 Introduction

Proteins are responsible for determining the required functions like repair, growth, or reduction of a cell. Prediction of protein function is indispensable for designing drugs and predicting many potential reasons and solutions for living being problems. Classification of proteins as a subject of research has its own significance in bioinformatics. A newly identified protein needs to be categorized for the behavior or function it is responsible for. Proteins feature values are available in expasy.org website. One of the ways of determining the new protein functionality is through sequence matching by using BLAST [1] or FASTA [2]. But highly functional matching proteins may not have higher alignment score [3] when they are tested for matching. Sequence alignment of the secondary structure for protein folding check [4] is another way for function characterization of a protein.

We choose the basic properties of a protein from the expasy.org for classifying the proteins. A protein is identified through one unique number, and for each protein the corresponding attribute values can be found in the expasy.org. Mostly the behavior and function of a protein are linked with the values of the attributes. For prediction of a new protein different methods are used, and popularly various machine learning methods are used like SVM [5], random forest [6], neural network deep learning methods like CNN [7], RNN, LSTM [8], etc. Classification of proteins into available classes is done by researchers with high accuracy. But classification accuracy of proteins to subclass levels is not yet satisfactory. Most of the study are considering same no and less number of proteins for prediction purposes. Different researchers used different methods for showing their results along with the comparison with others.

We try to study the outcome by using variant number of protein datasets with different number of classes. The subclasses 1, 4, 5, 6, 7, and 8 of class 1, subclasses 1, 3, 4, and 8 of class 2, and subclass 1 from class 3 and class 4 are considered to build models. For prediction we have used logistic regression [9], SVM classifiers, and random forest for multiclass classification, and detailed method description is given in Sect. 3. We first try with the well-known simple method logistic regression, as its already stated in the book [10] that for multiple classes it will not give more accurate prediction, and we propose to build the multiclass logistic regression classifier by using $\log_2(n)$ number of binary logistic regression classifiers. The detailed algorithm is given in fourth section. Many researchers tried with SVM and show the improved results with less number of fixed size protein samples. This motivates us to test by taking increasing size of samples for train and test and to see the impact when the data size and class size increase.

One important thing is to use SVM as a binary classifier but should predict multiple classes, to achieve that we have used $\log_2(n)$ number of SVM classifiers in the same way as in case of logistic regression. In testing random forest model comes out as a successful method providing consistent good accuracy over different size filtered datasets with varying number of subclasses. The detailed comparison is given in Sect. 6. This paper consists of 8 sections starting with Sect. 1 as introduction. In

Sect. 2 the prescription gist from others work is presented in nutshell. Section 3 contains the brief description of the existing methods used for the study. Preparation of data for study is present in Sect. 4. Section 5 contains the detail algorithm to use binary classifiers for multiclass prediction. Section 6 presents results and discussion. The paper concludes with conclusion and future work along with reference in Sects. 7 and 8.

2 Background

To predict the class a protein belongs to is studied in various ways, and researchers represent results of their study in the respective publications. In this section some studies are reflected in brief which motivated us to carry out the comparative study over varying size of samples and subclasses. Marcotte et al. [11] used domain fusion analysis with the concept that interacting proteins may fuse to formulate another one having the same functions. The proteins with similar functions interact with each other. Gene cluster idea by Overbeek et al. [12] is another approach as the genes with same functions form a cluster. Protein subcellular location by Cai et al. [13] formulates protein groups using Kohonen's self-organized neural network. Structural feature similarity by Stawiski et al. [14] is taken into consideration with an example of structural similarity on O-glycosidase function. Dobson et al. [15] try by taking 1178 proteins and representing them by secondary structure, amino acid propensities, content and surface properties, and ligands determining enzyme or not then predicting them using SVM by reducing the 52 features to 36 by applying adoptive search for feature selection. A 3-step predictor is designed by Shen et al. EzyPred [16] for subclassification after testing enzyme or not then the class label.

Both FunD and pseudo-position-specific scoring matrix (Pse-PSSM) approach put together to develop the model. After sequence comparison is done through RPS-BLAST, KNN classifier is used to predict protein subclasses from classes after filtering enzymes. Debasmita et al. [17] devised one optimal SVM classifier for predicting subclasses using $\log_2(n)$ classifiers for n subclasses for 63 classes; they used 7 classifiers: first one to predict enzyme then use the respective bit values in its position to categorize the subclasses. They also used orthogonal subset selection to reduce the features from 32 to 20 numbers. Modified teaching learning-based optimization (MTLBO) is used to tune the hyperparameters for RBF kernel used in SVM. They have trained and tested with 3336 enzymes with a varying accuracy percentage from 60 to 98.25.

This motivates us to train and test more number of proteins and find the suitable method for protein subclass classification. Kumar et al. [18] used random forest and make a three-layer model first to determine enzyme, followed by primary classes and subclasses determination. They have taken 4731 enzyme sequences and used parameters value as mtry=7 and ntree=200. But their accuracy is 83.98 for subclasses prediction. Wang et al. [19] used total 49 subclasses from 6 classes and get the accuracy from 90.86 (18 subclasses for class 1) to 98.03 (6 subclasses for class 6).

All their works tried to justify their results over a fixed set of data and subclasses which motivate us to conduct a study on varying sample size and number of subclasses using three well-known machine learning methods namely logistic regression, SVM, and random forest.

3 Brief Description of Methods Used in the Study

3.1 Experiment Using Logistic Regression Model

Logistic regression is used to solve classification problem for two class classification problems. Instead of directly predicting the class of a given pattern x , this method estimates the probability of x that belongs to the class. Let the two classes be labeled 1 and 0.

$$\text{Let } P(Y = 1|X = x; \beta) = p(x; \beta) = \frac{1}{e^{-\beta^T \tilde{x}}}, \tilde{x} = \begin{pmatrix} 1 \\ x \end{pmatrix}, x \in \mathbb{R}^p \quad (1)$$

$$\text{where } \beta = (\beta_0, \beta_1, \dots, \beta_p)^T, \beta \in \mathbb{R}^{p+1}. \text{ From (1)} P(Y = 0|X = x; \beta) = 1 - p(x; \beta) \quad (2)$$

These two class conditional probabilities can be written as a single expression:

$$P(Y = y|X = x; \beta) = p(x; \beta)^y (1 - p(x; \beta))^{1-y}, y \in \{0, 1\} \quad (3)$$

Using the given dataset $D = \{(x_i, y_i)\}_{i=1}^N$, $x_i \in \mathbb{R}^p$ and $y_i \in \{0, 1\}$ which are independent and identically distributed (iid), the parameter vector β is estimated by maximizing the log likelihood function $l(\beta)$ of D , where $l(\beta) = \log L(\beta) = \sum_{i=1}^N [y_i \log p(x_i; \beta) + (1 - y_i) \log(1 - p(x_i; \beta))]$

So β_j 's, $j = 0, 1, 2, 3, \dots, p$ are estimated by solving the problem: $\max_{\beta} l(\beta)$

One of the methods of estimating β_j 's is to solve the optimization problem (5) using gradient ascent method. For this the gradient of $l(\beta)$ is computed as $\frac{\partial l(\beta)}{\partial \beta} = X^T(y - p)$

$$\text{Where } \tilde{x}_i = \begin{pmatrix} 1 \\ x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix} \text{ with } x_{i0} = 1, .y = \begin{pmatrix} 1 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{12} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1N} & \dots & x_{Np} \end{pmatrix} \quad (4)$$

$$p(x_i; \beta) = p_i, \text{ and } p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ \vdots \\ p_N \end{pmatrix} \quad (5)$$

we can write (5) as

$$\frac{\partial l(\beta)}{\partial \beta} = X^T(y - p) \quad (6)$$

The above derivation yields the following algorithm:

Algorithm for estimating by Gradient Ascent Method

Input: $\{(x_i, y_i)\}_{i=1}^N$

1. Initialize β^{old} , choose θ , and $\alpha \in (0, 1)$
2. Compute y, X # use (4)
3. for $i = 1, 2, 3, \dots, N$ $p_i = \frac{1}{1+e^{-\beta^{old}T\chi_i}}$ # use (1) and (5)
4. Compute p # use (5)
5. $\beta^{new} \leftarrow \beta^{old} + \alpha X^T(y - p)$ # use (6)
6. if $\Delta\beta < \theta$ for 5 iterations in a sequence where $\Delta\beta = \|\beta^{new} - \beta^{old}\|$, return β
else $\beta^{new} \leftarrow \beta^{old}$ and go to Step 3.

3.2 Experiment Using SVM

Support vector machine (SVM), a binary classifier by Vapnik in 1995, works by constructing a hyperplane and tries to separate the closest samples that may be positive or negative from the hyperplane to the maximum by using the decision function $f(x)$. The detail discussion on SVM is given in appendix of the paper by Pradhan D et al. on enzyme classification [20]. We implement this technique, and the findings are explained in Sect. 4 of Computational procedure.

3.3 Experiment Using Random Forest

Based on ensemble learning and decision tree concepts random forest is one of the robust methods which successfully removes the limitation of decision tree by generating multiple trees. It avoids overfitting and handles the missing data well. It performs successfully without even tuning the hyperparameters.

Algorithm

1. for $b = 1$ to B (B = Number of trees to be built)
 - a. Draw a bootstrapped sample Z_b^* of size N from the given training set of size N .
 - b. Grow a random forest tree T_b to the boot strapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached ($n_{min} = 1$).
 - i. Select m predictors out of p the predictors at random where $m = \lfloor \sqrt{p} \rfloor$
 - ii. Select the best variable among m (Here best implies maximum information gain).
 - iii. Split the node into two daughter nodes.
2. Output the ensemble of trees $\{T_b\}_{b=1}^B = \{T_1, T_2, \dots, T_B\}$ (grown in Step 1)
3. Make a prediction for a point not in the training set as follows: $\hat{c}_{rf}^B(x) =$ Majority vote of $\{c_b(x)\}_{b=1}^B$ where $C_b(x)$ means the class of x determined by T_b . Since the expectation of an average of B identically distributed trees is the same as the expectation of any one of them, the bias of averaged tree is the same as that of the individual trees. So, the prediction accuracy can be improved by reducing the variance of the average. An average of B independent and identically distributed (iid) random variables, each with variance σ^2 , is σ^2/B . Hence, the variance of the average can be reduced by taking large B . If the variables are identically distributed but not necessarily independent with positive correlation ρ , the variance of the average is $\rho\sigma^2 + \frac{1-\rho}{B}\sigma^2$. When B is very large, the second term vanishes, if ρ is small, then the variance of average reduces.
Since the process of randomization used in Step 1: (a) and Step 1: (b) (1). results in trees which are less correlated (ρ is small), the variance of the average reduces.

4 Computational Procedure

Scikit learn existing tools for logistic regression, SVM, and random forest are used in the program for building the model in Python. Program is designed to use the binary classifiers to prepare the model for predicting multiple classes. The following algorithm is developed for logistic regression and used in SVM for tuning it to satisfy the requirement.

Step 1: The no of subclasses used for study on a particular set of samples are dynamically transformed into one unique number. The numbers assigned to each subclass follows an incremental order starting from 0.

Step 2: Each number is converted into equivalent binary and the no of bits is same for all unique numbers by placing zeros before the bit values if required.

Step 3: Each bit value is treated as a class label either 0 or 1 for one binary classifier. The total no of binary classifiers will be $\log_2(n)$ where n is the bit size as determined in step 2.

Step 4: For any given sample input $\log_2(n)$ SVM binary classifiers will output p0 to psize-1 predictions in form of 0 or 1.

Step 5: The no generated by combining p0 to psize-1 will generate the no which is the unique no representing a corresponding transformed subclass as the final prediction.

Step 6: Test the samples and predict the class label and find out the prediction accuracy along with precision, recall and f1 score.

Step 7: Continue executing the same from step 1 to step 6 for all varying classes and generate the summary report.

5 Data

Enzyme protein data is collected from expasy.org using ProtParam tool. The attributes of a protein are protein name, protein number, EC number (class and subclass number), number of amino acids, molecular weight, theoretical pI, amino acid compositions Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y), Val (V), Pyl (O), and Sec (U), (B), (Z), (X), total number of negatively charged residues (Asp + Glu), total number of positively charged residues (Arg + Lys), atomic composition: (carbon, hydrogen, nitrogen, oxygen, sulfur), formula, total number of atoms, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). The collected attribute values of each protein are then programmatically placed in a csv file. Features that are not required for computation purposes like protein name, protein number, formula, etc. are filtered out. The subclass labels are extracted and placed as class labels merging with the class it belongs, which represent the label name uniquely. After verifying the protein samples for individually collected subclasses, data is split programmatically to make the required sets for the desired number of subclasses. There are five number of partitions done like 4 subclasses, 6 subclasses, 8 subclasses, 10 subclasses, and 12 subclasses. For each subclass the total 6 sets of data samples are formed with size of 500, 1000, 1500, 2000, 2500, and 3000. It is taken care that each sample set contains equal number of samples from the specified class labels. Finally for building the model and testing 5 number of subclasses and for each subclass groups 6 sets of datasets are available for comparative study.

6 Results and Discussion

In this comparative study the behavior of chosen three classifiers is present in this section after implementation. The detailed outcomes in the form of reports are precisely discussed with help of tables and graphs. Experiment of these three methods on each six sets of data samples is described in data section and produces summary reports containing the accuracy, precision, recall, and F1-score along with weighted averages. The following Tables 1, 2, 3, 4, and 5 present the accuracy, precision, recall, and F1-score for each classifier over individual subclasses on a given sample size from four subclasses having 500 samples to 12 subclasses having 3000 samples. Similarly Figs. 1, 2, 3, 4, and 5 represent all three methods accuracy for the corresponding selected subclasses, and Fig. 6 represents the overall accuracy of all methods used on different set of subclasses for comparison.

It is observed that SVM and random forest both behave consistently well, whereas LR fails to predict well when data size increases and poorly performs for higher number of classes. SVM's result slightly decreases when the sample size increases. As the number of classes increase SVM and random forest both accuracy level slightly decrease but remain above and around 95%. It is observed that random forest does better when the size increases like reaches 100% for class 4 with data size 2500 and 3000. When the number of classes increased the number of samples is reduced for

Table 1 Detailed result for four subclasses using three methods (LR, SVM, and RF)

Size	Accuracy			Precision			Recall			F1-score		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
500	0.91	0.97	0.99	0.92	0.97	0.99	0.91	0.97	0.99	0.91	0.97	0.99
1000	0.85	0.97	0.99	0.87	0.97	0.99	0.85	0.97	0.99	0.85	0.97	0.99
1500	0.84	0.97	0.98	0.85	0.97	0.98	0.84	0.97	0.98	0.84	0.97	0.98
2000	0.74	0.97	0.98	0.75	0.97	0.99	0.75	0.97	0.99	0.74	0.97	0.99
2500	0.62	0.95	1	0.62	0.95	1	0.62	0.95	1	0.62	0.95	1
3000	0.8	0.93	1	0.82	0.93	1	0.8	0.92	1	0.79	0.93	1

Table 2 Detailed result for six subclasses using three methods (LR, SVM, and RF)

Size	Accuracy			Precision			Recall			F1-score		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
500	0.74	0.97	0.97	0.67	0.97	0.97	0.64	0.97	0.97	0.62	0.97	0.97
1000	0.7	0.96	0.96	0.55	0.73	0.96	0.53	0.72	0.97	0.52	0.73	0.96
1500	0.67	0.98	0.99	0.53	0.84	0.99	0.5	0.84	0.99	0.5	0.84	0.99
2000	0.71	0.95	0.98	0.54	0.82	0.98	0.54	0.82	0.98	0.54	0.82	0.98
2500	0.72	0.93	0.97	0.56	0.81	0.97	0.54	0.8	0.97	0.54	0.8	0.97
3000	0.52	0.94	0.98	0.43	0.81	0.98	0.39	0.8	0.98	0.39	0.8	0.98

Table 3 Detailed result for eight subclasses using three methods (LR, SVM, and RF)

Size	Accuracy			Precision			Recall			F1-score		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
500	0.77	0.94	0.97	0.78	0.71	0.98	0.75	0.71	0.97	0.75	0.71	0.97
1000	0.67	0.95	0.97	0.67	0.95	0.98	0.67	0.95	0.97	0.66	0.95	0.97
1500	0.68	0.96	0.98	0.7	0.97	0.97	0.69	0.96	0.98	0.68	0.96	0.98
2000	0.71	0.95	0.98	0.71	0.95	0.98	0.71	0.95	0.98	0.69	0.95	0.98
2500	0.62	0.96	0.98	0.61	0.96	0.98	0.62	0.96	0.98	0.61	0.96	0.98
3000	0.61	0.96	0.98	0.61	0.96	0.98	0.6	0.96	0.98	0.6	0.96	0.98

Table 4 Detailed result for ten subclasses using three methods (LR, SVM, and RF)

Size	Accuracy			Precision			Recall			F1-score		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
500	0.67	0.96	0.96	0.52	0.97	0.96	0.48	0.96	0.96	0.48	0.96	0.96
1000	0.69	0.94	0.98	0.59	0.79	0.98	0.57	0.78	0.98	0.57	0.78	0.98
1500	0.63	0.94	0.98	0.49	0.94	0.98	0.45	0.94	0.98	0.46	0.94	0.98
2000	0.67	0.94	0.98	0.5	0.86	0.98	0.48	0.86	0.98	0.48	0.86	0.98
2500	0.64	0.95	0.98	0.5	0.86	0.98	0.49	0.86	0.98	0.49	0.86	0.98
3000	0.56	0.94	0.98	0.38	0.86	0.98	0.36	0.86	0.99	0.35	0.86	0.98

Table 5 Detailed result for twelve subclasses using three methods (LR, SVM, and RF)

Size	Accuracy			Precision			Recall			F1-score		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
500	0.7	0.96	0.96	0.6	0.96	0.96	0.56	0.96	0.96	0.56	0.96	0.96
1000	0.61	0.95	0.96	0.54	0.95	0.96	0.53	0.95	0.96	0.52	0.95	0.96
1500	0.52	0.92	0.98	0.45	0.92	0.98	0.43	0.92	0.98	0.42	0.92	0.98
2000	0.53	0.92	0.97	0.44	0.85	0.97	0.43	0.85	0.97	0.41	0.85	0.97
2500	0.46	0.92	0.96	0.38	0.85	0.96	0.36	0.85	0.97	0.35	0.85	0.96
3000	0.49	0.95	0.98	0.39	0.82	0.98	0.4	0.81	0.98	0.38	0.81	0.98

each class label which affects the accuracy of random forest though acceptable. It means SVM accuracy diminishes with the sample size increased, and random forest gives little less result when number of classes increased for same size. But overall random forest gives more accurate result throughout the experiments conducted over all variations in sample size and number of subclasses. F1-score of random forest remains consistent with accuracy whereas SVM F1-score decreases with size. This experiment result shows that random forest may be used as one of the best classifiers for protein subclass prediction in real environment.

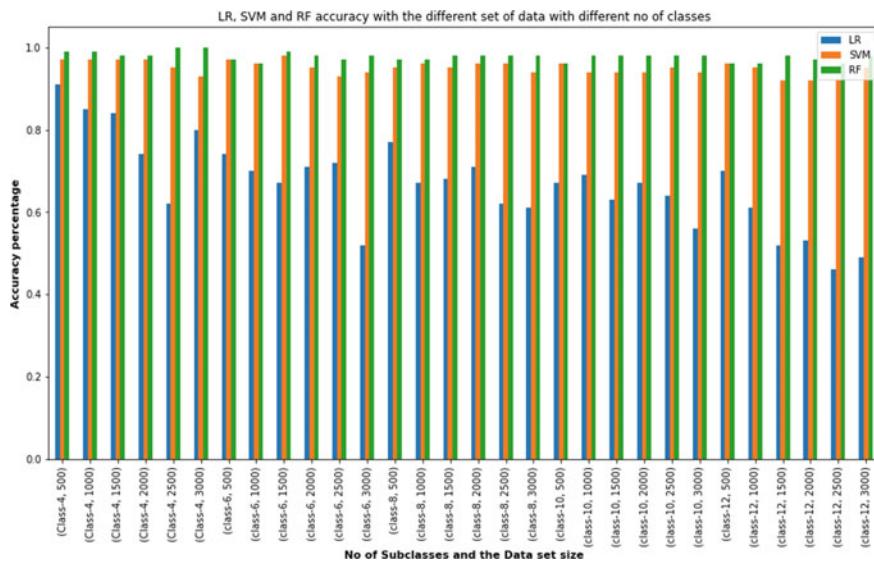


Fig. 1 Overall accuracy for set of subclasses

7 Conclusion and Future Work

The objective of this study is to check and select the consistent classification model for enzyme data subclassification. We test three models with varying data size and number of subclasses to test applicability and sustainability of the methods. The proposed way of using binary classifiers for multiple class prediction in case of logistic regression is not successful but SVM sustained to maintain consistent accuracy as a molded multiclass classifier giving diminishing result with the increased sample size. Random forest classifier is found as the most suitable one in fact when the size of data increased for all class labels it reaches cent percent for small number of classes and never goes below 96%. Based on this study, we would like to suggest random forest as one better classifier to deal with enzyme data. In our future work we would like to extend this work by replacing logistic regression with some other suitable methods and try to cover all the existing subclasses with all or large number of samples for predicting and suggesting the best method for protein subclass classification.

References

- Hackett G, Cole N, Bhartia M, Kennedy D, Raju J, Wilkinson P, Saghir A (2014) Blast study group the response to testosterone undecanoate in men with type 2 diabetes is dependent on achieving threshold serum levels (the BLAST study). *Int J Clin Pract* 68(2):203–215

2. Donkor ES, Dayie N, Adiku TK (2014) Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). *J Bioinf Sequence Anal* 1:1–6
3. Jones NC, Pevzner PA, Pevzner P (2004) In: An introduction to bioinformatics algorithms, MIT Press
4. Wallqvist A, Fukunishi Y, Murphy LR, Fadel A, Levy RM (2000) Iterative sequence/secondary structure search for protein homologs: comparison with amino acid sequence alignments and application to fold recognition in genome databases 16(11):988–1002. <https://doi.org/10.1093/bioinformatics/16.11.988>
5. Tian Y, Shi Y, Liu X (2012) Recent advances on support vector machines research. *Technol Econ Dev Econ* 18(1):5–33
6. Fawagreh K, Gaber MM (2014) Random forests: from early developments to recent advancements. *Syst Sci Control Eng* 2(1):602–609
7. Tian Y, Shi Y, Liu X (2012) Recent advances on support vector machines research. *Technol Econ Dev Econ* 18(1):5–33
8. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network, *Physica D: Nonlinear Phenomena*, March 2020: special issue on machine learning and dynamical systems, vol 404. Elsevier
9. Peng J, Lee KL, Ingersoll GM (2002) An introduction to logistic regression analysis and reporting. *The J Educat Res* 96(1):3–14. <https://doi.org/10.1080/00220670209598786>
10. Hastie, Tibshirani, Friedman (2009) In: Elements of statistical learning. Springer, pp 763
11. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D (1999) Detecting protein function and protein-protein interactions from genome sequences. *Science* 285(5428):751–753
12. Overbeek R, Fonstein M, D'souza M, Pusch GD, Maltsev N (1999) The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci* 96(6):2896–2901
13. Cai YD, Liu XJ, Chou KC (2002) Artificial neural network model for predicting protein sub-cellular location. *Comput Chem* 26(2):179–182
14. Stawiski EW, Mandel-Gutfreund Y, Lowenthal AC, Gregoret LM (2002) Progress in predicting protein function from structure: unique features of O-glycosidases. *Biocomputing* 637–648
15. Dobson PD, Doig AJ (2003) Distinguishing enzyme structures from non-enzymes without alignments. *J Mol Biol* 330(4):771–783
16. Shen HB, Chou KC (2007) EzyPred: a top-down approach for predicting enzyme functional classes and subclasses. *Biochem Biophys Res Commun* 364(1):53–59
17. Debasmita P, Biswajit S, Misra BB, Padhy S (2020) A multiclass SVM classifier with teaching learning based feature subset selection for enzyme subclass classification. *Appl Soft Comput.* <https://doi.org/10.1016/j.asoc.2020.106664>
18. Kumar C, Choudhary A (2012) A top-down approach to classify enzyme functional classes and sub-classes using random forest. *EURASIP J Bioinform Syst Biol* 1
19. Ying W, Xiuzhen H, Lixia S, Zhenxing F, Hangyu S (2014) Predicting enzyme subclasses by using random forest with multicharacteristic parameters protein and peptide letters. 21(3):275–284(10); Bentham Science Publishers
20. Pradhan D, Padhy S, Sahoo B (2017) Enzyme classification using multiclass support vector machine and feature subset selection. *Comput Biol Chem* 70:211–219. <https://doi.org/10.1016/j.combiolchem.2017.08.009>. Epub 2017 Aug 31. PMID: 28934693

Iris Recognition Method for Non-cooperative Images



Zainab Ghayyib Abdul Hasan

Abstract When iris images are collected under optimal circumstances, traditional iris segmentation algorithms provide accurate findings. However, an iris identification system's success is heavily dependent on the precision with which it segments iris pictures, particularly when dealing with irises that are non-cooperative. This research investigates the challenge of recognizing irises in low-quality photos taken under challenging lighting and other imaging situations. In order to reduce processing time and eliminate noise caused by eyelashes and eyelids, the system first acquires an iris image, then improves the image quality, detects the iris boundary and the pupil, detects the eyelids, removes the eyelashes and the shadows, and converts the iris coordinates from Cartesian to polar coordinates. The iris's characteristics are extracted with the use of a Gabor filter and then compared with the help of Euclidean distance. After using the suggested technique, we compared the outcomes with those found in the literature and found that the proposed method yields significant improvements in segmentation accuracy and recognition performance.

Keywords Biometrics · Iris recognition · Iris segmentation · Hamming distance · Iris code · Iris normalization

1 Introduction

During daily basis activities, people usually use names, passwords, and IDs, whenever using computers, ATM, and passing by an airport to verify their identities [1].

With the evolution of technology and the nature of life being so busy people intend to forget passwords and lose IDs or even worse get stolen. Preventing such things can be done using easier authentication methods such as biometrics specs that do not change such as the face, fingerprint, or even voice recognition [2].

Z. G. A. Hasan (✉)

Department of Computer Science, Faculty of Education for Girls, University of Kufa, Kufa, Iraq
e-mail: zainabg.alhatimy@uokufa.edu.iq

Biometrics is defined as physical or behavioral characteristics. Authentication based on biometrics such as face, voice, fingerprints, or iris has been a major topic for computer vision and signal processing. These types of biometric system depend on image processing and signal. The substructure of all biometric features is acquiring the image/input signal and applying several algorithms like fuzzy logic, wavelet transform, neural network, etc., of obtaining the standing out features. Biometrics of physical characteristics incorporates recognizing or confirming traits such as voice, iris, retina, facial highlights, fingerprints, hand geometry, unique finger impression, manually written marks, and palm-prints [3–5].

Iris recognition, one of these areas, has pulled in a part of consideration since it has different profitable components as more noteworthy speed, straightforwardness, and precision matched to other biometric areas. To verify and identify the identity of an individual, iris recognition depended on the special iris patterns of the person.

Personal identification and verification of the iris through automatic recognition are essential. The human iris has a structure that is one of a kind to each person. Even indistinguishable twins have one of a kind iris design. Appropriate division of the iris is a necessity for accurate classification, otherwise bad images will give weak results to the iris recognition system (IRS) [6].

In this paper, we attempt to address the problem of accurately locating the iris area. To that end, we propose a more robust and effective multi-stage algorithm. The suggested technique divides the process of iris segmentation into sections capable of predicting the boundary points properly and enhancing the iris boundary points from non-saliency points.

This paper's contents are as follows. Sections 2 and 3 discuss the structure of the iris and iris segmentation briefly. Section 4 describes a review of the literature. Section 5 presents the proposed methodology for IRS steps in detail. Section 6 will be stated the experimental results, and the conclusions are given in Sect. 7.

2 Structure of Iris

The person's iris has a unique pattern and is distinct for everyone, even twins have various patterns and never change no matter how life changes. Technology using permanent patterns gives positive results and accurate identification levels [7]. Figure 1 shows several specimens of the iris images.

Fig. 1 Sample iris images



The iris is a thin circular diaphragm that stands between the lens and the cornea of the human eye. The pupil is the circular aperture in the iris's center. The iris's role is to restrict the quantity of light that enters the pupil, which it does via the dilator and sphincter muscles, which control pupil size. The pupil size may vary between 10 and 80% of the iris diameter, with an average diameter of 12 mm [8, 9].

The iris has numerous layers, including the stromal layer, which houses the two iris muscles, blood vessels, and pigment cells. The epithelium is the lowest layer and is made up of pigmentation cells with a high density. The density of stromal pigmentation determines iris color. The visible exterior of a multilayered iris consists of two zones that often change color. An inner pupillary zone and an outer ciliary zone are separated by a zig-zagged collarette [10, 11].

Iris formation starts throughout the 3 months of embryonic development [9]. The creation of the individual pattern on the iris surface is instituted through the life's first year, and in first few years the pigmentation stroma is taking place [12]. Iris individual patterns are formed randomly and it is not associated with any embryonic factors [12]. The iris pigmentation is the only distinguishing that is dependent on genetics to determine the iris color. Due to the iris patterns' epigenetic nature, the two eyes of any person include iris patterns are perfectly independent, and also iris patterns are uncorrelated of identical twins [13, 33].

3 Iris Segmentation

It is a critical process in any IRS. Various factors (eyelid hinder, eyelashes, and light intensity) play a role in a faster and more accurate segmentation process. Iris segmentation starts with detecting the iris, any image then goes through the system marking the outer and inner boundaries of the sclera and pupil, and determines if the lower and upper eyelid boundaries close. Inexactness in the representation, identification, and modeling of these boundaries can cause several mappings in the iris extraction. Better quality images result in better iris segmentation [14].

4 Literature Review

Hu et al [15] an algorithm is proposed for segmenting color iris, which leverages the sparsity induced by l_1 -norm to conquer degradation in the images of color iris and noise. The algorithm fits the pupillary and limbic boundary as well as the eyelids by employing l_1 -norm regression on a collection of determined border points. By super pixel established correlation histogram methodology, a coarse iris region is chosen, this method is capable of locating the iris region in images captured at a distance, even when there is noise present (glasses and specular reflection). The suggested segmentation method's usefulness is demonstrated by experimental results on the FRGC and UBIRIS v2 datasets.

Jillela and Ross [16] explore the suitability of using iris texture for biometric recognition in mobile devices. The paper briefly describes the complete iris recognition process and concentrates on iris segmentation in the visible spectrum. The authors discuss different approaches (the clustering and semantic rules, the boundary regularization, the multistage refinement, the Zernike moments, and the color component analysis). Iris segmentation in the visible spectrum challenges is summarized with future directions.

Maria Frucci et al. [17] in this paper present a Watershed transform-based Iris Recognition system (WIRE) for noisy images obtained in visible wavelength. The preprocessing stage for color and illumination adjustment, which is essential for irises with dark pigmentation since corneal specular reflections would dominate their albedo; the criteria used for the watershed transform binarization, resulting in an initial segmentation that is refined by taking into account the watershed regions at least partially included in the best iris fitting circle; the introduction of a new cost function to score the circles detected as potentially delimiting limbus and pupil. The positive effect regarding iris code is mainly the advantage of using a high precision of WIRE in iris segmentation, in which more accurate results are computed moving forward with the execution of iris acknowledgment. Tests carried out on the UBIRIS.v1 session 2 and the subset of UBIRIS.v2 are utilized as training set for the international challenge NICE II.

Liu et al. [18] improvement in the noisy iris, (e.g., blurry images captured at a distance and/or on the move, visible light images together with reflections, etc.) is explored using hierarchical CNNs and multiscale fully CNNs. The authors assembled a HCNN employing organized speaks as input, scales scoping from small to large getting both global and local iris information. The CASIA.v4- distance and UBIRIS.v2 databases are used for experiments.

Radman et al. [19] the segmentation process of the iris images taken in a non-cooperative environment improved in two ways. First, fewer pseudo segmentation outcomes by the HOG-SVM method are proposed for bounding the iris structure, contributing significantly to the diminution of the non-iris region segmentation. Second, the GrowCut technique Abduljalil used in iris texture extraction from surrounding structures by only labeling a few pixels of the iris. Under unconstrained imaging settings, iris images captured using this technique are shown to be very resilient. The best segmentation accuracy enhancement is up to 11% using the suggested method for NICE. II images. In addition, the proposed method is developed and tested on MICHE iris database to reflect the challenges in recognition of unconstrained images taken by mobile devices.

Muhammad et al. [20] the study of choosing defined iris boundaries in noisy eye images capturing visible spectrum from flaw surroundings using densely connected fully CNN. This environment makes the taken iris image exhibits blurriness, unusual glint, low resolution, off-angles, and occlusion. These restraints prevent usual segmentation algorithms from coping. Also, the lack of near-infrared (NIR) light made iris segmentation more braving within the visible light noise. In face of bad quality images densely connected fully CNN proper iris boundary can be determined utilizing better information between the dense blocks of gradient flow.

Sardar et al. [21] this paper proposes utilizing a rough-entropy for iris segmentation, with using the circular sector analysis (CSA) for localization. By incorporating the rough-entropy concept, the paper carefully minimizes the impact of complex types of ambiguities and uncertainties. The performance of the above algorithm was compared with that of the circular Hough transform, despite being computationally intensive, it is state-of-the-art in approximating the region of the iris. The MMU1, IITD, and CASIA-Iris-V3-Interval databases are used for experiments, and the accuracy of this method is 97%.

Zainab and Ebtesam [22] the proposed method is based on truncated total-variation that can keep the strong structures and preserve the good balance between texture removal. Image structure is slowly changing elements like limbic, eyelid boundaries, etc., and different penalization from noise and surrounding textures. This method dependably implies more accurate localizing of pupil circles and iris from eye structure for further segmentation. Eliminating truncated TV texture in this model is used to object segmentation problems or solve other textures that involve noise removal in addition to iris segmentation. Adaptive decision-making methods may be effectively used in an alternative difficult problem in remote sensing and surveillance suffering unwanted occlusions and illumination stability in different conditions. The developed framework during the method proposed provides an effective and robust requirement for applications and researchers attempt to execute iris recognition accuracy capturing noisy images from imperfect environment. The proposed segmentation method that experiments performed on the UBIRIS.v2 has shown effectiveness.

Zhang et al. [23] four feasible network designs are suggested, and through testing and training on identical datasets, the Fully Dilated Convolution combining U-Net (FD-UNet) is identified as the optimal network model. To extract features that are more global, the FD-UNet utilizes dilated convolution instead of original convolution for better image detail processing. In the near-infrared illumination iris datasets, the proposed method tested of iris dataset captured with the visible light illumination (UBIRIS.v2) and (ND-IRIS-0405 and CASIA-iris-interval-v4.0). The experiments results show the network model decreases the error rate and improves accuracy, that performed well of iris datasets that captured with visible light illumination and near-infrared illumination also has good robustness.

Meenakshi et al. [24] in this article the author utilized deep neural networks for classification, feature extraction, and segmentation. Iris segmentation is accomplished using “the Dense Fully Convolutional Network (DFCN)”, which is subsequently followed by normalization of the segmented iris. Then it’s transformed into Gabor wavelets, which are utilized for the purpose of feature extraction. The process of extracting Gabor features involves the use of a multiclass Support Vector Machine classifier (svm). The suggested approach is used on the IITD and CASIA-Iris 1000 datasets.

Bharadwaj et al. [25] in this study, two feature extraction methods are presented, namely the Gabor filter and CNN, and then utilized neural networks and SVM which are various classification algorithms used to analyze the impact of the extracted features on the resulting accuracies. In addition, used the circular Hough transform algorithm and also applied median, Gaussian, and Bilateral filters to iris localization.

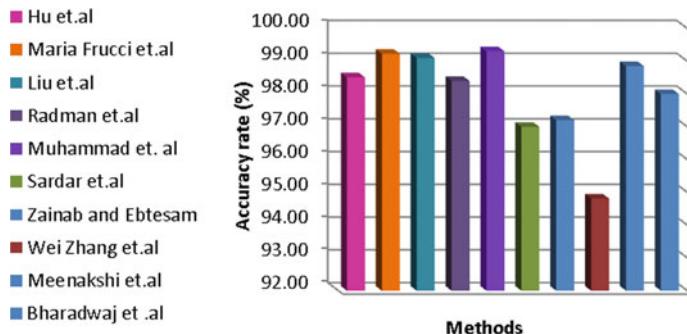


Fig. 2 Accuracy rate of segmentation algorithms

Ultimately, the proposed method obtained the best combination of techniques CNN-NN. To execute experiments for both comparison and testing utilized the CASIA.V1 database which achieved an accuracy of 98% (Fig. 2 and Table 1).

5 Methodology

This section showcases the primary methodology of the proposed system, which is illustrated as a flowchart in Fig. 3. Specifically, segmentation is utilized to improve the image quality, detect the iris boundary and the pupil, detect the eyelids, and remove the eyelashes and the shadows. The normalization stage is utilized to convert the iris coordinates from Cartesian to polar coordinates, whereas the iris's characteristics are extracted with the use of a Gabor filter in the feature extraction stage, then the matching process using Euclidean distance.

5.1 Image Acquisition

UBIRIS. v1 [26] database images are utilized for the input image. Noisy images were taken at long distances and visible light of the UBIRIS. v1 database.

5.2 Segmentation

5.2.1 Converting into Grayscale Image

It removes all of the colors from an image to get a grayscale image.

Table 1 Overview of iris segmentation techniques

Authors	Database	Techniques
Hu et al. [15]	FRGC, UBIRIS-v2	Pupillary and limbic boundary segmentation of iris localization is done using ℓ_1 -norm which induces sparsity. Postprocessing and eyelid fitting are implemented
Jillela and Ross [16]	–	Investigate appropriateness of utilizing recognition of iris texture in mobile devices. Iris segmentation in obvious spectrum is focus of study
Fracci et al. [17]	UBIRIS v1-session 2, UBIRIS v 2	The algorithm used watershed segmentation. Iris segmentation using this technique offers high precision
Liu et al. [18]	CASIA.v4-distance and UBIRIS.v2	Explored hierarchical CNNs and multiscale fully CNNs of noisy iris images for the purpose of enhancing segmentation
Radman et al. [19]	UBIRIS.v2, MICHE	The algorithm used HOG-SVM and Grow Cut. The proposed algorithm doesn't need parameter modification for the different database. It reduces false segmentation
Muhammad et al [20]	UBIRIS.v2, MICHE-I, IIT Delhi v1.0, CASIA v4.0 distance, CASIA v4.0 interval	Iris boundaries are acquired using defined densely connected fully CNN without preprocessing noisy images of the eye obtained in the uncooperative environment
Sardar et al. [21]	MMU1, IITD, and CASIA-Iris-V3-Interval	Utilizing a rough-entropy for iris segmentation, with using the circular sector analysis (CSA) for localization
Zainab and Ebtesam [22]	UBIRIS.v2	The proposed method based on truncated total-variation that can keep the strong structures and preserving the good balance between texture removal
Zhang et al. [23]	CASIA-iris-interval-v4.0, ND-IRIS-0405, UBIRIS.v2	Better detail processing due to dilated convolution to extract more global features instead of the original convolution used in FD-UNet
Meenakshi et al. [24]	IITD and CASIA Iris-1000	Iris segmentation is accomplished using “the dense fully convolutional network (DFCN)”
Bharadwaj et al. [25]	The CASIA.V1	To perform iris localization, employ the circular Hough transform (CHT) algorithm. This algorithm is capable of detecting the circular shape of both the pupil and iris. Before applying the Hough transform algorithm, various filters, such as median, Gaussian, and bilateral are utilized to preprocess the image

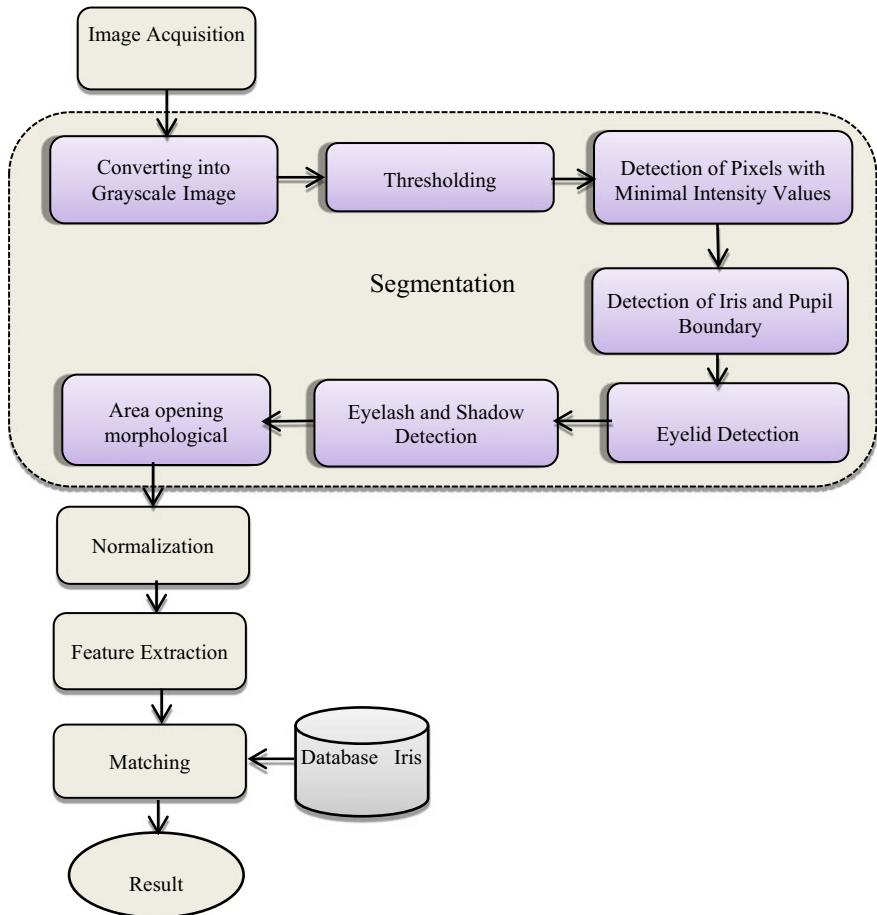


Fig. 3 Methodology of proposed IRS

5.2.2 Thresholding

It is a widely common technique for segmenting images. The notion of below thresholding is used to successfully construct the Daughman's Integro-Differential Operator (DIO) method [27]. Here, pixels that are darker than the background pixels are designated as object pixels using below thresholding, and they are referred to as center pixels. Both the iris and pupil's center pixels are situated within the black pupil area. In certain cases, the center pixels may be located in the iris's comparably brighter part but not in the sclera's white portion. In the eye images, the intensity values of the pixels have been set to a range of $[0, 1]$, with 0 representing black pixels and 1 representing white pixels. To optimize the output, all objects with pixel values less than 0.5 are identified prior to the DIO algorithm implementation.

5.2.3 Detection of Pixels with Minimal Intensity Values

In order to identify local minima in the immediate 3-by-3 neighborhood of a pixel, a threshold image is scanned pixel by pixel. This indicates that the brightness of each pixel is compared with the brightness of its nine closest neighbors. Among these nine pixels, the pixel with the lowest value of intensity is further analyzed. Those remaining pixels are wiped clean. The iris identification process is accelerated by reducing the pixel number of the object on which the Daughman's operator is used.

5.2.4 Detection of Pupil and Iris Boundary

Daughman's equation is used to compute the iris and pupil's center coordinates and radius, respectively. The integro-differential operator equation lies at the core of Daughman's theory of border recognition [27–29].

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (1)$$

$I(x, y)$ specifies the intensity of the pixel at (x, y) in an iris image.

r represents the radius of several circular areas centered on (x_0, y_0) .

σ is the Gaussian distribution's standard deviation.

(x_0, y_0) are iris center coordinates.

$G\sigma(r)$ is a Gaussian filter with the scale sigma (σ).

s denotes the circumference of the circle defined by the parameters (r, x_0, y_0) .

Use selects in which circle path is most effective by changing the radius and center points of its circular shape [29].

5.2.5 Eyelid Detection

In many iris segmentation algorithms, the use of a parabola to simulate the eyelids is a frequent strategy that has been shown to outperform other methods [30]. Our method for searching the parabola is straightforward and very effective in terms of speed and accuracy.

5.2.6 Eyelash and Shadow Detection

The next step is to mask away the pixels related to the shadow under the eyelid and the eyelashes after finding the upper eyelid. This step is identical to the one described in [31]. The threshold was used to determine the points on an eyelash.

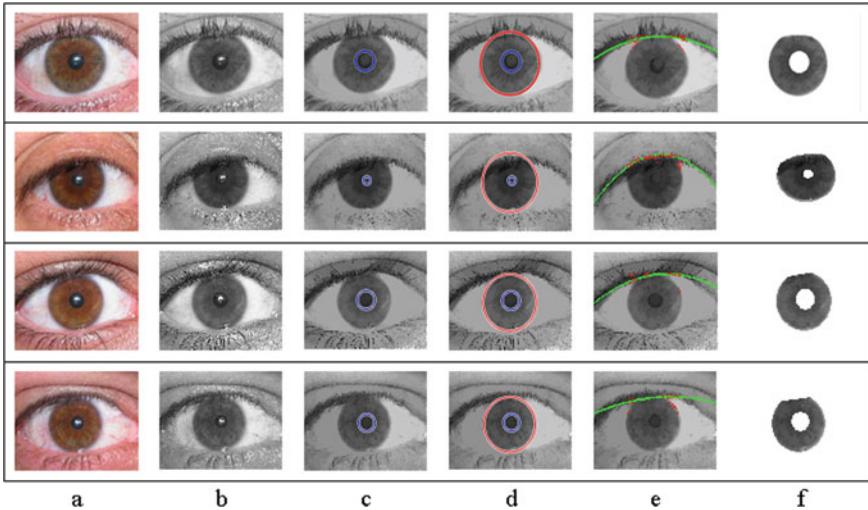


Fig. 4 Sample images from the results of the proposed segmentation method: **a** Original image, **b** Grayscale image, **c** Inner pupil boundary detection, **d** Outer iris boundary detection, **e** Eyelid detection, and **f** Final segmentation results

5.2.7 Area Opening Morphological

By using an opening procedure, isolated noisy pixels in the iris mask may be removed. Figure 4 depicts the sample iris segmentation results from the database used in this investigation.

5.3 Iris Normalization

Normalized images may vary in terms of size and form. It may have an influence on the system's categorization. As a result, the image is adjusted to reduce this impact. The circular iris picture is transformed into a rectangular picture during the normalizing procedure. Daughman's rubber method [13] is utilized to transform the circular intensities in the iris image to polar intensities. The normalizing procedure is shown in Fig. 5a, where the iris point is remapped from the pair of polar coordinates denoted by (r, θ) to a Cartesian position denoted by (x, y) .

$$I(X(r, 0), y(r, 0) \rightarrow I(r, 0) \quad (2)$$

where $y(r, \theta) = y_p(\theta)(1 - r) + ry_i(\theta)$ and $x(r, \theta) = x_p(\theta)(1 - r) + rx_i(\theta)$ and the center coordinates of the iris and pupil are denoted by (x_i, y_i) and (x_p, y_p) , respectively, while the corresponding normalized polar coordinates is represented by (r, θ) . Additionally,

Fig. 5 Example of **a** Iris normalization image, and **b** Iris code



(x, y) signifies the original Cartesian coordinates, and the variable $I(x, y)$ represents the intensity value located at coordinates (x, y) within the iris region image.

5.4 Features Extraction (Iris Code)

Features from $p(r, \theta)$ were extracted using $G(f)$, a one-dimensional Log-Gabor filter, in this study. According to this definition, the $G(f)$ function [28] is:

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right) \quad (3)$$

where σ and f_0 denote the filter's bandwidth and center frequency, respectively. These parameters were determined empirically to be $\sigma = 0.5$ and $f_0 = 0.125$. To save space, however, to get the Iris Code out of $p(r, \theta)$, we followed the method provided in [28]. For this aim, the output of $G(f)$ was quantized into four levels. Iris Code is a bitwise template that was created as a result of this activity. The generated Iris Code is seen in Fig. 5b.

5.5 Matching

It is a recognition procedure in which the iris template is compared with an iris image from the database. The hamming distance (HD), weighted Euclidean distance, and normalized correlation are all utilized as matching techniques. Weighted Euclidean distance [32] requires a great deal of computation, and this metric also requires a great deal of integer values. Additionally, normalized correlation requires a significant amount of calculation [33]. Hamming distance is a more efficient and straightforward approach for matching. The hamming distance is calculated by comparing the newly preprocessed iris image with the previously recorded iris image in the database. The hamming distance between matched images is 0. The hamming distance is computed using the following formula:

$$HD = \frac{\text{Number of differing bits in two iris templates}}{\text{Length of vector}} \quad (4)$$

Table 2 Performance comparison of the recommended techniques with previous methods

Name of the authors	Accuracy rate (%)
Ours	98.92
Meenakshi [24]	98.85
Zhang et al. [23]	98.05
Bharadwaj et al. [25]	98
Sardar et al. [21]	97

6 Results

We tested our method on 500 images from the UBIRIS.v1 database to evaluate the performance of the recommended technique. We achieved a recognition rate of 98.92%. The suggested method's findings are compared with recent approaches such as Meenakshi [24], Zhang et al [23], Bharadwaj et al [25], and Sardar et al [21], which are discussed in Sect. 4. Table 2 summarizes a comparison between the findings of our method and recent works. The suggested technique produced the highest identity verification accuracy.

7 Conclusions

In any biometric authentication system that relies on iris segmentation, if the iris region of an eye image cannot be recognized, the whole procedure fails. This study focuses on establishing an efficient and accurate iris segmentation approach for use in the development of more robust biometric identification systems in a variety of application areas. The technique tackles the problem of processing iris pictures with non-circular pupil and iris borders, as well as non-ideal ocular images. To address these issues, we consider various techniques such as generating grayscale images, selecting image thresholding, removal of the small connected component, finding of iris and pupil boundary detection, eyelid detection, eyelash and shadow removal, and intensity level transformation. This system has a recognition accuracy of 98.92 percent when used with the UBIRIS.V1 database. Our suggested solution outperformed previous IRS in terms of accuracy and performance.

References

1. Bowyer KW, Hollingsworth K, Flynn PJ (2008) Image understanding for iris biometrics: a survey. *Comput Vis Image Underst* 110(2):281–307
2. Amin M, Mohamed N (2021) The evolution of wi-fi technology in human motion recognition: concepts, techniques and future works. In: International computer engineering conference

3. Jain AK, Nandakumar K, Ross A (2016) 50 years of biometric research: accomplishments, challenges, and opportunities. *Pattern Recognit Lett* 79:80–105. <https://doi.org/10.1016/j.patrec.2015.12.013>
4. Jain AK, Ross A, Prabhakar S (2004) An introduction to biometric recognition. *IEEE Trans circuits Syst video Technol* 14(1):4–20
5. Jain AK, Ross A, Pankanti S (2006) Biometrics: a tool for information security. *IEEE Trans Inf forensics Secur* 1(2):125–143
6. Abidin ZZ, Manaf M, Shibghatullah AS, Yunus SHAM, Anawar S, Ayop Z (2012) Iris segmentation analysis using integro-differential and hough transform in biometric system. *J Telecommun Electron Comput Eng* 4(2):41–48
7. Hollingsworth K, Bowyer KW, Lagree S, Fenker SP, Flynn PJ (2011) Genetically identical irises have texture similarity that is not detected by iris biometrics. *Comput Vis Image Underst* 115(11):1493–1502
8. Huang Y-P, Luo S-W, Chen E-Y (2002) An efficient iris recognition system. In: Proceedings. international conference on machine learning and cybernetics, 2002, vol 1. pp 450–454
9. Bodade RM, Talbar SN (2014) Iris analysis for biometric recognition systems. Springer
10. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Alkwai LM, Kumar S (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electr Eng* 107:108617. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
11. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerg Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>
12. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electri Eng* 108:108715. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
13. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun* ISSN 0140–3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
14. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM (2023) An efficient hybrid deep learning model for denial of service detection in cyber physical systems. *IEEE Trans Netw Sci Eng*. <https://doi.org/10.1109/TNSE.2023.3273301>
15. Gupta U, Sharma R (2023) Analysis of criminal spatial events in india using exploratory data analysis and regression. *Comput Electri Eng* 109(Part A):108761. ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108761>
16. Goyal B et al. (2023) Detection of fake accounts on social media using multimodal data with deep learning. *IEEE Trans Comput Soc Syst*. <https://doi.org/10.1109/TCSS.2023.3296837>
17. Sneha, Malik P, Sharma R, Ghosh U, Alnumay WS (2023) Internet of Things and long-range antenna's; challenges, solutions and comparison in next generation systems. *Microprocessors and Microsyst* 104934. ISSN 0141–9331. <https://doi.org/10.1016/j.micpro.2023.104934>
18. Vohnout R et al. (2023) Living lab long-term sustainability in hybrid access positive energy districts—a prosumager smart fog computing perspective. *IEEE Internet of Things J*. <https://doi.org/10.1109/JIOT.2023.3280594>
19. Yu X, Li W, Zhou X et al (2023) Deep learning personalized recommendation-based construction method of hybrid blockchain model. *Sci Rep* 13:17915. <https://doi.org/10.1038/s41598-023-39564-x>
20. Yadav S et al. (2018) Video object detection from compressed formats for modern lightweight consumer electronics. *IEEE Trans Consum Electron*. <https://doi.org/10.1109/TCE.2023.3325480>
21. Sardar M, Mitra S, Shankar BU (2018) Iris localization using rough entropy and CSA: a soft computing approach. *Appl Soft Comput* 67:61–69
22. Ghaib Z, Alshemmary EN (2019) A robust segmentation of non-ideal iris images. *J Adv Res Dyn Control Syst* 11(10):99–103. <https://doi.org/10.5373/JARDCS/V11I10/20193011>

23. Zhang W, Lu X, Gu Y, Liu Y, Meng X, Li J (2019) A robust iris segmentation scheme based on improved U-net. *IEEE Access* 7:85082–85089
24. Meenakshi D (2021) Iris segmentation and recognition using dense fully convolutional network and multiclass support vector machine classifier. *Turkish J Comput Math Educ* 12(13):5418–5428
25. Bharadwaj R, Sujana S (2021) Iris recognition based on Gabor and deep convolutional networks. In: 2021 international conference on communication, control and information sciences (ICCISc), 2021, vol 1. pp 1–6
26. Proen  a H, Alexandre LA (2005) UBIRIS: a noisy iris image database. In: Image analysis and processing–ICIAP 2005: 13th international conference, Cagliari, Italy, September 6–8, 2005. Proceedings 13, 2005, pp 970–977
27. Daugman J (2009) How iris recognition works. In: The essential guide to image processing, Elsevier, pp 715–739
28. Daugman JG (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 15(11):1148–1161
29. Daugman J (2007) New methods in iris recognition. *IEEE Trans Syst Man Cybern Part B* 37(5):1167–1175
30. Min T-H, Park R-H (2008) Comparison of eyelid and eyelash detection algorithms for performance improvement of iris recognition. In: 2008 15th IEEE international conference on image processing, 2008, pp 257–260
31. Tan C-W, Kumar A (2013) Towards online iris and periocular recognition under relaxed imaging constraints. *IEEE Trans Image Process* 22(10):3751–3765
32. Birgale L, Kokare M (2012) Iris recognition using ridgelets. *J Inf Process Syst* 8(3):445–458
33. Khan MT, Arora D, Shukla S (2013) Feature extraction through iris images using 1-D Gabor filter on different iris datasets. In: 2013 sixth international conference on contemporary computing (IC3), 2013, pp 445–450
34. Al-asadi TA, Obaid AJ (2016) Object-based image retrieval using enhanced SURF. *Asian J Inform Technol* 15:2756–2762. <https://doi.org/10.36478/ajit.2016.2756.2762>

An Exploration: Deep Learning-Based Hybrid Model for Automated Diagnosis and Classification of Brain Tumor Disorder



Kamini Lamba and Shalli Rani

Abstract Reproduction of abnormal tissues within the brain due to any damage can cause major concerns for an individuals' health which can be identified by radiologists after examining cell structure of brain that clarifies whether it belongs to benign, i.e., non-cancerous or malign, i.e., cancerous. Although it cannot be treated properly, identifying abnormal growth of tissue at very initial phase can definitely help in preventing from major issues. Most of the researchers described automated brain tumor diagnosing methods in their publications which also received the most attention to provide significant contribution in the healthcare. Authors achieved the highest accuracy of 93.72% via deploying deep learning-based models while predicting brain tumor disease as these models have ability to analyze vast amount of data and able to extract significant features accurately and efficiently as compared to the existing approaches in short duration to provide improved patient outcomes and timely treatment in the healthcare.

Keywords Brain tumor · Health · Increased life expectancy · Deep neural network · Healthcare

1 Introduction

Brain is responsible for controlling each process [1] which regulates an individual's human body such as thought, emotions, vision, breathing, temperature, hunger, touch, and motor skills. It also acts as the control hub of central nervous system due to which it may give the worst impact on people's health if any kind of abnormality exists in

K. Lamba · S. Rani (✉)

Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India

e-mail: shalli.rani@chitkara.edu.in

K. Lamba

e-mail: kamini.2400@chitkara.edu.in

the brain. It is also noticed that whenever there will be a mass or growth of abnormal cells in the brain, it can cause brain tumor in an individual which is not curable yet. Spinal tumor in collaboration with brain tumor can lead to nervous system tumor. For instance, according to the National Brain Tumor Society (NBTS), 1 million Americans have already been suffering from brain tumor, whereas 94,390 people will be influenced by brain tumor in 2023 [2]. Their study also gave estimation based on gender that influenced rate of brain tumor in females is 59%, whereas 41% males are at a risk from such disease. It can also be characterized as benign which is known as non-cancerous and malignant which is known as cancerous depending on the size of its growth in the brain [3].

In benign brain tumor, healthy cells do not come in influence of the infected cells in most of the cases as they remain inactive in the brain, whereas infected cells give the worst impact on nearby healthy cells at a faster pace in malignant brain tumor. Other than this, brain tumor can also be categorized based on the origin [4]. Abnormal tissues occur inside brain in primary brain tumor, whereas secondary tumor may lie anywhere in body except brain. Moreover, it is also categorized in the form of four stages depending upon its boundary and growth rate [5]. The infected cells in the brain do not affect neighboring healthy cells due to the bounding nature in stage 0, whereas infected cells give worst impact on the nearby cells in rest stages. Especially in stage 4, it becomes very complicated to save an individual's life suffering from brain tumor as infected cells in the brain harm other parts of the body too that may lead to the death of a person if not identified and treated timely [6].

According to the researchers, development of brain tumors can be the result of damaged genes on the chromosomes of a cell [7] due to injury or pressure build inside the cranium [8]. However, an individual does not notice any of the symptoms at an early stage due to asymptotic behavior of brain tumor [9] or due to its small size which results in delay in providing treatment and may result in major illness which becomes impossible to cure. Thus, an individual needs to go to radiologist for computerized tomography, magnetic resonance imaging, etc.; if any of the following symptoms as shown in Fig. 1 experienced by an individual without any delay. There is a need to utilize deep learning models[10, 11] for developing an automated computer diagnosis system to provide quick result in terms of accuracy while diagnosing brain tumor as existing techniques consume a lot of time and sometimes may also give inaccurate results. Consequently, early detection of such disease is required at the earliest by radiologist for the survival of a patient. Figure 2 represents samples of healthy brain images, whereas tumor within brain images has been shown in Fig. 3.

Paper has been structured in a way that description of approaches proposed by various researchers has been given in Section 2 followed by Section 3 which provides methodology and obtained results from the existing models have been specified in Section 4 with necessary discussion. Section 5 concludes the study.



Fig. 1 Common symptoms of brain tumor

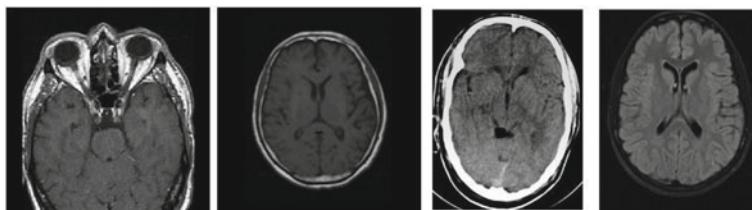


Fig. 2 Healthy brain samples

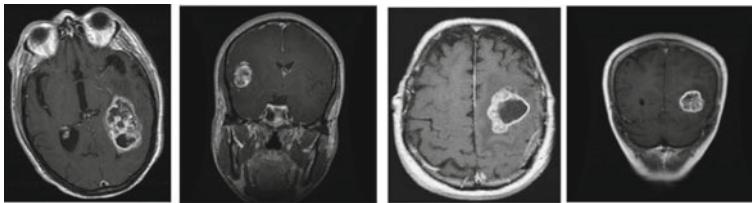


Fig. 3 Tumor within brain samples

2 Literature Review

Most of the researchers provided several surveys to make us aware about the techniques which can be utilized in the detection of brain tumor disease. In this literature survey, description of such techniques has been specified along with its authors. An automated segmentation method for detection of brain tumor has been proposed by Montemurro et al. [12] in which normalization has been performed as a preprocessing step on the basis of convolutional neural network. For performing so, BRATS 2013 and BRATS 2015 datasets have been utilized, whereas in-depth study has been done by Acharya et al. [13] to provide the possibility to put on convolutional neural network directly while segmentation of brain tumor tissues and preprocessing in the form of

standard intensity has been performed while it does not include any post-processing at the output of convolutional neural network for brain tumor detection.

To make it more convenient to achieve better result than existing approaches for brain tumor diagnosis, Badza et al. [14] collaborated convolutional neural network with algorithm of genetic and also used ensemble algorithm for reduction of prediction error variance. A design has been introduced by urban et al. [15] in which multimodality three-dimensional patches, voxel blocks have been used in collaboration with convolutional neural network so that visibility of tissue mark lies within the middle voxel can be made possible for the detection of brain tumor disease. Additionally, utilization of maxout nonlinearity has been done Havaei et al. [16] along with the relevant segment approach for achieving better brain tumor classification accuracy than existing approaches.

Although selection of multiple plane fixes performed by Rao et al. [17] to go through every pixel of considered images while utilizing convolutional neural network having $5*5*5*4$ channel capacity for preparation of a classifier to diagnose brain tumor. A pre-trained network has been proposed by Rehman et al. [18] for classification of images into their corresponding classes which further used fine-tuned model comprising AlexNet, GoogleNet, and VGGNet and achieved more than 95% accuracy in respective approaches.

3 Methodology

While diagnosing brain tumor, basic methodology has been followed in general as shown in Fig. 4. At first, collection of dataset comprising of infected and healthy images has been done. To train our model, split-operation is performed on available dataset which results in 80% training data and 20% testing data is kept for performing validation on test data to ensure efficient results from automated system. Ratio of splitting can be varied depending upon the requirement. Then preprocessing is done to remove noise from the image and also to normalize the data. Other than this, augmentation can also be performed to balance the count of dataset. Then, raw data is transformed into integral form for better representation and extracted features are forwarded to classifier.

4 Results and Discussion

Illustration of outcomes achieved from the existing approaches in terms of accuracy to identify disease has been given in Table 1. Various approaches have been utilized by most of the researchers to achieve better and quick result than traditional approaches for diagnosis of brain tumor.

Moreover, several neural networks have also been integrated with each other as for achieving efficient outcomes during brain tumor detection. However, it is

Fig. 4 Basic methodology for diagnosis and classification of brain tumor

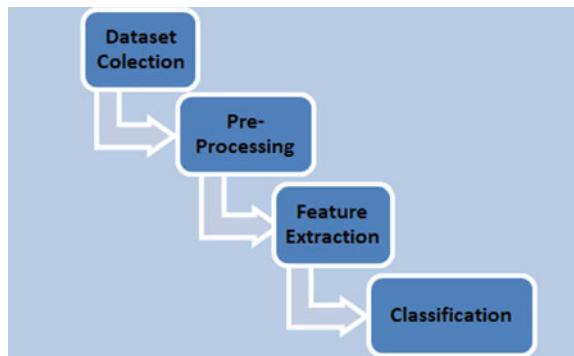


Table 1 Results of existing approaches

Authors/reference no	Approach	Accuracy (%)
Kang et al. [19]	Pre-trained CNN	93.72
Irmak et al. [20]	25-layer CNN	92.66
Kumar et al. [21]	Convolutional neural network	92.30
Siar et al. [22]	Inception model	92
Xie et al. [23]	Multi-view knowledge-based collaborative extensive design	91.60
Anaraki et al. [24]	Genetic algorithm and CNN	90.90
Selvy et al. [25]	Image processing, artificial neural network, histogram equalization	90.90
Kurup et al. [26]	Portable neural network	90
Boustani et al. [27]	Convolutional neural network	90
Pan et al. [28]	Convolutional neural network	87
Ari et al. [29]	ResNet-50	87
Abiwinanda et al. [30]	13-layer CNN	84.19
Sharmila et al. [31]	CNN	80.30
Athency et al. [32]	3D-CNN	75.40

observed that systems proposed by various researchers have their own advantages as well as limitations. For instance, apart from accuracy percentage, a few approaches also experience vanishing gradient issues, overfitting, computational as well as time complexity. In future, focus can be done on the techniques which can overcome such limitations and provide improved and efficient outcomes while diagnosing brain tumor.

5 Conclusion

To monitor various diseases depending upon the utilized techniques, researchers' focus is on developing an automated system for identifying diseases such as brain tumor at an early stage via employing the above-specified techniques as all existing traditional approaches completely rely on experts which are prone to human errors. To eradicate chances of human error, preventing overfitting, time-complexity issues and increase the accuracy of brain tumor disease detection, there is a high requirement for development of an automated computer-aided diagnostic system. Utilization of such system has been suggested by many researchers in which deep learning techniques make it convenient to get accurate result during diagnosis of brain tumor while considering large amount of data for performing any of the analysis which basically train the considered model for identification of any type of anomaly. In the same way, artificial neural network has been considered as one of the popular techniques of machine learning models especially in processing of images as it performs image segmentation and classification efficiently. Moreover, convolutional neural network has also been deployed for brain tumor detection by most of the researchers due to its property of fully connected feed-forward neural network which helps in the reduction of parameters while maintaining models' actual quality.

References

1. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
2. National brain tumor society, braintumor.org/brain-tumors/about-brain-tumors/braintumor-facts/. Last accessed 10 July 2023
3. Tandel GS, Biswas M, Kakde OG, Tiwari A, Suri HS, Turk M, Laird JR, Asare CK, Ankrah AA, Khanna N et al (2019) A review on a deep learning perspective in brain cancer classification. *Cancers* 11(1):111
4. Shah V, Kochar P (2018) Brain cancer: implication to disease, therapeutic strategies and tumor targeted drug delivery approaches. *Recent Pat Anti-cancer Drug Discovery* 13(1):70–85
5. Ahmed S, Iftekharuddin KM, Vossough A (2011) Efficacy of texture, shape, and intensity feature fusion for posterior-fossa tumor segmentation in MRI. *IEEE Trans Inf Technol Biomed* 15(2):206–213
6. Deorah S, Lynch CF, Sibenaller ZA, Ryken TC (2006) Trends in brain cancer incidence and survival in the United States: surveillance, epidemiology, and end results program, 1973 to 2001. *Neurosurg Focus* 20(4):E1
7. Rehni AK, Singh TG, Jaggi AS, Singh N (2008) Pharmacological preconditioning of the brain: a possible interplay between opioid and calcitonin gene related peptide transduction systems. *Pharmacol Rep* 60(6):904
8. Thapa K, Khan H, Singh TG, Kaur A (2021) Traumatic brain injury: mechanistic insight on pathophysiology and potential therapeutic targets. *J Mol Neurosci* 71(9):1725–1742
9. Maharjan S, Alsadoon A, Prasad P, Al-Dalain T, Alsadoon OH (2020) A novel enhanced softmax loss function for brain tumour detection using deep learning. *J Neurosci Methods* 330:108520
10. Srinivas C, KS NP, Zakariah M, Alothaibi YA, Shaukat K, Partibane B, Awal H, et al (2022) Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images. *J Healthc Eng* 2022:1–17

11. Anaya-Isaza A, Mera-Jiménez L (2022) Data augmentation and transfer learning for brain tumor detection in magnetic resonance imaging. *IEEE Access* 10:23217–23233
12. Montemurro N, Condino S, Cattari N, D’Amato R, Ferrari V, Cutolo F (2021) Augmented reality-assisted craniotomy for parasagittal and convexity en plaque meningiomas and custom-made cranio-plasty: A preliminary laboratory report. *Int J Environ Res Public Health* 18(19):9955
13. Acharya UR, Fernandes SL, WeiKoh JE, Ciaccio EJ, Fabell MKM, Tanik UJ, Rajinikanth V, Yeong CH (2019) Automated detection of alzheimer’s disease using brain mri images—a study with various feature extraction techniques. *J Med Syst* 43:1–14
14. Bad’za MM, Barjaktarović MC. (2020) Classification of brain tumors from mri images using a convolutional neural network. *Appl Sci* 10(6):1999
15. Urban G, Bendszus M, Hamprecht F, Kleesiek J, et al (2014) Multi-modal brain tumor segmentation using deep convolutional neural networks. MICCAI BraTS (brain tumor segmentation) challenge. Proceedings, winning contribution, pp. 31–35
16. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H (2017) Brain tumor segmentation with deep neural networks. *Med Image Anal* 35:18–31
17. Rao V, Sarabi MS, Jaiswal A (2015) Brain tumor segmentation with deep learning. MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS) 59:1–4
18. Rehman A, Naz S, Razzak MI, Akram F, Imran M (2020) A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circ Syst Signal Process* 39:757–775
19. Kang J, Ullah Z, Gwak J (2021) Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors* 21(6):2222
20. Irmak E (2021) Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework. *Iran J Sci Technol Trans Elec Eng* 45(3):1015–1036
21. Anilkumar B, Kumar PR (2020) Tumor classification using block wise fine tuning and transfer learning of deep neural network and KNN classifier on MR brain images. *Int J Emerg Trends Eng Res* 8(2):574–583
22. Siar M, Teshnehlab M (2019) Brain tumor detection using deep neural network and machine learning algorithm. In: 2019 9th international conference on computer and knowledge engineering (ICCKE), pp. 363–368. IEEE
23. Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M, Cai W (2018) Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. *IEEE Trans Med Imaging* 38(4):991–1004
24. Anaraki AK, Ayati M, Kazemi F (2019) Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybernetics Biomed Eng* 39(1):63–74
25. Selvy PT, Dharani V, Indhuja A (2019) Brain tumour detection using deep learning techniques. *Int J Sci Res Comput Sci Eng Inf Technol* 169:175
26. Vimal Kurup R, Sowmya V, Soman K (2020) Effect of data pre-processing on brain tumor classification using capsulenet. In: ICICCT 2019—System Reliability, Quality Control, Safety, Maintenance and Management: Applications to Electrical, Electronics and Computer Science and Engineering, pp. 110–119. Springer
27. El Boustani A, Aatila M, El Bachari E, El Oirrak A (2020) MRI brain images classification using convolutional neural networks. In: Advanced Intelligent Systems for Sustainable Development (AI2SD’2019) Volume 4-Advanced Intelligent Systems for Applied Computing Sciences, pp. 308–320. Springer
28. Pan Y, Huang W, Lin Z, Zhu W, Zhou J, Wong J, Ding Z (2015) Brain tumor grading based on neural networks and convolutional neural networks. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp. 699–702. IEEE
29. Ari A, Hanbay D (2018) Deep learning based brain tumor classification and detection system. *Turk J Electr Eng Comput Sci* 26(5):2275–2286

30. Abiwinanda N, Hanif M, Hesaputra ST, Handayani A, Mengko TR (2019) Brain tumor classification using convolutional neural network. In: World Congress on Medical Physics and Biomedical Engineering 2018: June 3–8, 2018, Prague, Czech Republic (Vol. 1). pp. 183–189. Springer
31. Sharmila A, Arun D, Venkatesh J, Sudarshan S, Pranav A (2019) Predicting survival of brain tumor patients using deep learning. *Int J Innovative Technol Explor Eng (IJITEE)* 8(6)
32. Athency A, Ancy B, Fathima K, Dilin R, Binish M (2017) Brain tumor detection and classification in MRI images. *Int J Innov Res Sci Eng Technol* 6:84–89

Recognition of Apple Leaves Infection Using DenseNet121 with Additional Layers



Shubham Nain, Neha Mittal, and Ayushi Jain

Abstract Apple is one of the most popular fruits all over the world, and it is also very liable to diseases like scabs, apple rot, and leaf blotch. These diseases majorly destroy the quality and lead to less healthy production; it is difficult to identify diseases in apples as they appear for a short interval of time so the most prominent way to identify infection is through the condition of their leaves. Most of the leaves are infected by scab, rust, bacteria, and viruses. Early detection is complex for farmers as they all appear the same in shape, color, and texture. Deep learning technology is contributing greatly in this area, addressing this we have proposed an accurately improved Segmentation with the CNN model using a transfer learning model, i.e., DenseNet121 with the weight of ImageNet, and by adding an extra top layer for accurate results. This study also includes the comparative analysis of Seg+ DenseNet121 and some integrated machine learning models. This experiment achieves 99.06% accuracy.

Keywords Apple · Plant disease · DenseNet121 · CNN · Transfer learning

1 Introduction

Deep learning is referring to several problems in the agricultural domain and contributes greatly to increasing the production of crops like smart farming, crop yield prediction, and crop disease detection applications. A convolutional neural network is one of the efficient algorithms contributing greatly to analyze diseases in plants. Leaves disease is a very crucial problem faced by farmers all over the world as it is greatly affecting crop yielding. The main challenge is to timely detect the

S. Nain (✉) · N. Mittal

Department of Electronics and Communication Engineering, Meerut Institute of Engineering and Technology, Meerut, India
e-mail: shubhamnain28@gmail.com

A. Jain

Excel Geomatics Pvt Ltd, Noida, India

disease in fruits and flower plants. Regular monitoring of its health growth is mandatory as it provides a technical basis for the cultivation and selection of a new variety of species [1]. One of the most economic crops in almost the country is Apple [2]. In 2020, apples are cultivated on 308,000 hectares of land and it is doubled in 2021. This usually grows in selected regions of India like Kashmir, Jammu, Uttarakhand's hills, Himachal Pradesh, and some Uttar Pradesh plains [3]. This crop needs regular maintenance as it easily gets infected with fungus like Botryosphaeria obtuse. Apples appear for a short time on trees so disease can be detected through their leaves. Initial symptoms of infection appear on leaves like reddish-brown, scab, rust, blights, black rot, spots, and mites. Leaves change their original texture when infected with any disease, it provides benefits for researchers to analyze images precisely. Identifying these infections with the naked eye needs vast experience and exclusive knowledge [4]. Such factors create challenges for farmers to shift the entire adoption of infected plants to another to control infection to spread [5]. Another, most found disease is rust which appears on leaves surface or stem. It appears black, yellow, orange, and mostly brown powder patches [6]. Conventional techniques used for this purpose are laborious, involve human intervention, and take consuming [7]. Image processing is then opted for by some researchers in which images are analyzed and segmented. Classification and prediction of disease are implemented using neural networks. Veins of leaves are important aspects to categorize and identify as each species has different characteristics, addressing this information Pushpa et al. [8] proposed a convolutional neural network-based method for picking exact features and accurate classification. They used canny edge detection on the Flavia dataset and achieved 95% of accuracy. This algorithm is powerful but still, there are some challenges to overcome like dealing with very small datasets and huge datasets. Kangchen Liu et al. [9] introduced a new method that is PiTLiD which uses the transfer learning model Inception-V3 CNN for small-size datasets. They achieve good results. The most important for good results are extraction and recognition. A study was proposed for extracting, that is LAD-Net, and it can accurately identify infection on apple leaf using a mobile phone in real time with 98.58% accuracy. This study concluded that LAD-Net can detect the affected leaf with the training model weight size of 1.25 MB [10].

Transfer learning is also opted for creating state-of-the-art CNN-based models, Sasikala Vallabhajosyula et al. [11] presented a detailed performance of an ensemble-based model called deep ensemble neural networks (DENN). Abirami T. et al. introduced a new tool that is LeafGAN which identifies disease image-to-image, and this tool modifies image by implementing data multiplication to improve efficiency; it upgrades the efficiency by a percent of 7.4; they also compared it with vanilla LeafGAN which concluded as high quality [12]. Many researchers worked on improving the CNN model for multiple purposes. A few of them added some mathematical expressions, preprocessing, and extra top layers to extract features more accurately. Xingyu Chen et al. proposed faster RCNN and improved faster region CNN, and for this, they utilize residual CNN and opted for some pyramid networks, which resulted in a map value of 83% [13]. Naidu et al. focused on the classification part of the model and proposed multi-layer CNN (MCNN). They used

canonical correlation analysis (CCA) fusion for picking and fusing features. They introduced an ultrasonic sensor for disease detection and achieved 90% accuracy [14]. Two types of feature extraction are implicit and explicit. Himanshu Sharma et al. [7] presented a D-KAP system that benefits some deep learning-based extraction of implicit features, and they achieved 92% of accuracy in identifying apple tree disease. Ahmed Abba Haruna et al. addressed the apple plant infected with foliar disease, they introduced a combination of two strong algorithms that is CNN and long short-term memory (LSTM). These both are deep learning-based models, and they have collected image data from the Kaggle named Plant Pathology 2020—FGVC7 dataset and achieved 98% accuracy [15].

This study is organized into five sections. Section 1 presents the initial idea of the problem addressed and the introduction of the technology with proposed approaches. Section 2 presents the proposed study including an expression of the mathematical operations performed in the CNN model and transfer learning model, and it also depicts some applied preprocessing techniques. Section 3 refers to information on the data collection and some experimental results. Section 4 concludes the results achieved with the comparative table of our model to other approaches. Section 5 complies with the work and challenges faced; it also reflects the future research demand.

2 DenseNet with Additional Layers

This DenseNet with additional layers model focuses on enhancing the basic CNN model by adding extra layers on DenseNet121, which was trained on the ImageNet dataset. The problem faced in a basic model of CNN is the issue of vanishing gradient. Our proposed idealized model overcame this limitation. Figure 1 shows the block-wise description of the idea. This workflow starts with preprocessing the dataset by segmentation, resizing, and converting to generate an appropriate database for picking prominent features and classifying the classes precisely.

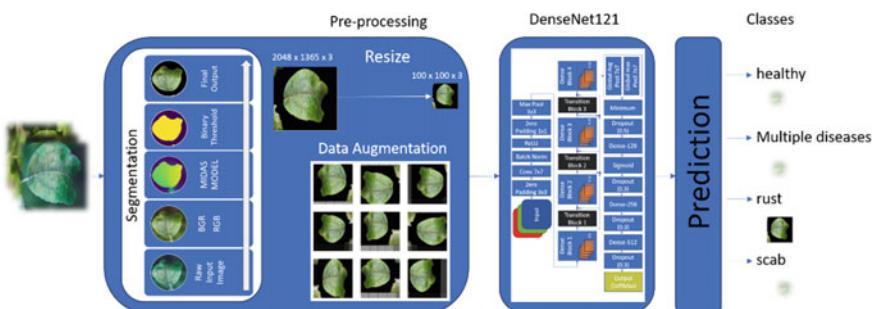


Fig. 1 Block diagram of the proposed system

2.1 Preprocessing

Preparing data is a crucial step before passing it to any kind of network, especially for feature classification. Figure 2 describes the preprocessing techniques, and we implemented them using a block diagram. The images used in this system were big which is $2048 \times 1365 \times 3$, so we resized it to $100 \times 100 \times 3$. Resizing images is important as it reduces the computational time and for a precise result.

Initially, this system also converted the BGR images to RGB and then segmented out foreground and background from the image using Monocular Depth Estimation in Real time with Adaptive Sampling (MIDAS) model algorithm [16] with integrated binary threshold masking to extract the ROI as shown in Fig. 3. This will reduce the complexity of the image as a model can learn easily from these images and reduce the image area on which the model needs to put extra effort for understanding. Normalization is performed to reduce the dimensions. We normalize the image pixel value from 0–255 to 0–1. The dataset was imbalanced and was difficult to distinguish between healthy and unhealthy leaves. Our system is applied augmentation techniques to overcome the issue and equalize all class images.

We applied multiple augmentation techniques like rotating images at a range of 0.35, zooming at 0.2, flipping the vertically as well as horizontally, shifting of 0.15 height and width, and sheared range of 0.2. The images after applying the augmentation technique are shown in Fig. 4.

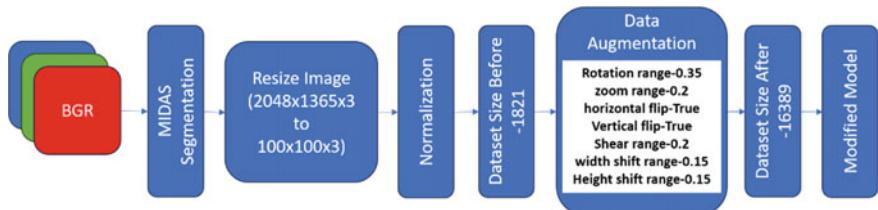


Fig. 2 Preprocessing of database



Fig. 3 BGR image to RGB image

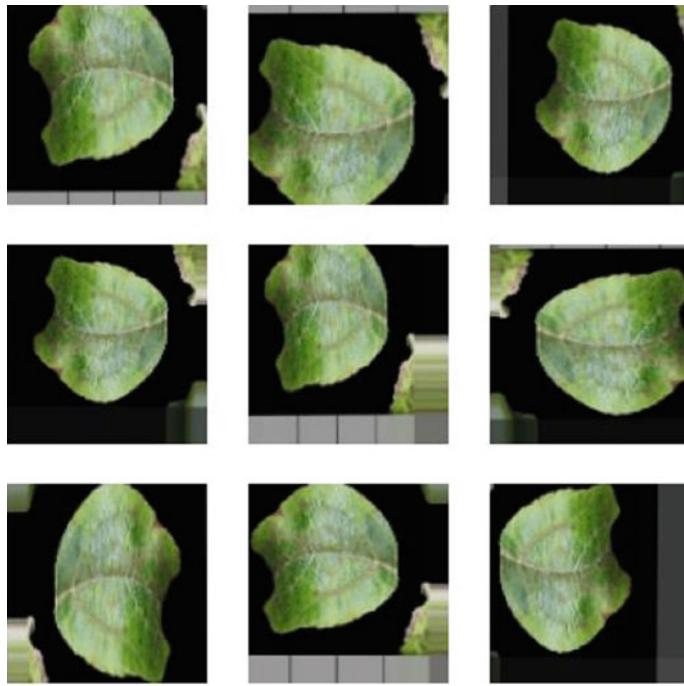


Fig. 4 Augmentation

2.2 Architecture

The convolutional neural network has several main and hidden layers, however when we utilize a transfer-based model like DenseNet. It is divided into three segments such as dense block, transition block, and output block.

Dense Block: In this, a dense block consists of batch normalization, activation function-ReLU, and conv2d 3×3 layers staging the same three blocks onto it means a total of six layers are there for this dense block. It performs the concatenation operation to add features of the topmost existing feature maps.

Transition Block: These blocks consist of batch normalization, activation function-ReLU, conv2d 1×1 , and average pooling layer 2×2 .

Output Block: It involves global Avg 7×7 , global max 7×7 , dropout (0.5 and 0.3), activation function- (sigmoid and softmax), and dense layer.

This model is strengthening the feature propagation and encourages the reusing of features by reducing the parameter number. The overall complexity is handled by the output classifier. Figure 5 displays the modified architecture of the DenseNet model.

This architecture utilizes the Adam optimizer for training the model with a batch size of 32 and an epoch of 30. We have also added some extra top pooling layers

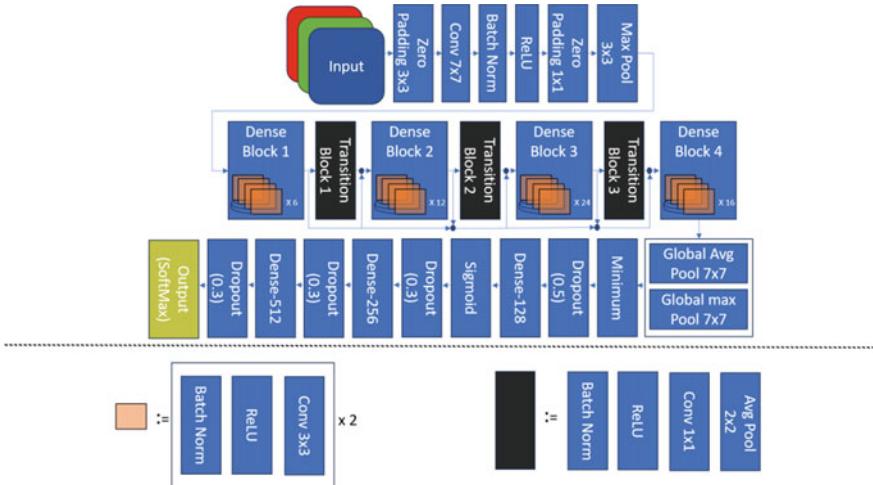


Fig. 5 The architecture of the modified model

like the global average and global max layer. Figure 6 lists all layers involved in the architecture of the model. Some minimum layers with dropouts, dense, and softmax layer. We have updated the system for classifying four classes. Mainly DenseNet121 was trained in 1000 classes.

global_average_pooling2d (Globa (None, 1024)	0	relu[0][0]
global_max_pooling2d (GlobalMax (None, 1024)	0	relu[0][0]
minimum (Minimum) (None, 1024)	0	global_average_pooling2d[0][0] global_max_pooling2d[0][0]
dropout (Dropout) (None, 1024)	0	minimum[0][0]
dense (Dense) (None, 128)	131200	dropout[0][0]
dropout_1 (Dropout) (None, 128)	0	dense[0][0]
dense_1 (Dense) (None, 256)	33024	dropout_1[0][0]
dropout_2 (Dropout) (None, 256)	0	dense_1[0][0]
dense_2 (Dense) (None, 512)	131584	dropout_2[0][0]
dropout_3 (Dropout) (None, 512)	0	dense_2[0][0]
dense_3 (Dense) (None, 4)	2052	dropout_3[0][0]
<hr/> <hr/> <hr/>		
Total params:	7,335,364	
Trainable params:	7,251,716	
Non-trainable params:	83,648	

Fig. 6 Additional layers in DenseNet121

3 Dataset

Dataset taken from [17] for referring to apple plant leaves, fine-grained visual categorization (FGVC7) is a pathology used for image analysis. It consists of four classes for apple leaves disease that is apple scab, rust, healthy, and multiple diseases (cedar, frog eye leaf spot, rot, alternaria leaf spot). All these disease images are shown in Fig. 7 of each class. The number of each class image is shown in Fig. 8 using a bar chart.

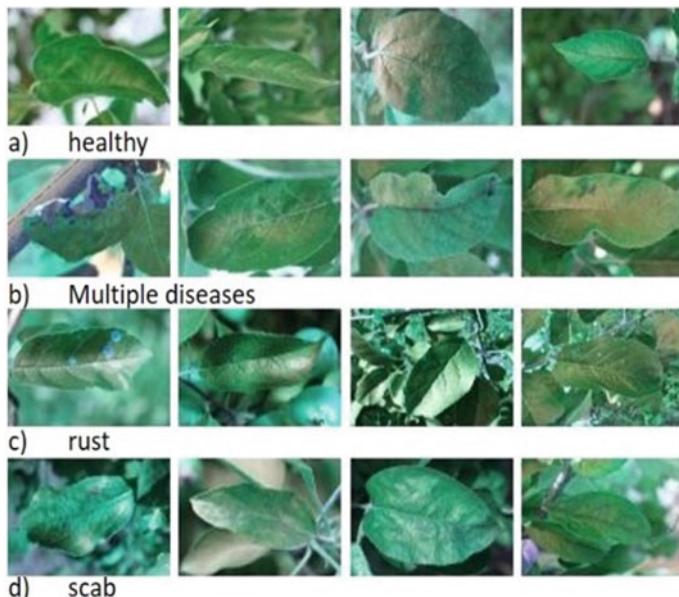
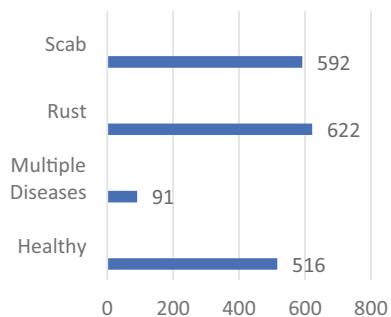


Fig. 7 Images of each class of the dataset

Fig. 8 Bar chart representing classes of database



This database consists 3642 of images that present apple fruit leaves. The data is divided into training and testing data in the ratio of 85:15. The number of images working for training and testing is 13,930 and 2459, respectively.

Fgvc7 is one of the pathologies of plant apple 2020. These images are focusing foliar disease on apples and were taken by DSLR (Canon Rebel T5i). the annotation of the images is also approved by an expert pathologist, especially the images with the infection that are difficult to discern like frogeye leaf spot, alternaria leaf spot, and some complex infection of apple leaves.

4 Results

From the experimental outcome, our modified DenseNet with an extra top layer outperforms well with most similar kinds of images in the dataset. As Fig. 9a, b graphically represent the overall outcome of the model, precision, recall, accuracy, and losses are performance parameters of the segmentation-modified DenseNet121 system.

We achieved 98.94% of accuracy, precision, and recall with 0.04 losses. The recognition rate of each class of the model is represented in Fig. 10.

The mathematical equation of each parameter is enlisted in Table 1. We have experimented with several models with and without segmentation as shown in Table 2. We compared our proposal with a few other systems that utilized the same dataset. The comparative results are listed in Table 3.

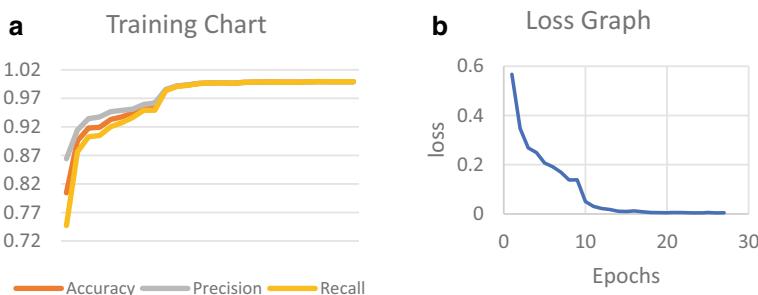
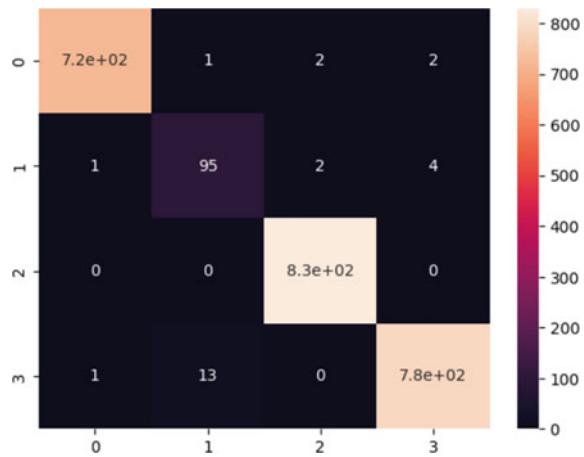


Fig. 9 a Training chart representing accuracy, precision, and recall. b Loss graph representing loss curve

Fig. 10 Confusion matrix**Table 1** Mathematical equation of parameter used in paper

$$\text{Convolution}(z^l) = h^{l-1} \times W^l \quad (1)$$

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (2)$$

$$\text{ReLU}(x) = \max(x, 0) \quad (3)$$

$$\text{maxPooling}(h_{xy}) = \max_{i=0..s, j=0..s} h^{l-1}(x+i)(y+j) \quad (4)$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TN} + \text{FN})} \quad (6)$$

$$F1\text{score} = \frac{2 \times \text{TP}}{(2 \times \text{TP} + (\text{FP} + \text{FN}))} \quad (7)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (8)$$

Table 2 Experimented results with and without segmentation

Model used	Accuracy (%)		Precision (%)		Recall (%)		F1 Score (%)	
	*	#	*	#	*	#	*	#
CNN + SVM	99.06	97.88	97	93	98	96	99	95
CNN + Random Forest	98.90	97.67	96	93	98	96	99	94
CNN + XGBOOST	98.73	97.88	96	92	97	96	99	94
CNN + DesNet121	98.94	98	96	94	98	96	99	95

*Result with segmentation

#Result without Segmentation

Table 3 Comparative analysis

References	Dataset	Methods	Accuracy (%)
[18]	FGVC7	ResNet50	97
[19]	FGVC7	GAN	77.5
[20]	FGVC7	Machine Learning Ensemble	95
[15]	FGVC7	CNN-LSTM	96
[Our]	FGVC7	Seg-DesNet121-Extra top layer+SVM	99.06

5 Conclusion

The growth of crops mainly fruits and flowers is affected due to leaf diseases; therefore, detection at the initial stage is crucial. The fruit yield is majorly affected by various infections like scabs, bacteria, and fungus. An efficient and precise model which overcame several limitations and time complexity in basic CNN architecture. The dataset experimented in this research involves leaves of apples that look almost similar and are difficult to distinguish small infections. The comparative analysis proved that the proposed state-of-the-art DenseNet model is accurate. This system reuses the features and reduces the number of parameters. This automation technique will detect it at an initial stage to prevent damage to plants. For the future scope, the author will further evaluate the same dataset by adding an automated post-process technique.

References

1. Cao L, Li H, Liu X, Chen G, Yu H (2022) Semantic segmentation of plant leaves based on generative adversarial network and attention mechanism. *IEEE Access* 10:76310–76317
2. Bashir S, Firdous F, Rufai SZ (2023) A comprehensive review on apple leaf disease detection. In: 2023 IEEE 8th International Conference for Convergence in Technology (I2CT). IEEE, April, pp 1–6
3. Gupta A, Sinha A, Poddar S (2022) Lesion isolation from apple plant leaves affected by black rot disease using optimized masking on various color channels. In: 2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). IEEE, November, pp 1–5
4. Thangaraj R, Anandamurugan S, Pandiyar P, Kaliappan VK (2022) Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion. *J Plant Dis Prot* 129(3):469–488
5. Wani JA, Sharma S, Muzamil M, Ahmed S, Sharma S, Singh S (2022) Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: methodologies, applications, and challenges. *Arch Comput Methods Eng* 29(1):641–677
6. Bora R, Parasar D, Charhate S (2022) Plant leaf disease detection using deep learning: a review. In: 2022 IEEE 7th International Conference for Convergence in Technology (I2CT). IEEE, April, pp 1–6
7. Sharma H, Padha D, Bashir N (2022) D-kap: a deep learning-based Kashmiri apple plant disease prediction framework. In: 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). IEEE, November, pp 576–581

8. Pushpa BR, Lakshmi P (2022) Deep learning model for plant species classification using leaf vein features. In: 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAIS). IEEE, November, pp 238–243
9. Liu K, Zhang X (2022) PiTLID: identification of plant disease from leaf images based on convolutional neural network. In: IEEE/ACM Transactions on Computational Biology and Bioinformatics
10. Zhu X, Li J, Jia R, Liu B, Yao Z, Yuan A, Huo Y, Zhang H (2022). LAD-net: a novel light weight model for early apple leaf pests and diseases classification. In: IEEE/ACM Transactions on Computational Biology and Bioinformatics
11. Vallabhajosyula S, Sistla V, Kolli VKK (2022) Transfer learning-based deep ensemble neural network for plant leaf disease detection. *J Plant Dis Prot* 129(3):545–558
12. Abirami T, Berlin SN, Johnson S (2022) Prediction of affected leaf of the plant using machine learning. In: 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, April, pp 247–252
13. Chen X, Ye X, Li M, Lou Y, Li H, Ma Z, Liu F (2022) Cucumber leaf diseases detection based on an improved faster RCNN. In: 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC). IEEE, vol 6, March, pp 1025–1031
14. Naidu GG, Ramesh GP (2022) Mango leaf disease detection using ultrasonic sensor. In: 2022 IEEE International Conference on Data Science and Information System (ICDSIS). IEEE, July, pp 1–5
15. Haruna AA, Badi IA, Muhammad LJ, Abuobieda A, Altamimi A (2023) CNN-LSTM learning approach for classification of foliar disease of apple. In: 2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC). IEEE, January, pp 1–6
16. Ranftl R, Bochkovskiy A, Koltun V (2021) Vision transformers for dense prediction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp 12179–12188
17. Kaggle, <https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7/>. Last accessed on 1 June 2023
18. Thapa R, Zhang K, Snavely N, Belongie S, Khan A (2020) The plant pathology challenge 2020 data set to classify foliar disease of apples. *Appl Plant Sci* 8(9):e11390
19. Van Marrewijk BM, Polder G, Kootstra G (2022) Investigation of the added value of CycleGAN on the plant pathology dataset. *IFAC-PapersOnLine* 55(32):89–94
20. Aich P, Ataby AA, Mahyoub M, Mustafina J, Upadhyay Y (2023) Automated plant disease diagnosis in apple trees based on supervised machine learning model. In: 15th International Conference on Developments in eSystems Engineering (DeSE), Baghdad & Anbar, Iraq, pp 160–165. <https://doi.org/10.1109/DeSE58274.2023.10099689>

Techniques for Digital Image Watermarking: A Review



Bipasha Shukla, Kalpana Singh, and Kavita Chaudhary

Abstract The digital revolution poses a significant challenge in terms of authenticating digital images due to the ease of image manipulation. In recent years, ensuring the validity of digital photographs has become a critical subject of research. Various watermarking methods have been devised to tackle this issue, tailored to specific applications. Nonetheless, creating a watermarking system that is both reliable and secure presents a formidable task. This article delves into the intricacies of common watermarking systems, offering comprehensive frameworks. Additionally, it presents a compilation of commonly employed specifications when designing watermarking methods for diverse purposes. The paper also explores the latest advancements in digital picture watermarking technologies, examining their strengths and weaknesses. Furthermore, it sheds light on potential future attacks employing conventional methods.

Keywords Copyright protection · Discrete wavelet transform · Watermarking scheme · Singular value decomposition

1 Introduction

Watermarks are inevitably present on paper banknotes, serving to prevent counterfeiting and secure personal information while indicating the legitimacy of legal documents. Watermarking, a method akin to steganography [1], has been employed for centuries and finds frequent usage on banknotes and postage stamps to detect counterfeit activities. The primary objective of watermarking is to establish authenticity by imprinting a transparent image on the paper. Failing to protect online assets

B. Shukla · K. Chaudhary (✉)
Department of ECED, MIET, Meerut, India
e-mail: kavita.choudhary@miet.ac.in

K. Singh (✉)
Department of UCRD, Chandigarh University, Punjab, India
e-mail: kalpanamnmit@gmail.com

can expose organizations to significant consequences, including asset misuse, erosion of brand reputation, and potential legal penalties [2]. It is concerning that some businesses do not utilize watermark protection despite investing considerable resources in generating original content and the potential risks of exploitation. Producing high-quality imagery requires a substantial investment of time and financial resources for any organization. The entire process encompasses valuable resources such as ideation, photoshoots, and graphic design. Even when collaborating with a company, the creation of an infographic, for instance, can consume up to a month from inception to completion [2, 3].

A watermark which is digital in nature is a type of identifying mark which is covertly incorporated into a signal that is noise-resistant, such as audio or video or that could be image data. To determine who owns the intellectual property rights for a certain signal, it is frequently utilized. Encoding digital data into a carrier signal is the “watermarking” process; the hidden data should, but need not, be connected to the conveyed signal. Watermarks can be used to observe the validity/integrity of the conveyed signal and reveal the names of their holders. It is frequently utilized to track violations of copyright and authenticate banknotes.

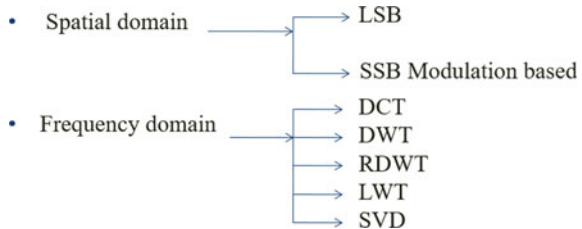
Depending on the use case, several functionalities may be required for a digital watermark. Watermark must be reasonably resistant to changes that can be made to the carrier signal in order to add copyright information to media files. Instead, a tiny watermark would be used to assure integrity [3]. Steganography aims to be imperceptible to the human eye, whereas watermarking prioritizes robustness control. A digital watermark serves as a passive security feature because a digital replica of data is identical to the original. Data is just noted; it is not altered in any way. It is argued that a watermark that is digital in nature is durable with order to the transformations when it can be consistently detected from the marked signal while being subjected to various transformations.

2 Watermarking Techniques

In this paper, we will see the techniques of watermarking. Digital watermarking encompasses a variety of techniques aimed at safeguarding data, which can be broadly categorized into two groups as mentioned in Table 1 and Fig. 1 with the comparison between them.

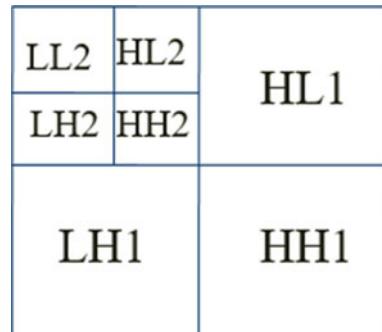
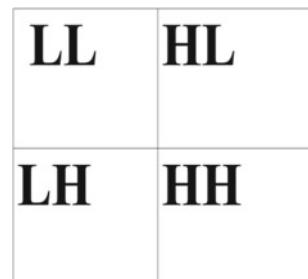
Table 1 Comparisons between spatial and frequency domain techniques

Terms	Spatial domain	Frequency domain
Computation cost	Low	High
Robustness	Fragile	More robust
Perceptual quality	High control	Low control
Capacity	High	low
Example of application	Authentications	Copyrights

Fig. 1 Various techniques

Discrete Wavelet Transform (DWT): The abbreviation DWT represents the discrete wavelet transform technique as shown in Fig. 2. A tool for the hierarchical decomposition of images in mathematics. Excellent localization of space and time. Shift variation is the main drawback [2–4].

RDWT: Redundant discrete wavelet transformation. Shift invariance is applied in it. So it could be able to identify appropriate places to incorporate the watermark. Designed to address the DWT's lack of translation invariance. RDWT is a redundant design by nature. In light of the best sub-band, it therefore distributes the watermark throughout the entire image. These techniques collectively contribute to the efficacy of digital watermarking by facilitating secure information embedding and offering resilience against diverse attacks. The decomposition is shown in Fig. 3.

Fig. 2 DWT's multi-resolution ability**Fig. 3** Image undergoes a single level of RDWT decomposition

3 Foundations of Presented Work

Image watermarking is an effective method for concealing information within an image, providing enhanced security for reliable communication. Various image watermarking techniques have been proposed to improve the security of these images. Several research papers have contributed to the study and comparison of different watermarking techniques, highlighting their uses and limitations.

In a comparative analysis conducted by Manpreet Kaur et al. [1], diverse watermarking techniques are examined, discussing their specific applications and limitations. The primary focus of the study revolves around investigating various image watermarking techniques with the aim of data protection. Jobenjit Singh et al. [5] provide a comprehensive description of the digital image watermarking process, shedding light on its applications and inherent properties. Evaluation of the system's robustness and imperceptibility is carried out using parameters such as peak signal-to-noise ratio (PSNR) and normalized cross-correlation (NCC). Chaturvedi et al. [6] compare two digital image watermarking methods, DWT and DWT-DCT, by analyzing their PSNR performance. The study concludes that the DWT-DCT method is the most suitable technique for level one watermark embedding. Mohan Durvey et al. [7] conduct a comprehensive investigation into various digital watermarking techniques and their contributions across various fields. The paper explores the features, applications, challenges, limitations, quality, and performance of these techniques. Jaishri Guru et al. [8] specifically focus on the study of diverse watermarking algorithms for digital images. These algorithms are designed to enhance factors such as robustness, capacity, security, and other crucial aspects of the watermarking process. Hai Tao et al. [9] carry out an analysis and evaluation of a watermarking system's performance in both the transform and geometric domain invariant regions. The study takes into consideration fundamental attributes of watermarking in their analysis.

These research papers collectively contribute to the overall understanding of image watermarking techniques, their practical applications, and the evaluation of their performance across different domains.

4 Process of Watermarking

Watermarking is a technique used to embed data, such as images or audio, into an image file, which serves to identify the copyright information of the file. The process of image watermarking involves two main steps: the insertion of the watermark into the sources like image or it could be a video, or an audio, and the subsequent extracting watermark from the carrier signal [10, 11].

- (a) The process of placing a watermark into a signal involves following steps also shown in Figs. 4 and 5 [12]:

Fig. 4 Embedding the watermark

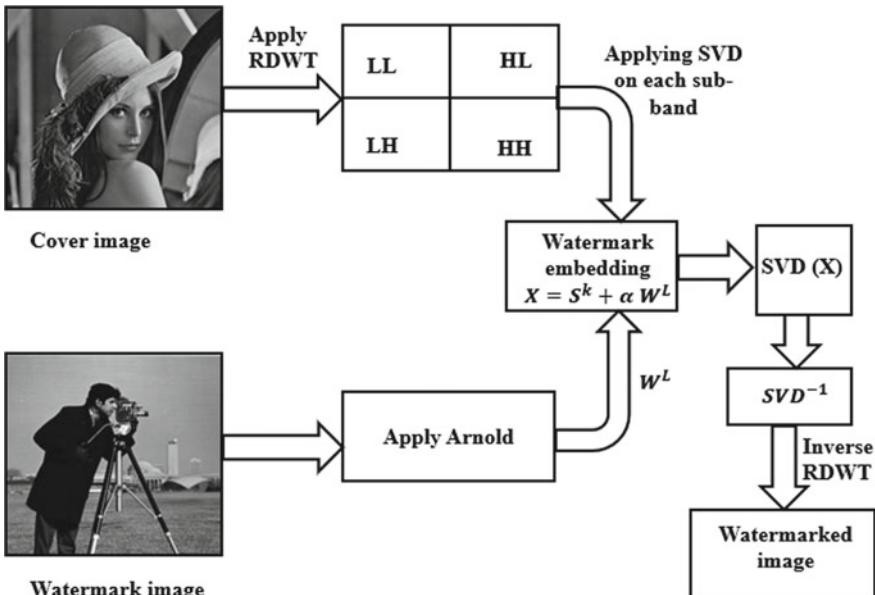
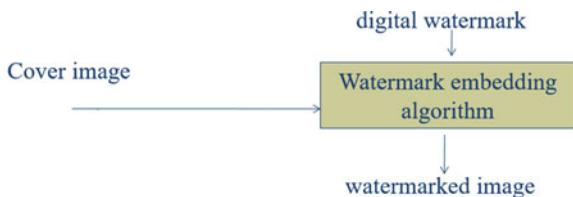


Fig. 5 Watermark embedding process

1. Selecting the carrier signal: Choose the signal, such as an image or audio file, in which the watermark will be inserted.
 2. Selecting the watermark: Choose the watermark that will be inserted into the carrier signal. The watermark could be in the configuration of an image or it could be a video, an audio, or text or any other desired format.
 3. Embedding the watermark: Use a watermarking algorithm or technique to embed the selected watermark into the carrier signal. This process modifies the signal in a way that incorporates the watermark while maintaining the integrity of the original content.
- (b) The process of extracting a watermark from an image involves the following steps:
1. Selecting the watermarked carrier signal: Choose the signal, such as an image or audio file, from which the watermark needs to be removed.

2. Watermark detection: Apply a watermark detection algorithm or technique to analyze the watermarked carrier signal and identify the presence and characteristics of the embedded watermark. This step involves examining the signal and identifying patterns or features specific to the watermark.
3. Watermark extraction: Once the watermark is detected, extract it from the carrier signal using the information obtained during the watermark detection process. This step involves isolating and separating the watermark from the original signal, resulting in the retrieval of the embedded watermark data.

It is important to note that the extraction process should be performed accurately to ensure the integrity and authenticity of the extracted watermark as shown in Figs. 6 and 7 [12].

(c) Arnold transform

- It is an encrypting tool used to muddle the watermark and increase the image's robustness while watermarking images.
- After a number of cycles, the original image can be reconstructed in accordance with the periodicity of the Arnold transform.

Fig. 6 Extraction of watermark

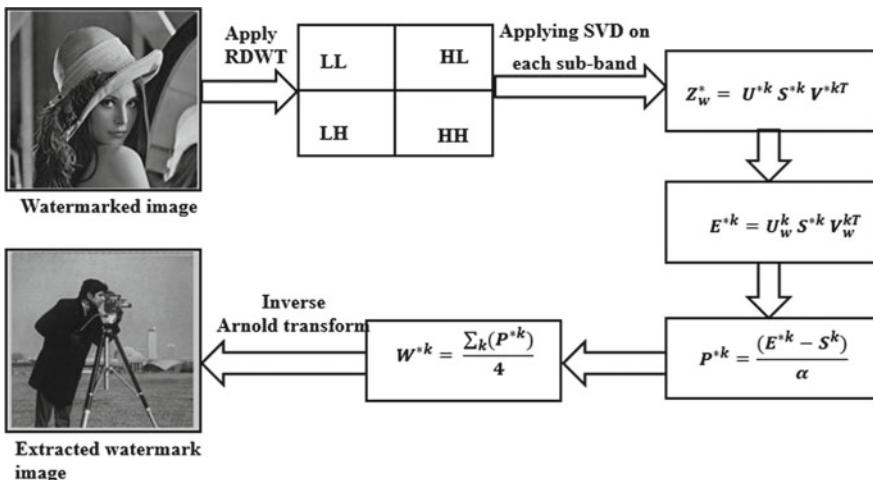
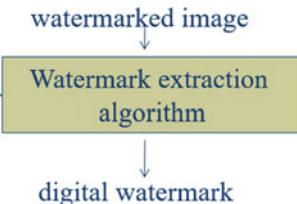


Fig. 7 Process of extracting the watermark

- The image recovery process can be time-consuming due to the reliance of Arnold scrambling on the size of the image, which affects its periodicity.
- The same number of iterations that were used during encoding must be employed to perform an inverse Arnold transformation on the image.

An image's security, dependability, and robustness are increased through scrambling. After transformation, the output is a completely useless image. Typically, it is a preprocessing step before information is hidden. There are several ways to scramble data, including the Arnold transform and pseudo-sequence (random). This thesis' use of the Arnold transform as a method of scrambling gave it the benefit of simplicity and periodicity. V. I. Arnold introduces the Arnold transform and suggests cat mapping as a different name for Arnold. Below is an example of an Arnold transform for a P-P picture.

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} (\text{mod } P) \quad (1)$$

P here stands for the size of the image. X' and Y' stand for the original pixels X and Y that have been jumbled. The equation is used to figure out the inverse Arnold transform.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \end{bmatrix} (\text{mod } P) \quad (2)$$

5 Characteristics of Watermarking

The attributes of image watermarking that need to be considered during the design of a watermarking system are as follows:

- Impalpability: The watermark should be imperceptible or minimally perceptible to human senses. It should not cause significant degradation or distortion to the quality of the image in which it is embedded. The embedded watermark should blend seamlessly with the original image.
- Security: The watermarking system should provide a high level of security to prevent unauthorized access, tampering, or removal of the watermark. Strong encryption and authentication mechanisms may be employed to ensure the integrity and authenticity of the watermark [13].
- Capacity: Capacity refers to the amount of data that can be embedded within the image as a watermark. It depends on the size and complexity of the watermarking algorithm. Higher capacity allows for the embedding of more information, but it may also affect the perceptual quality of the image.

- Robustness: Digital watermarks are deemed “fragile” [14] if they become undetectable even after minor changes. For tamper detection (integrity proof), fragile watermarks are frequently utilized. Alterations performed on a unique piece which are immediately recognizable are frequently referred to as generalized barcodes rather than watermarks. If a digital watermark can withstand benign alterations but not malignant ones, it is said to be semi-fragile. Detecting malignant changes frequently makes use of semi-fragile watermarks.
- Localization: The ability to localize the watermark within the image is important for applications where specific regions or objects need to be protected or identified. Localization allows for precise extraction and verification of the watermark from the desired regions.

These factors play a crucial role in determining the effectiveness, reliability, and usability of an image watermarking system. The balance between these factors needs to be carefully considered based on the specific requirements and constraints of the application.

6 Parameters of Quality Evaluation

The quality parameter is evaluated by assessing various factors or metrics that determine the overall quality of a system or process:

- Evaluation of Peak Signal-to-Noise Ratio (PSNR)

$$\text{SNR} = 20 \times \log\left(\frac{255}{\sqrt{\text{MSE}}}\right) \quad (3)$$

where SNR is signal-to-noise ratio and MSE is the mean square error.

- Mean Square Error:

$$\text{MSE} = \frac{1}{P \times Q} \sum_i \cdot \sum_j [\mathbf{A}(i, j) - \mathbf{A}_w(i, j)]^2 \quad (4)$$

where $\mathbf{A}(i, j)$ is the original image and contains $P \times Q$ pixels and $\mathbf{A}_w(i, j)$ is the watermarked image

7 Applications of the Image Watermarking

Image watermarking has various applications in different domains. Here are some examples:

1. Tamper Detection: Image watermarking can be used to detect any tampering or unauthorized modifications made to an image. By comparing the extracted watermark with the original watermark, the integrity of the image can be verified.
2. Telecast Monitoring: Watermarking can be applied to monitor the telecast of content, particularly in advertisements, to ensure compliance with contractual agreements between advertisers and broadcasters. Watermarks can help identify the source and ownership of the content.
3. Software Clipping: Watermarking is utilized in software to provide a trial version with limited functionality. Certain features, such as saving or printing, may be disabled until the user purchases a registration key or license. Watermarks can indicate whether the software is fully licensed or not.
4. Copyright Protection: Image watermarking is commonly used to protect copyrighted material. Copyright information or ownership details can be embedded as a watermark in the image, allowing for easy identification and proving ownership in case of copyright infringement [15].
5. Authentication and Integrity Validation: Watermarking can be used to verify the authenticity and integrity of an image. Fragile watermarks, which are highly sensitive to modifications, can be applied to detect any unauthorized changes made to the image.
6. Medical Applications: In the medical field, image watermarking can play a role in maintaining patient data confidentiality. Watermarks can be used to protect medical images and ensure that only authorized personnel can access and view sensitive patient information.

These are just a few examples of the diverse applications of image watermarking. The specific use cases and requirements may vary depending on the industry and context in which watermarking is applied.

8 Conclusion

Image watermarking plays a crucial role in secure data transmission, particularly on Internet. While various techniques of watermarking have been kept foreword for embedding watermarks in images, the security of the embedded watermark itself has often been overlooked in existing literature. It is important to consider the security of the embedded watermark to ensure the overall effectiveness of the watermarking system.

To address this gap, further research and implementation can focus on enhancing the security of embedded watermarks. This includes exploring encryption and authentication mechanisms to protect the integrity and confidentiality of the watermark image within a media file [16]. By incorporating robust security measures, the watermarking system can provide enhanced protection against unauthorized access, tampering, and counterfeiting. By considering both the embedding process and the security of the embedded watermark, the overall effectiveness and reliability of the

watermarking system can be improved [17]. This will contribute to more secure data transmission and protection of intellectual property rights in various domains.

References

1. Kaur M (2014) Review paper on digital image watermarking technique for robustness. *Int J Adv Res Comput Sci Softw Eng* 4(5):948–952
2. Thapa M (2014) Digital image watermarking technique based on different attacks. *Int J Adv Comput Sci Appl (IJACSA)* 2(4):14–19
3. Radhika V (2013) Comparative analysis of watermarking in digital images using DCT & DWT. *Int J Sci Res Publ* 3(2):1–4
4. Yuefeng Z (2015) Digital image watermarking algorithms based on dual transform domain and self-recovery. *Int J Smart Sens Intell Syst* 8(1):199–219
5. Chahal J (2013) A review on digital image watermarking. *Int J Emerg Technol Adv Eng Website* 3(12):482–484
6. Navnidhi C (2012) Comparison of digital image watermarking methods DWT and DWT-DCT on the basis of PSNR. *Int J Innovative Res Sci Eng Technol (IJIRSET)* 1(2):147–153
7. Durvey M (2014) A review paper on digital watermarking. *Int J Emerg Trends Technol Comput Sci (IJETTCS)* 3(4):99–105
8. Guru J (2014) A review of watermarking algorithms for digital image. *Int J Innovative Res Comput Commun Eng* 2(9):5701–5708
9. Tao H (2014) Robust image watermarking theories and techniques: a review. *J Appl Res Technol* 12(1):122–138
10. Savakar DG, Pujar S (2018) Digital image watermarking using DWT and FWHT. *Int J Image Graph Signal Process* 11(6):50
11. Shukla B, Singh K (2023) Based on RDWT and SVD Arnold transform: a strong and secure image watermarking method. In: 2023 International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)
12. Reza S (2012) An approach of digital image copyright protection by using watermarking technology. *IJCSI Int J Comput Sci Issues* 9(2):280–286
13. Gupta V, Barve A (2014) A review on image watermarking and its techniques. *Int J Adv Res Comput Sci Softw Eng* 4(1):92–97
14. Thawkar S (2012) Digital image watermarking for copyright protection. *Int J Comput Sci Inf Technol* 3(2):3757–3760
15. Porwal S (2013) Data compression methodologies for lossless data and comparison between algorithms. *Int J Eng Sci Innovative Technol (IJSIT)* 2(2):142–147
16. Yasmeen F, Uddin MS (2021) An efficient watermarking approach based on LL and HH edges of DWT–SVD. *SN Comput Sci* 2(2):82
17. Tang M, Zhou F (2022) A robust and secure watermarking algorithm based on DWT and SVD in the fractional order Fourier transform domain. *Array* 15:100230

Improved Traffic Sign Recognition System for Driver Safety Using Dimensionality Reduction Techniques



Manisha Vashisht and Vipul Vashisht

Abstract Recent developments in the field of emerging technologies including Artificial Intelligence and Machine Learning have led to wide interest in designing and developing innovative solutions in the area of traffic management and human safety. Multiple researchers have used these technologies to propose solutions for traffic sign image management that can lead to driver safety and reduction in number of accidents [10, 12, 14]. To prevent road accidents, traffic signages on roads are vital parameters that can help drivers to take timely decisions and preventive measures to avoid accidents. There is a strong need to improve the traffic sign identification and detection so that irrespective of weather conditions and degradation of sign boards, still driver navigation system is able to identify the correct signs and help in decision making ([13]; Bhatt and Tiwari, Smart traffic sign boards (STSB) for smart cities [Bhatt DP, Tiwari M (2019) Smart traffic sign boards (STSB) for smart cities. In: 2nd Smart Cities Symposium (SCS 2019), pp 1–4, March. IET]). Researchers have been using public traffic sign datasets to find ways to enhance the precision of image recognition approaches. In the previous published research, Vashisht and Kumar [21], have proposed a 3D color texture-based approach for detecting the traffic sign images by making use of ML algorithms and ANN. In this paper, the research is further extended using dimensionality reduction techniques used for feature reduction on Mapillary traffic sign image dataset. In terms of organizing the rest of the paper, the next section of background covers the previous research work done in dimensionality reduction. Next, the authors have explained the proposed methodology for feature selection and ANN design. Then, results from the implementation using ranking algorithms, classifier algorithms, and their comparisons are described. In the end, authors concluded the paper with suggested future direction of work.

Keywords Image processing · Traffic sign management · Machine learning algorithms

M. Vashisht ()
Lingayas Vidyapeeth, Faridabad, Haryana, India
e-mail: manishavashisht@gmail.com

V. Vashisht
Lagozon Technologies, Delhi, India

1 Introduction

In most cases, data mining applications such as traffic sign images require remarkably high-dimensional dataset that could lead to phenomenon of “Curse of Dimensionality”. Studies have shown that not all the features of a dataset are required for accurate predictions, there could be features which could either be redundant, irrelevant, or a direct derivable of any existing feature. Processing of high dimensions could lead to reduction in the accuracy level of classification algorithms since there could be several dimensions which though termed as insignificant but became part of processed large image dataset. In order to overcome this challenge, researchers make use of feature reduction approaches and ranking algorithms to identify only the most significant dimensions from the image dataset for processing purposes. This approach has led to considerable increase in accuracy of classification algorithms and has also resulted in reduction in the hardware computation cost budget [1, 7, 18, 19]. This research work has been executed using the Mapillary traffic sign images dataset. This dataset contains 100,000 images captured from across the world. The dataset has 300 classes available along with the annotations [15].

2 Literature Review

In this section, authors have summarized the earlier research work performed on dimensionality reduction and feature selection approaches.

In this paper, authors have proposed an unsupervised method while making use of neural networks to achieve dimensionality reduction. Authors have used wavelet decomposition along with the principal component analysis (PCA) technique, which resulted in considerable reduction of 64% in dimensionality and achieved classifier accuracy of 88.5% [9]. Migenda et al. [16] proposed neural network-based PCA technique, where they used linear independent variable in place of correlated variables to auto-update with the optimal number of principal components providing highly competitive results. In this paper, authors evaluated a convolutional neural network with a large dataset of Chinese roads containing 10,500 images and achieved accuracy of 84.22% [2]. Ayesha et al. [3] presented a detailed study of various dimensionality reduction techniques and their applicability in specific application areas. Authors also detailed out the challenges faced due to high dimensions impacting accuracy and results.

In this paper, authors raised a new feature selection approach where they combined the two approaches, chi-square test and minimum redundancy test. Chi-square test was used for sample data selection which is closely related to the classes and reduces the data scale. The minimum redundancy algorithm was then implemented to eliminate redundancies. The classification of multi-dimensional data for prediction was performed using the K-nearest neighbor and support vector machine. The result showed that the approach was effective in causing improvement to classification

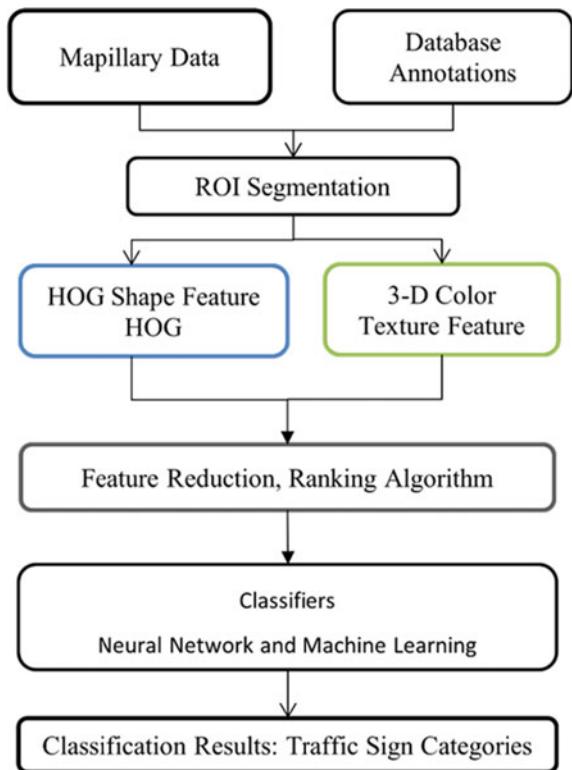
accuracy [25]. Niu et al. [17] proposed feature selection approach to determine the optimal feature set using deep learning model. The corresponding sensitivity analysis results demonstrated the superiority of the proposed model in comparison with the sixteen benchmarks research work. In this paper, authors have put forward a slicing view of multiple approaches applied to feature selection and feature extraction as applied to dimensionality reduction. The review work includes details of datasets, classifiers, algorithms, and comparative results. It was found that amongst SVM and KNN algorithms, SVM achieved better accuracy. The optimized PCA demonstrated improved accuracy and computational time, with the reduced number of dimensions [26]. In this paper, Chao et al. [5] explained the feature reduction as applied to principal component analysis methods. The authors provided the pros and cons of all approaches and concluded by proposing various problem statements. Cheng and Lu [6] proposed a system for scoring features using an enhanced optimization algorithm and was used in gene-selection research in bioengineering. The experimental results showed a high disease recognition accuracy of 0.99.

Gui et al. [8] provided their perspective of feature selection through a detailed survey paper that relationship among different formulations when combined with other machine learning algorithms for specific applications such as classification and clustering. The paper provided detailed comparative analysis based on regression and regularization approaches. Khalid et al. [11] analyzed various feature selection techniques with reference to evaluation on effectiveness of performance on learning of algorithm which results in enhancing the predictive accuracy of classifier.

3 Proposed Methodology for Model Design

In this section, the approach used for implementing the dimensionality reduction and feature selection approach applied to Mapillary traffic sign image dataset is described. As shown in Fig. 1, images from the publicly available Mapillary traffic sign image dataset have been taken with the database annotation with labeled classes, which would later be used by supervised machine learning algorithm for object detection. The region of interest in the image frame is defined during segmentation, which would be used for pre-processing. Histogram of Oriented Gradients (HOG) and 3-D color texture feature descriptors are then used for object detection. These approaches use structure or the shape, color, and texture of an object for the purpose of object detection [18]. Then, feature selection is executed using various ranking algorithms along with ANN used as a classifier. The best ranking algorithm is then implemented along with multiple machine learning classification algorithms for performing analysis to identify the high affinity approach.

Fig. 1 Proposed methodology diagram



3.1 Feature Selection and Dimensionality Reduction Approach

The Feature Selection Library available from MATLAB tool library has been used for electing only a subset of measured features (predictor variables) to create a model. Authors have used three ranking algorithms for feature reduction. These algorithms will then choose a subgroup of reduced but relevant set of features established on the degrees of relevance of dimension (Table 1).

Under Mapillary dataset, each image in region of interest has nearly 6528 data points. After the HOG feature implementation, 324 data points were obtained from the images. Then, once the 3D color texture feature was executed, the date points reduced to 48. At an overall level, an image of 6528 data points gets reduced to 372 data points. The 372 features were preprocessed and after removing the less significant dimensions the date points got reduced to 300. The proposed network has 300 input neurons and includes 10 number of hidden layers and a single output layer. Weights are stated as w and bias are represented as b are the learnable parameters of proposed neural network model. Authors have used Levenberg–Marquardt Algorithm (LMA) for preparing the ANN model using the training technique. LMA

Table 1 Ranking algorithms used for feature selection

Ranking algorithm	Description
Chi-Square	fscchi2 method from MATLAB tool library has been used for univariate feature ranking for classification using chi-square tests. This formula is used to evaluate two or more numerical data sets, in cases where the data comprises of variables spread through several categories and is represented by χ^2 . The formula is denoted as: $\chi^2 = \sum (O_i - E_i)^2 / E_i$, where O_i = observed value and E_i = expected value. Furthermore, the correlation of the significant value can be calculated based on χ^2 value referring to the chi-square distribution table. If the signed value is smaller than a crisis points, that is 0.05, then the feature has a strong relevance in data, in other words, the feature is an important feature
MRMR ranking	The minimum redundancy maximum relevance (MRMR) algorithm is used to identify features that are ranked using the Laplacian scores for determining the feature importance. In this process, at every stage, the feature having the maximum feature importance score will be included to the selected feature set represented by S. Considering a total of m features, assuming a feature represented as X_i ($i \in \{1, 2, \dots, m\}$) can be stated as: $f^{\text{mRMR}}(X_i) = I(Y, X_i) - \frac{1}{ S } \sum_{X_s \in S} I(X_s, X_i)$ where Y is response variable, S denotes the set of selected features, $ S $ represents the size of feature set, and X_i denotes a feature that has presently not been chosen
RRelieff algorithm	The algorithm is used to rank the importance of predictors by determining weights of the feature by increasing the prediction accuracy with respect to the pairwise distance. RRelieff algorithms are best suited for assessing the quality of attributes, considering the performance as well as how well the element is able to differentiate among occurrences in close proximity to each other $W[A] := W[A] - \text{diff}(A, R_i, H)/m + \text{diff}(A, R, M)/m;$ Initially, the weight of each of the attributes is configured to zero. Random instance, represented as R_i is designated and its two nearest neighbors are then established. The neighbors thus selected are taken as, one belonging to the same class of random instance, R_i and the other is picked from a distinct class known as nearest miss M

algorithm is known for its high speed and effective capability for network optimization. The system formulation has been done by using a feed-forward network with. The neural network undergoes the training process using back-propagation training approach. The neural network architecture view is displayed in Fig. 2. ANN has been used by multiple researchers in past for proposing solution to problems that requires pattern recognition similar to traffic sign detection, that also necessitates model to learn from historical image pattern [20, 23, 24]. The execution output of the purported methodology is assessed using ANN-based and by multiple ranking algorithms.

Authors have assessed the analytical execution of their recommended approach by utilizing ANN and a subgroup of best-suited ML-based algorithms to arrive at the experimental results and comparative outcome based on metrics of R-squared and MSE values. Authors have utilized the MATLAB software package for applying

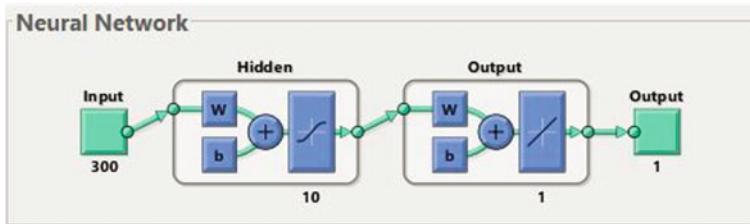


Fig. 2 The architecture view of ANN

various ML algorithms for ranking and classification. MSE has been used to measure the performance of the network with respect to the average of the squares of the “errors”.

Figure 3 illustrates the outcome of measurement metrics using ANN with the chi-square test. Figure 4 shows the experimental results outcome while using ANN with MRRMR test, and Fig. 5 depicts the accuracy results using ANN with RRelief test. It is observed that implementation of ranking algorithms with ANN demonstrates that chi-square test has the best results. The ANN (LM Algorithm) **with chi-square ranking** classification model, has demonstrated the best experimental results outcome having the R-squared value of 0.90, while **MRRMR ranking classification** model, has value of 0.78 and **RRelief ranking classification** model has value of 0.82. Further, the chi-square test for ranking was executed along with multiple machine classifier algorithms for performing comparative analysis.

4 Experimental Results Using ML Classifier Algorithms

Authors used Microsoft Windows machine configured on Azure cloud for setting up the infrastructure to run the set up for executing ANN and ML-based algorithms. The objective was to make the necessary computer power available for algorithm processing on a large dataset. The same algorithms can also be executed by researchers in the laboratory, though it may take more time depending on the configuration of the hardware used. In the proposed approach, different techniques/algorithms were used to get the dimensionality reduction of the Mapillary traffic image dataset, reduction in computation time, and overall improvement in the classification accuracy. Since, as observed in earlier experiments, chi-square outperforms as ranking algorithm, this result has further led to further analysis of combining chi-square with several other best-suited machine learning algorithms considered as classifier for detecting traffic sign images from Mapillary public dataset.

Authors compared the results using multiple classification algorithms combined with the ranking algorithms as evident from Fig. 6. Experimental results of combining chi-square ranking algorithm with multiple ML algorithms has found that SVM Cubic algorithm-based approach had the maximum R-squared value of 0.82. It is further

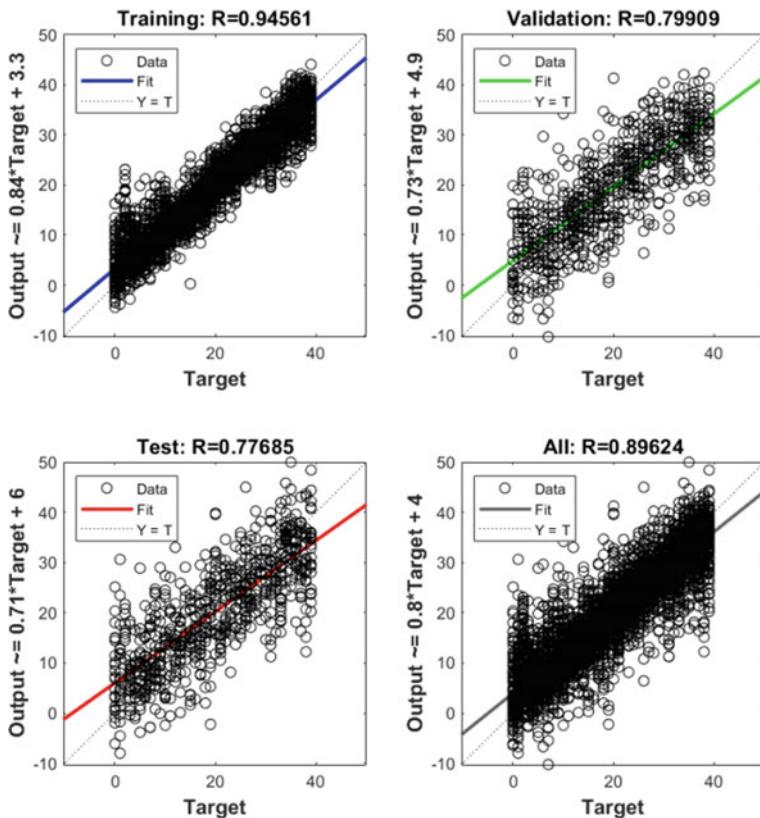


Fig. 3 NN with chi-square results

observed that the best value of MSE has been identified as 24.78, which has been achieved using the same SVM Cubic algorithm. The results of our proposed approach are better than that of Hidalgo et al. [9] and Ayachi et al. [2], who also used neural network-based approach for achieving dimensionality reduction.

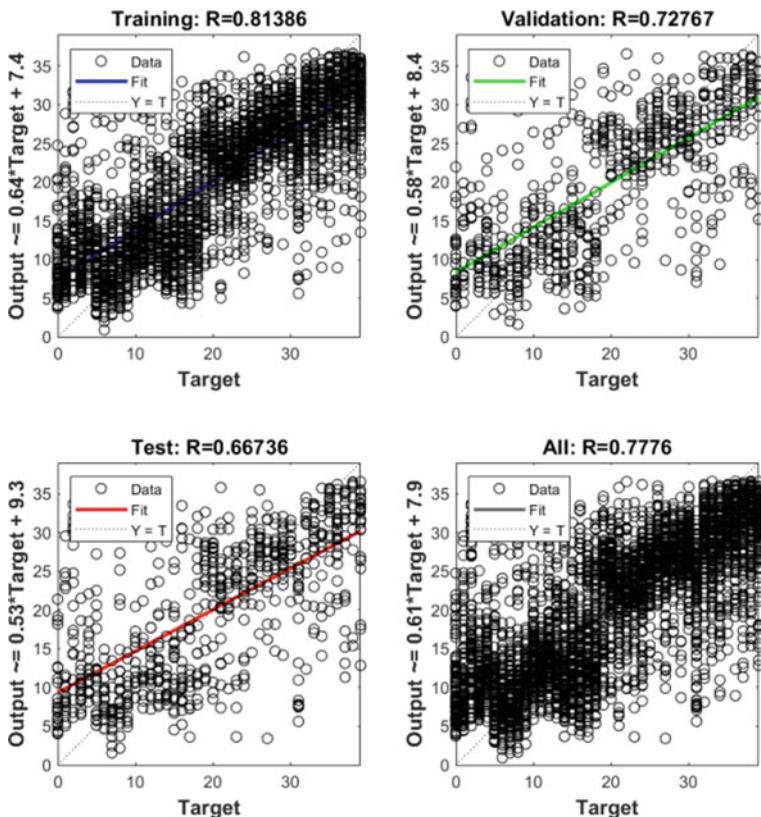


Fig. 4 NN with MRMR results

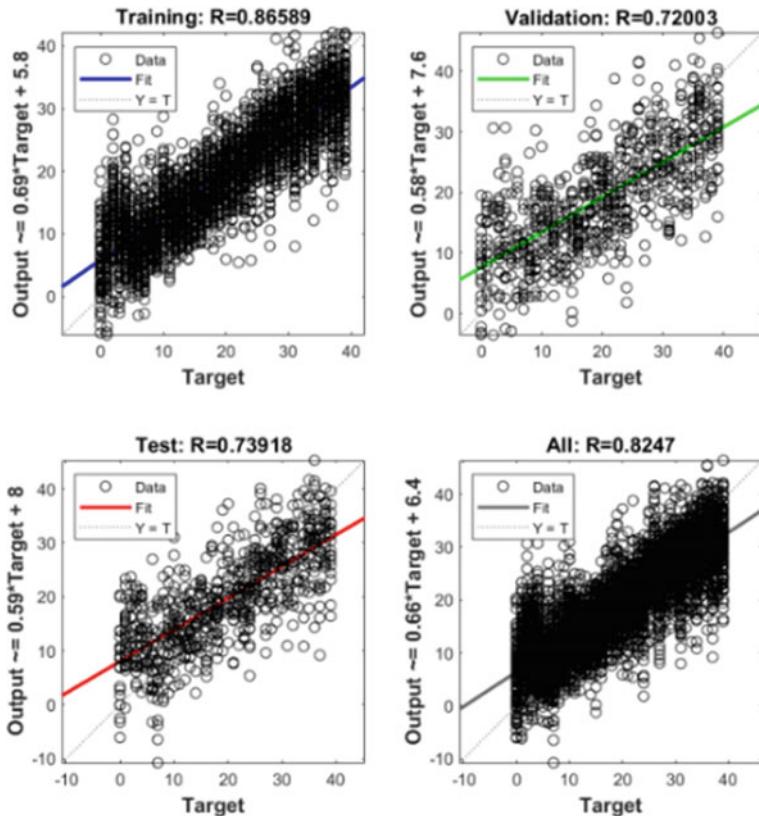


Fig. 5 NN with RRelief results

S.No	Model Name	Type	Ranking	RMSE	R-Squared	MSE	MAE
1.1	Neural Network	Levenberg–Marquardt algorithm	Chi-Square	0.90			
1.2	Neural Network	Levenberg–Marquardt algorithm	Minimum Redundancy Maximum Relevance	0.78			
1.3	Neural Network	Levenberg–Marquardt algorithm	Rrelief	0.82			
1.11	Linear Regression	Linear	Chi-Square	8.20	0.49	67.18	6.34
1.2	Linear Regression	Interactions Linear	Chi-Square	Failed			
1.3	Linear Regression	Robust Linear	Chi-Square	8.28	0.48	67.48	6.29
1.4	Stepwise Linear Regression	Stepwise Linear	Chi-Square	Failed			
1.5	Tree	Fine Tree	Chi-Square	8.79	0.42	77.26	4.29
1.6	Tree	Medium Tree	Chi-Square	8.77	0.42	76.86	4.98
1.7	Tree	Coarse Tree	Chi-Square	8.85	0.41	78.30	5.77
1.8	SVM	Linear	Chi-Square	8.64	0.44	74.57	6.31
1.9	SVM	Quadratic	Chi-Square	5.48	0.77	29.98	3.76
1.10	SVM	Cubic	Chi-Square	4.94	0.82	24.37	3.54
1.11	SVM	Fine Gaussian	Chi-Square	10.79	0.12	116.34	9.28
1.12	SVM	Medium Gaussian	Chi-Square	4.98	0.81	24.78	3.07
1.13	SVM	Coarse Gaussian	Chi-Square	8.95	0.40	80.13	6.18
1.14	Ensemble	Boosted Trees	Chi-Square	7.79	0.54	60.75	5.93
1.15	Ensemble	Bagged Trees	Chi-Square	6.46	0.69	41.69	4.49
1.16	Gaussian Process Regression	Squared Exponential GPR	Chi-Square	11.52	0.00	132.74	9.98
1.17	Gaussian Process Regression	Matern 5/2 GPR	Chi-Square	Failed			
1.18	Gaussian Process Regression	Exponential GPR	Chi-Square	22.73	-2.88	516.51	5.31
1.19	Gaussian Process Regression	Rational Quadratic GPR	Chi-Square	Failed			

Fig. 6 Experimental results

5 Conclusion

In this work, we proposed an approach for effective implementation of using dimensionality reduction approach for feature selection applied to traffic sign images taken from Mapillary public dataset. The primary contribution is the using ranking algorithms combined with ANN and multiple classification algorithms. The comparative studies showed that our proposed approach has given extremely optimistic experimental results. With respect to future research, there is a scope for extending the proposed approach by using other public datasets and combination of Neuro-Fuzzy methods.

References

1. Abd-Alsabour N (2018) On the role of dimensionality reduction. *J Comput* 13(5):571–579
2. Ayachi R, Afif M, Said Y, Atri M (2020) Traffic signs detection for real-world application of an advanced driving assisting system using deep learning. *Neural Process Lett* 51(1):837–851
3. Ayesha S, Hanif MK, Talib R (2020) Overview and comparative study of dimensionality reduction techniques for high dimensional data. *Inf Fusion* 59:44–58
4. Bhatt DP, Tiwari M (2019) Smart traffic sign boards (STSB) for smart cities. In: 2nd Smart Cities Symposium (SCS 2019), pp 1–4, March. IET
5. Chao G, Luo Y, Ding W (2019) Recent advances in supervised dimension reduction: a survey. *Mach Learn Knowl Extr* 1(1):341–358
6. Cheng Z, Lu Z (2018) A novel efficient feature dimensionality reduction method and its application in engineering. *Complexity*
7. Chetana VL, Kolisetty SS, Amogh K (2020) A short survey of dimensionality reduction techniques. In: Recent advances in computer based systems, processes and applications (pp 3–14). CRC Press
8. Gui J, Sun Z, Ji S, Tao D, Tan T (2016) Feature selection based on structured sparsity: a comprehensive study. *IEEE Trans Neural Netw Learn Syst* 28(7):1490–1507
9. Hidalgo DR, Cortés BB, Bravo EC (2021) Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps. *Inf Process Agric* 8(2):310–327
10. Javanmardi M, Song Z, Qi X (2021) A fusion approach to detect traffic signs using registered color images and noisy airborne LiDAR data. *Appl Sci* 11(1):309
11. Khalid S, Khalil T, Nasreen S (2014) A survey of feature selection and feature extraction techniques in machine learning. In: 2014 science and information conference (pp 372–378). IEEE
12. Krishan TT, Alkhwaldeh RS, Al-Hadid I, Al Azawi R, Al-Sharaeh SH (2020) The impact of smart traffic sensing on strategic planning for sustainable smart cities. In: sustainable development and social responsibility—Volume 2 (pp 25–31). Springer, Cham
13. Lira ER, Fynn E, Coelho PR, Faina LF, Camargos L, Villaça RS, Pasquini R (2016). An architecture for traffic sign management in smart cities. In: 2016 IEEE 30th international conference on advanced information networking and applications (AINA) (pp 580–587). IEEE
14. Liu H (2020) Key issues of smart cities. In: Smart cities: big data prediction methods and applications (pp 3–24). Springer, Singapore
15. Ma D, Fan H, Li W, Ding X (2019) The state of mapillary: an exploratory analysis. *ISPRS Int J Geo-Inf* 9(1):10. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/ijgi9010010>
16. Migenda N, Möller R, Schenck W (2021) Adaptive dimensionality reduction for neural network-based online principal component analysis. *PLoS one* 16(3):e0248896

17. Niu T, Wang J, Lu H, Yang W, Du P (2020) Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Syst Appl* 148:113237
18. Sharma N, Saroha K (2015) Study of dimension reduction methodologies in data mining. In: International conference on computing, communication & automation (pp 133–137). IEEE
19. Suryakumar D, Sung AH, Liu Q (2012) Dependence of critical dimension on learning machines and ranking methods. In: 2012 IEEE 13th International conference on information reuse & integration (IRI) (pp 738–739). IEEE
20. Vashisht V, Lal M, Sureshchandar GS (2016) Defect prediction framework using neural networks for software enhancement projects. *J Adv Math Comput Sci*, 1–12
21. Vashisht M, Kumar B (2021) Effective implementation of machine learning algorithms using 3D colour texture feature for traffic sign detection for smart cities. *Expert Syst* e12781
22. Vashisht M, Kumar B (2021) Traffic sign recognition using multi-layer color texture and shape feature based on neural network classifier. In: Micro-electronics and telecommunication engineering (pp 479–487). Springer, Singapore
23. Vashisht V, Kamya S, Vashisht M (2020) Defect prediction framework using neural networks for business intelligence technology-based projects. In: 2020 international conference on computer science, engineering and applications (ICCSEA) (pp 1–5). IEEE
24. Vashisht V, Lal M, Sureshchandar GS (2015) A framework for software defect prediction using neural networks. *J Softw Eng Appl* 8(08):384
25. Wang Y, Zhou C (2020) Feature selection method based on chi-square test and minimum redundancy. In: International Conference on Intelligent and Interactive Systems and Applications (pp 171–178). Springer, Cham
26. Zebari R, Abdulazeez A, Zeebaree D, Zebari D, Saeed J (2020) A comprehensive review of dimensionality reduction techniques for feature selection

Detection of Fake Reviews in Yelp Dataset Using Machine Learning and Chain Classifier Approach



Lina Shugaa Abdulzahra and Ahmed J. Obaid

Abstract Over the past few years, e-commerce has led to a significant shift in business activities from traditional methods to online platforms. Nowadays, consumers heavily rely on online reviews to guide their purchasing decisions, prompting businesses to adapt to this new reality. However, fake reviews pose a critical challenge in online reviews. Fake reviews can have serious consequences, including misleading customers and damaging the reputation of organizations. To tackle this issue, various approaches, such as natural language processing, machine learning, and sentiment analysis, have been proposed as potential solutions for detecting fake reviews. These strategies typically involve analyzing the content of reviews along with associated metadata, such as the language used, review timing, and ratings. However, differentiating between fake and genuine reviews can be challenging, as fake reviewers often employ tactics to make their reviews appear more legitimate. Despite the complexities involved, significant progress has been made in developing effective strategies for detecting fake reviews. These techniques play a crucial role in ensuring that consumers can make informed decisions based on trustworthy information while safeguarding online review systems' integrity. The main focus of this paper is to combine textual elements with other related behavioral parameters, which leads to a higher rate of perception and detection compared to other existing methods. By incorporating new behavioral variables, the proposed model enhances the accuracy of detecting fake reviews. The Elmo Model is utilized for encoding result vectors and reducing computational overhead, while the VADER model helps to determine the polarity of review text, enabling individual reviews to be evaluated. To classify real-time reviews in the Yelp dataset, a stack model is constructed using the Multinomial Naive Bayes method (MNA) and the Gradient Boosting Classifier. Compared to similar studies, the proposed model demonstrates exceptional accuracy, achieving an AUC of 82% and overall accuracy of around 98%. Overall, the integration of textual components with behavioral parameters in the proposed model offers a promising

L. S. Abdulzahra · A. J. Obaid (✉)

Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

e-mail: ahmedj.aljanaby@uokufa.edu.iq

L. S. Abdulzahra

e-mail: linas.alabdali@student.uokufa.edu.iq

approach to effectively detect fake reviews and improve the authenticity of online review systems.

Keywords Fake reviews' detection · Yelp dataset · Fake detection · Fake user behavior · Reviews metrics analysis

1 Introduction

In the modern era, e-commerce has revolutionized the business landscape, with a significant shift toward online platforms. This has led to the emergence of electronic purchasing systems, which enable consumers to explore products and services, read user reviews, and make informed decisions. User reviews are critical to the decision-making process, as they provide opinions, experiences, and recommendations about specific products or services [1]. As a result, they can significantly impact the success of an organization. The internet marketing and shopping websites provide people with the ability to buy, sell, and share opinions on a wide range of products and services, including cell phones, laptops, hotels, restaurants, and airline tickets. Users can rate, evaluate, remark, recommend, or suggest products, and their opinions can heavily influence others' purchasing decisions [2]. These reviews can be positive or negative and may significantly impact customer behavior. However, there is a risk of false information being disseminated on behalf of product manufacturers by agents to deter potential purchasers. Identifying fraudulent evaluations is crucial for both new buyers and manufacturers of high-quality products [3]. Therefore, it is essential to have systems in place to detect and address these issues. Overall, the emergence of e-commerce and online platforms has revolutionized how people buy and sell products and services. User reviews have become a crucial aspect of the decision-making process, and ensuring that they are trustworthy and reliable is essential. Detecting inaccurate and fraudulent evaluations automatically is a crucial challenge for online review systems [4]. This issue, also known as "fake review detection," is essential for numerous finance, healthcare, security, and review management applications. However, conventional detection methods fail when fraudsters use uncomplicated concealment techniques. For example, spammers may publish genuine reviews alongside fake ones, making it difficult to distinguish between them. As a result, it is crucial to develop techniques that can accurately identify fake reviews [5, 6].

This paper explores the different categories of false reviews and the contextual and behavioral aspects of reviews. It also addresses the challenges of detecting false opinions and reviews, recent research, problem statements, and study objectives. In recent years, several techniques have been developed for identifying fake reviews. However, these techniques may not be effective when fraudsters use advanced techniques to mask fraudulent behavior. For instance, a fake review may go unnoticed if the account behind it behaves like a regular user. Therefore, it is essential to continue developing new techniques to address this issue and ensure that online review systems remain trustworthy and reliable.

2 Fake Reviews' Detection Challenges

Fraud detection is a crucial task because the number of fraudsters (spammers/fake reviewers) is typically much smaller than the number of legitimate users. However, out of millions or billions of reviews or comments, it will be a difficult process to identify fraudulent reviews or remarks. Most well-known online service providers for example (Amazon, Yelp, Alibaba, Netflix, etc.) attempt to increase their productivity by preventing fake reviews from affecting their services [7]. However, to address the most well-known challenges that refer to the task of fake reviews' detection:

- **Lined Reviews:** These reviews, also known as (camouflaged Reviews), are made when spammers pose as normal users by posting legitimate reviews mixed with fraudulent ones [8].
- **Redundant Account:** When spammers use origin accounts to post false reviews, it will display ties between fake comments and those accounts [9].
- **Fraudsters Trading:** This challenge is also called lack of necessary information, and it happens when fraudsters (spammers) will avoid trading with each other in order to be not detected [10].
- **Noise Identity:** When fraudsters (spammers) fabricate noise information about them (such as personal photo, name, affiliation) to make them hard to identify them [11].
- **Lack of Traceability:** This challenge pretends that when users do not disclose correct or sufficient information about themselves, they limit their access from other businesses and users [12].

Many research works have been presented to detect fake reviews by taking individual criteria related to user attributes and object characteristics into account. Unfortunately, most of the well-known traditional methods fail to distinguish between genuine and spammer reviewers. However, this chapter chose one of the unresolved problems in the Yelp dataset (Yelpchi) and evaluated it before recommending a solution based on our practical method.

3 Fake Reviews' Categorization

The task of categorizing a set of reviews from a pre-defined number of users, groups, voting, rating, or labeling based on specific attributes is a highly tough procedure because we cannot recognize fake reviews by reading a large number of online reviews. Generally, fake (spam) reviews can be categorized based on their actions into three types [10]:

- **Untruthful Reviews:** Fake reviews can raise or lessen the overall suspicion of fraud associated with a particular target product. There is the potential for false content reviews to make claims about various product attributes.

- **Fake Promotions:** The act of posting false reviews with the intent to anger sellers, makers, or distributors of a product.
- **Irrelevant Contents:** Content that has not received any comments, with the exception of reviews that are unrelated to the subject they were posted about. Other sorts of material include possible inquiries and responses as well as advertising.

Fake reviews can also be categorized based on their “polarity,” which is the proportion of reviews that are either “positive” or “negative” on the content scale. Positive reviews typically highlight features that customers appreciate, and negative reviews describe features that users dislike [11].

Finally, based on extracted features, there are two types of fake reviews [8, 12]:

- **Behavioral Features:** are sometimes called “non-verbal features.” These features show how each reviewer and review are different. Real-world data, on the other hand, have natural graph structures. Objects (nodes) represent users or customers; connections (edges) between them are encoded as a graph G.
- **Contextual Features:** Verbal features are taken from review-centered features and show how to review material can be seen from different points of view. Text mining algorithms can be used to draw out this information.

4 Why We Chose the Yelp Dataset

The Yelp dataset, available through the Yelp dataset challenge, is a publicly accessible dataset that presents a significant challenge for detecting fake reviews using new methodologies due to its regular updates [4]. We utilized a specific version of this dataset obtained from the Yelp Dataset Challenge webpage. The Yelp dataset comprises a wealth of information on reviews, users, businesses, and check-ins. Our research focused on the review data, which consists of 1,125,458 user reviews spanning various business categories. This dataset has also been widely used as a benchmark for state-of-the-art challenges.

The Yelp dataset includes multiple versions, with the primary task in the Yelp-Fraud Dataset being a binary classification problem that involves distinguishing between suggested and filtered reviews. In addition, we categorized users into two distinct groups: spammers, users with fabricated reviews that have been filtered, and benign users, whose reviews have not been filtered. Previous research has indicated that review manipulators leave traces in various aspects, such as user behavior, product associations, review text, and timing. Fraudulent reviews often specific products and show distinctive patterns in the review text. To analyze the Yelp dataset, we represent users as nodes in a graph and employ text mining techniques to determine whether the reviews are legitimate or spam.

This alternative representation of the Yelp dataset, where users are treated as graph nodes and their hidden connections as edges, presents an opportunity to address a real-time dataset challenge and contribute to research on an unresolved problem. In this thesis, we specifically utilized the Yelpchi dataset, a current challenge, which

focuses on reviews from hotels and restaurants in Chicago, USA [4]. By tackling this real-world problem, our research aims to generate novel insights and contribute to the advancement of knowledge in this field.

5 The Problem of Research Work

The prevalence of web platforms is rapidly growing, extending to all types of businesses, with buyers and sellers shifting toward the e-market. The abundance of e-commerce websites has broadened users' product selection options, expanding the merchant/seller business on these platforms. Prior to purchasing everyday products, it is common for individuals to seek advice from their friends and family. However, millions of reviews, opinions, and suggestions from users worldwide are available online. As a result, a customer's decision-making process is heavily influenced by user reviews, which can impact the financial profits or losses of the product provider. A review is always honest in the eyes of the purchaser, but the buyer is unaware that reviews can be fake [4, 5, 8, 13, 14].

Yelp dataset considers one of the best practical datasets for research works, as it is collected from the Yelp.com website from actual users in some business categories. The spammer (user)/fake (fraud) detection task has been created as a benchmark on the State-of-the-Art website.

6 Related Work

Several studies have been conducted to address the issue of identifying fake (spam) reviews within the Yelp dataset. These research efforts have explored various approaches to detect fraudulent reviews. Some of these studies have focused on uncovering connections between users to identify groups of spammers based on their behaviors. Others have delved into the analysis of comments, aiming to identify deceptive words hidden within the reviews. This section will discuss the most noteworthy research endeavors undertaken to detect fake reviews.

Carbon et al. [12] use K-means clustering and sentiment analysis to improve feature selection. They use classification techniques to identify fraudulent comments, including Logistic Regression, Support Vector Machine (SVM), Naive Bayes, Random Forest, Decision Tree, and Gaussian Discriminant Analysis (GDA). The authors found that the sentiment classifier might improve feature selection accuracy and sentiment analysis would improve prediction accuracy, with an accuracy AUC of 45.00 [12].

Xu et al. [13], first preprocess the review text; then, they experiment with different feature selection algorithms to pick the best features; and finally, they use a supervised learning algorithm (Naive Bayes) to extract the contextual meaning of each review text according to pre-defined categories. This identifies contextual meaning from

textual characteristics and combines it with the review rating (Star Rating) to extract positive and negative user comments. Finally, they determine that a higher suggested rating will make individuals feel better, yet the reverse will happen. Their 10,000-sample Yelp dataset approach has an accuracy AUC of 58.89 [15].

Wang and Qiu [16] aimed to enhance the prediction of user ratings based solely on review text. They utilized a Logistic Regression (LR) Model, incorporating sentiment analysis and Part of Speech Recognition (POS) techniques to identify the keywords in each review comment. To configure these important words, they employed n-gram word granularity features (Unigram, Bigram, and Trigram). Additionally, they used TF-IDF to extract the most frequent words in the review comments [16].

Shebuti and Akoglu [17], users, products, and reviews are nodes in a G graph network with two types of edges. User review and review product edges are the two types of edges. Their network detects opinion fakes by classifying each node (user, review product) as benign or fake, targeted or non-targeted, and genuine or fake. The active interface procedure determined the class of opinion by picking nodes and computing their Weighted Uncertainty Score (WUS) and Probability of Reaching. They used Expected Uncertainty Reach (EUCR) to determine WUS by finding the nodes with the greatest score at each iteration. When the number of nodes (users) increased, the prediction of bogus opinions decreased, and they attained an accuracy of around AUC 60.00 by selecting review levels from 100 to 1000 [17].

Wu [13], authors applied feature objects to categorize users as spam or legitimate. The main goal is to calculate the normal distribution of user reviews using Gaussian Distribution (GD) to classify them as spam or genuine and then identify the relationship between each user and these categories. The authors also established Reliable Fake Review Detection (RFRD) to determine object review distributions and user dependability using Yelp's online review data. Expectation Maximization (EM)-based learning identifies users based on their review distribution with an Accuracy AUC 75.60 [13].

Tay et al. [18] developed a modified Neural Network (NN) named Neural Network-based Multi-pointer Learning (MPCN). They encode user groups and review group inputs for a business category at six levels. Next, embedded vectors will link to the most relevant reviews' category for each user [18].

Sihombing and Fong [19], this paper used the Gradient Boosting-based XGBoost Classification Framework1. Fast, versatile, and portable distributed gradient-boosting library XGBoost [20]. They want to verify Yelp reviews. They used Naive Bayes, SVM, and Logistic Regression to identify user reviews and attained an average accuracy AUC of 78.00 [19].

Dou et al. [11] investigate fraudster camouflage using user connections. They consider CAmouflage-REsistant GNN (CARE-GNN), a modified Graph Neural Network (GNN). Their model starts with a label-aware similarity metric to find meaningful neighboring nodes. Influence reinforcement learning (RL) determines the ideal number of neighbors. Finally, neighbors from different relationships are grouped. Feature-based spam detectors cannot find the scammers' special characters. Their proposed model had an accuracy of 74.00 [11].

Chen and Xia [21] utilized the Restaurant Reviews from the Yelp dataset to estimate the rating of each restaurant, providing insights into their services and quality. They aimed to predict review star ratings on a scale from 0 to 5 by considering both textual and non-textual elements. Regression Model, Naive Bayes, Decision Tree, and Neural Network were employed to determine the most accurate rating prediction. Among these models, Decision Tree yielded the highest accuracy with an AUC of 82.50. It is crucial to address fake (spam) reviews as they can significantly impact the ranking of products and influence customers' perceptions and decisions [21].

Lim et al. [22] introduced Non-Homophilous Graphs (NHGs) as a framework for evaluating graph machine learning and capturing the complexity of label-topology relationships. This model presented a novel method for quantifying the presence or absence of homophily, which proved to be more effective in certain scenarios. A fully homophilous network exclusively connects nodes of the same class. The proposed approach achieved an accuracy of 90.04%, establishing it as a benchmark in the State of the Art [22].

Liu et al. [10] introduced a novel model called Pick and Choose Graph Neural Network (PC-GNN) for addressing imbalanced supervised learning on graphs. The authors achieved an AUC of 77.65, demonstrating the effectiveness of their proposed model [10].

Peng et al. [23], RioGNN, a Reinforced, Recursive, and Flexible multi-relational Graph Neural Network (GNN) architecture, handles Neural Network complexity while keeping relation-dependent representations. RioGNN creates a multi-relational graph with varied nodes, edges, properties, and labels based on the situation. RioGNN solves multi-relational graph problems in a novel and adaptable way. Using a relation-aware neighbor selection approach, they obtained the final node embedding with an AUC of 83.54 [23]. The following Table 1 summarizes the related research works that were achieved on the dataset.

7 Proposed System

The approach that has been presented is comprised five steps, each involving conducting an analysis of sentiment and extracting textual elements. These attributes are then integrated with behavioral features depending on the comments and activities previously generated by users. The primary goal of this combination is to identify comments or reviews that may convey masked ideas or feelings about the product. Determining how a person feels about a piece of writing is frequently utilized; nonetheless, we believe that the proposed system will benefit from this tweak and produce better results. In addition to this, well-known classifiers such as Naive Bayes and Gradient Boosting Classifier have been incorporated into the system. The following Fig. 1 illustrates the proposed system's architecture and the procedures that will be taken to process data.

Table 1 Related work of fake reviews' detection in Yelpchi dataset

T	Ref. no.	Year of study	Model/method name	Dataset version	Accuracy result	
					ACC	AUC
1	[12]	2014	SVM	Yelp Phoenix		0.45
2	[15]	2015	Naïve Bayes	Yelp challenge		0.58
3	[16]	2015	Logistic Regression	Yelp Charlot		~
4	[17]	2016	WUS	Yelpchi/NYC		0.60
5	[13]	2017	RFRD	Yelp Phoenix Yelp Charlot		0.78 0.82
6	[18]	2018	MPCN	Yelp 17		0.43
7	[19]	2019	XGBoost	Yelpchi		0.78
8	[11]	2020	CARE-GNN	Yelpchi / NYC		0.72
9	[24]	2023	Decision Tree	Yelp challenge		0.82
10	[21]	2021	NHG	Yelpchi	0.90	–
11	[10]	2021	PC-GNN	Yelpchi		0.79
12	[23]	2021	RioGNN	Yelpchi	0.83	–
13	[3]	2023	Bilot	Yelpchi		0.93
14	[4]	2013	SVM	Yelpchi	0.85	–
15	[5]	2016	XGBoost	Yelp	0.95	
16	Proposed method	2023	MNB + GBC	Yelpchi	0.98	0.82

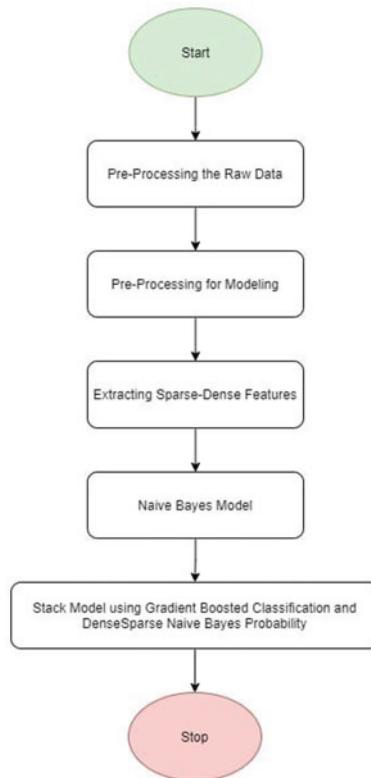
8 Experimental Results

In the experimental part, the proposed system was evaluated using a comprehensive dataset consisting of various user reviews and feedbacks. The dataset was carefully selected to represent various products and services. The system's performance was assessed based on several measures of accuracy, precision, recall, and F_1 -score.

Table 2 presents the statistics of the dataset used in the experiments. It provides information on the total number of reviews, the number of positive and negative reviews, and the distribution of reviews across different categories. The dataset showcased a balanced distribution of positive and negative reviews, allowing for a fair evaluation of the system's effectiveness in handling both types of feedback.

8.1 Preprocessing Raw Data

Preprocessing raw data is crucial to any data analysis or machine learning process. Preprocessing is the process of preparing data for analysis or use in a machine learning model. But reading the comments table from the Yelp dataset comes first. The review

**Fig. 1** Flowchart of proposed system**Table 2** Overall statistics of Yelp_{chi} dataset

Metric name	Restaurants	Hotels
No. of restaurants	242,652	283,086
No. of reviews	788,471	688,336
No. of reviewers	16,941	5123
Size in MB	953	914
No. of fake reviews (Y)	8303	781
No. of non-fake reviews (N)	58,716	5078
No. of filtered reviews (NR) ^a	402,774	415,707
No. of non-filtered reviews ^b (YR)	318,678	266,764
No. of reviews considered for (training/testing/validation)	67,019	5859
No. of reviews need to be filtered, analyzed, and classified	721,452	682,471

^a**No. of filtered reviews (NR):** the total number of reviews filtered from a company's business Yelp page due to automated review filtering

^b**No. of non-filtered reviews (YR):** the total number of reviews not filtered from a company's business Yelp page due to automated review filtering

Table 3 Content of review table in Yelp dataset

Review table	Feature type	description	Preprocessing action
	Date	Date of the review	The date will be excluded
	ReviewID	Review id in the Database	No preprocess required
	ReviewerID	Reviewer Id who made this review	No preprocess required
	ReviewContent	Textual content of the review	Text features (need to preprocess)
	Rating	Rating of the review made by the user	Behavioral features (already presented in tabular form) and no need for preprocessing
	UsefulCount	Clicks count of useful by others	
	CoolCount	Clicks count of cool by others	
	FunnyCount	Clicks count of funny by others	
	Flagged	Expression status of each review	

table contains both the comments that users have left on the Yelp.com website and the opinions that other users have already seen those comments. This tabular data contain two different types of information that can be used to extract features that later help to recognize fraud (fake reviews). The following Table 3 shows the content of the review table in the Yelp dataset.

8.2 Preprocessing for Modeling

This level involved the polarity calculation of the review text. Text polarity is the degree to which it is positive or negative regarding the reader's emotional response to the subject matter. It may be studied with NLP methods because it is typically stated as positive, negative, or neutral. A sentence with a positive polarity might read something like, "I love football," while one with a negative polarity might read something like, "I don't like Sunday nights." While neutral polarity does not express positivity or negativity, like, "The sky is blue." Text polarity analytics aims to automatically determine the level of positivity or negativity in a given sentence. Many fields, including marketing research, social media monitoring, customer feedback analysis, and political analysis, could benefit from this kind of interpretation of the text tone.

reviewerID	reviewContent	rating	usefulCount	coolCount	funnyCount	flagged	restaurantID	label	doc_embeddings	Polarity
Z944s6lJYowOnB0IA	unlik next wed eaten previou night dish comple...	5	0	0	0	N	pbEixam9YJL3neCYHgwlUA	0	[0.6125606, 0.40937138, 0.99674743, -0.6138403...	0.9912
LC3y-ZvP45e5iiIMtw	probabil one best meal ive ever perform food gr...	5	0	0	0	N	pbEixam9YJL3neCYHgwlUA	0	[1.0684175, 1.4197484, 0.65383476, -0.86704063...	0.8625
AqXZzyhxNpL4M9g	servic impecc experi present cool eat balloon ...	3	2	0	0	N	pbEixam9YJL3neCYHgwlUA	0	[0.7680903, 0.7185146, 0.45236242, -0.5638112...	0.7964
z88ymAzu45skODw	problem place like given exhibit cost media a...	3	8	0	3	N	pbEixam9YJL3neCYHgwlUA	0	[0.9151588, 0.49309808, 0.5408095, -0.38833123...	0.9451
Wihmh3g7k9N2G8A	idea write review dine alinea bring whole diff...	5	1	2	0	N	pbEixam9YJL3neCYHgwlUA	0	[0.7532564, 0.67473817, 0.1330135, -0.29704955...	0.8639
xDTfUqVfhPf9MBQ	despit firstworld tragedi endur effort get res...	5	3	1	1	N	pbEixam9YJL3neCYHgwlUA	0	[0.6754131, 0.34726316, 0.6446862, -0.29560024...	0.9637

Fig. 2 Embedding of preprocessed review text in the review table of Yelp

8.2.1 Extracting Sparse Dense Matrix

Dense features are a type of representation in which each text is mapped to a continuous vector with fewer dimensions. This is often done with word embeddings or other methods of feature engineering. Dense features are good for capturing a word's meaning and how it relates to other words, but they can be harder to understand and may need more computing power. However, dense feature models are a useful tool that may be used for a variety of natural language processing (NLP) applications. These models are able to capture the meaning of natural language text as well as the relationships between individual words. Hence, there are several models used for dense feature generations, such as Word2vector, Glove, BERT, FastText, and ELMo. This paper used Word2vec embedding model to make the text ready for the next models, which includes training models for the classification of the review text. The flowing Fig. 2 shows the embedding of review text in the Yelp review table of (restaurants).

8.3 Classification Model

The proposed system adopts Multinomial Naive Bayes (MNB) algorithm (reviews). MNB is a regularly used variation of the Naive Bayes method for text classification applications; in the context of text classification, the algorithm computes the probability that a review belongs to a certain class (such as spam or non-spam, fraud or illegible) based on the frequency with which each term appears in the review. However, based on the experimental analysis of the proposed system in this paper and the nature of the text in the selected dataset, we have combined the MNB with

another classifier (Gradient Boosting Classifier) in the way of a chain classifier. A chain classifier, a sequential classifier, is a machine learning model that sequentially combines multiple classifiers to make predictions. In a chain classifier, the output of one classifier is used as input for the next classifier in the chain. Gradient-Boosted Classifier (GBC) is a machine learning algorithm used to sort things into different groups. It is a type of learning called “ensemble learning,” which combines several weak classifiers into one strong one.

The first step in implementing a GBC is to choose a base classifier. However, in this paper, the basis classifier is Naive Bayes, and it will then be used as a weak classifier. The ensemble is created after several of these classifiers have been trained using various subsets of the data. Each new classifier is trained to correct the errors caused by the prior ones when the process is applied iteratively. The following Fig. 3 shows the average accuracy of predicting labels for the Yelpchi dataset’s test part using the Chain Classifier Model.

Figure 4 shows the **(0.82)** \pm 0.23 curve for the model at the prediction phase.

	precision	recall	f1-score	support
0	0.99	1.00	0.99	156042
1	0.03	0.00	0.00	1652
accuracy			0.99	157694
macro avg	0.51	0.50	0.50	157694
weighted avg	0.98	0.99	0.98	157694

Fig. 3 Average accuracy of the proposed model

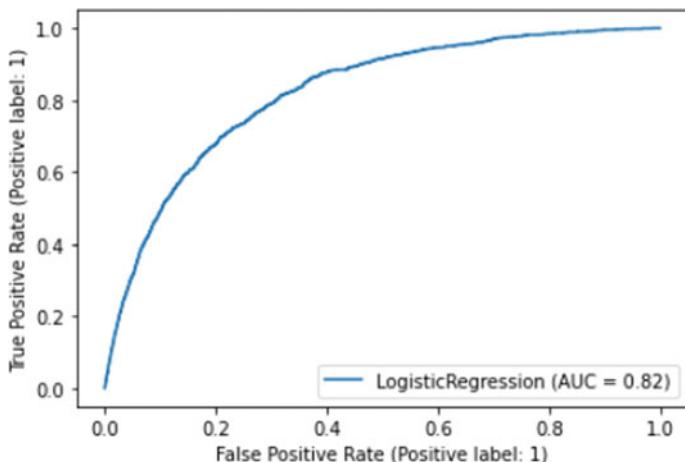


Fig. 4 AUC curve of the proposed model

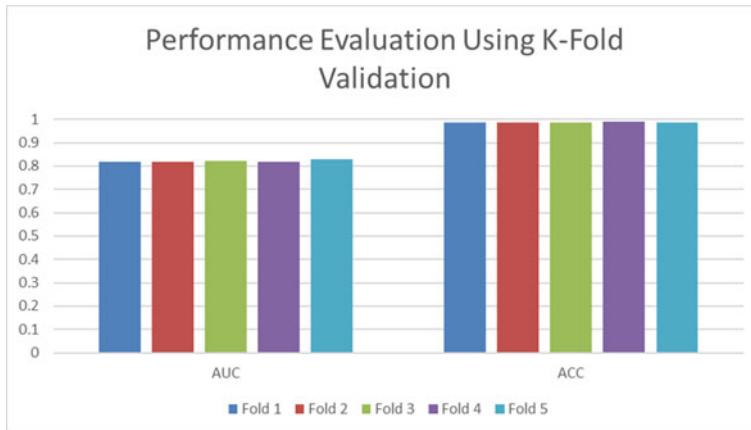


Fig. 5 Accuracy and AUC score of the proposed model when $K = 5$

Cross-validation (CV) is used to evaluate the efficacy of the model. CV tests how well a model developed using machine learning performs with just a tiny amount of data. K-fold cross-validation is a popular method for measuring the efficacy of an ML model. It includes splitting a dataset into k equal-sized folds, with the first time through each fold serving as validation data and the remaining $k-1$ folds serving as training data, and then training and evaluating the model k times; the following Figure 5 shows how we applied K-fold validation to test the proposed system performance, where $K = 5$ and the folded size is 126,155 records as shown below.

The proposed approach for detecting fake reviews significantly improves on earlier solutions. Unlike existing systems that rely on basic rules-oriented heuristics or basic traditional ways, the proposed approach detects fake reviews using powerful machine learning algorithms. The methodology can evaluate a wide range of data points, including review language, reviewer activity, and polarity calculation, to accurately determine the legitimacy of reviews. Furthermore, the suggested system can continuously adapt and increase its accuracy as new fake reviews are encountered, making it a more robust and effective solution than past systems. Overall, the suggested methodology offers a more dependable and effective method of combating fraudulent reviews, which can considerably increase the reliability and trustworthiness of online reviews; the following Figure 6 shows the confusion matrix of testing part of our selected dataset.

Figure 7 shows the classifying of Real-time User reviews from Starbucks New York Page on Yelp.Com Website (Case Study).

Fig. 6 Confusion matrix of testing part in Yelpchi dataset

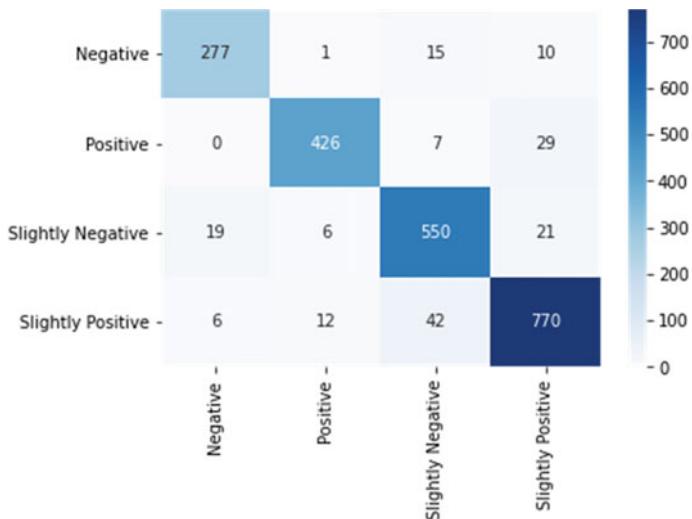
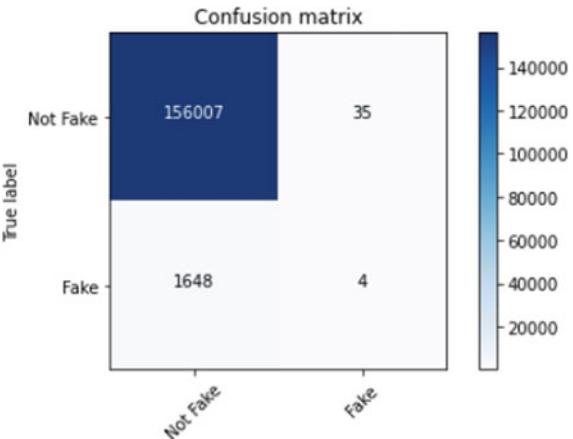


Fig. 7 Confusing matrix of Starbucks review page (case study)

9 Conclusion

Finding fake reviews on Yelp is essential for keeping the site's credibility and ensuring that users can make well-informed choices about the businesses they visit. Yelp has taken steps to deal with this problem, but the fact that fake reviews are still common is still a worry for many companies and customers. Yelp can help to ensure that its platform stays a trusted source of information for users and a valuable tool for businesses trying to get more customers by developing better ways to find and get rid

of fake reviews. However, based on our proposed methodology and research works achieved in this paper, we concluded the following points:

- Considering only the text content in the Yelp reviews dataset could not be sufficient to identify and detect fake reviews as many users could be used normal text in their comments.
- The committee's behavior can be better understood, especially when compared to individuals who appear to be harmful, by analyzing the content of Yelp reviews in real-time.
- The combination of textual content and behavioral features provides better perception and develops the proposed system's performance.
- The newly developed measures of user behavior help to increase the precision for detecting fake reviews.
- Text polarity calculation is one crucial step that we have found to know the text's tone, and it helped to understand the positive or negative presence of user feedback.
- Using the MNB classifier and the GBC classifier provides better performance when compared to the result with other classifiers.
- User activity and its correlation with others can be considered the main features to identify fraudsters and monitor user behavior, especially for users who change their feedback over time.
- Most Fraud Reviews or comments got negative feedback from the committee (Useful Count, Funny Count, Cool Count, etc.), and most fake reviews got zero ratings based on the mentioned metrics.
- Improvements have been shown in the classification of fake reviews after combining "reviewer deviation" with additional contextual and behavioral features.
- The findings indicate that "reviewer variation" is one of the highest essential qualities when ranked according to its significance.

References

1. CaoD et al (2017) Cross-platform app recommendation by jointly modeling ratings and texts. 35(4):1–27
2. Hand DJJDs (2007) Principles of data mining. 30(7):621–622
3. Aljur S, Hiremath E, Patil A, Shivashankar S (2010) Spam detection of customer reviews from web pages. In: Proceedings of the 2nd international conference on IT and business intelligence, pp 1–13
4. Mukherjee A, Venkataraman V, Liu B, Glance N (2013) What yelp fake review filter might be doing? In: Proceedings of the international AAAI conference on web and social media, vol 7, no 1
5. Akoglu L, Chandy R, Faloutsos C (2013) Opinion fraud detection in online reviews by network effects. Proc Int AAAI Conf Web Soc Media 7(1):2–11
6. Hooi B et al (2013) Birdnest: Bayesian inference for ratings-fraud detection. In: Proceedings of the 2016 SIAM international conference on data mining: SIAM, pp 495–503

7. Günnemann S, Günnemann N, Faloutsos C (2014) Detecting anomalies in dynamic rating data: a robust probabilistic model for rating evolution. In: Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining, pp 841–850
8. Hooi B, Song HA, Beutel A, Shah N, Shin K, Faloutsos C (2016) Fraudar: Bounding graph fraud in the face of camouflage. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 895–904
9. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WPJJoir (2002) SMOTE: synthetic minority over-sampling technique. 16:321–357
10. Liu Y et al (2021) Pick and choose: a GNN-based imbalanced learning approach for fraud detection. Proc Web Conf 2021:3168–3177
11. Dou Y, Liu Z, Sun L, Deng Y, Peng H, Yu PS (2020) Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In: Proceedings of the 29th ACM international conference on information & knowledge management, pp 315–324
12. Carbon K, Fujii K, Veerina PJ (2014) Applications of machine learning to predict yelp ratings
13. Wu X, Dong Y, Tao J, Huang C, Chawla NV (2017) Reliable fake review detection via modeling temporal and behavioral patterns. In: 2017 IEEE international conference on big data (big data). IEEE, pp 494–499
14. Zhang D, Zhou L, Kehoe JL, Kilic IYJJo MIS (2016) What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. 33(2):456–481
15. Xu Y, Wu X, Wang QJSU (2015) Sentiment analysis of Yelp's ratings based on text reviews. 17:117–120
16. Wang M, Qiu R (2015) Text mining for yelp dataset challenge. Comput Sci 1–5
17. Rayana S, Akoglu L (2016) Collective opinion spam detection using active inference. In: Proceedings of the 2016 SIAM international conference on data mining: SIAM, pp 630–638
18. Tay Y, Luu AT, Hui SC (2018) Multi-pointer co-attention networks for a recommendation. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, pp 2309–2318
19. Sihombing A, Fong ACM (2019) Fake review detection on Yelp dataset using classification techniques in machine learning. In 2019 International conference on contemporary computing and informatics (IC3I). IEEE, pp 64–68
20. Brownlee JJMLM (2016) A gentle introduction to xgboost for applied machine learning, pp 1–20
21. Chen Y, Xia F (2020) Restaurants' rating prediction using Yelp dataset. In: 2020 IEEE international conference on advances in electrical engineering and computer applications (AEECA). IEEE, pp 113–117
22. Lim D, Li X, Hohne F, Lim S-N (2021) New benchmarks for learning on non-homophilous graphs
23. Peng H, Zhang R, Dou Y, Yang R, Zhang J, Yu PSJAToIS (2021) Reinforced neighborhood selection guided multi-relational graph neural networks 40(4):1–46
24. Mohammed R, Hammoodi AJ, Obaid D, Kumar SR, Sharma G, Jeon RK (2023) Data analytics for smart grids applications—A key to smart city development fake user account detection in online social media networks using machine learning and neural network techniques. Springer Nature, Switzerland Cham, pp 199–215

Data Governance Framework for Industrial Internet of Things



Mohammed Alaa Al-Hamami and Ahmed Alaa Al-Hamami

Abstract Industrial Internet of Things (IIoT) is a major element for industrial systems future. Since IIoT applications are the natural evolution of the IoT, cybersecurity is the main consideration related to IIoT adoption. Because of the similar basic architecture, IIoT inherits some security challenges from IoT. IIoT is not standing alone application. It depends on other applications such as: Cloud network and Big Data. In Cloud, the devices can be connected with each other at anytime from anywhere. Big Data application plays a very important role in storing, analyzing, security, and safety of data. This research aims to design a data governance framework for IIoT. Internet of Things (IoT) governance data security is depended on the good governance of the Cloud Computing and Big Data governance. IIoT security is very important because devices are interconnected, and if there is any hack for one of the devices, it will transfer the problem to the rest of the connected devices. The final design is to use the Encryption algorithm such as Data Encryption Standard (DES) to secure the communications between the devices. Also, the authentication algorithm (two-factor method) has been used to prove the identity of the devices and to mitigate the communication between IIoT devices. Big Data that transferred between devices may be stored or retrieved from the data centers (Cloud) to solve the storage limitation problem of the devices. The results of the practical implementation show an improvement in the data security, integrity, and accuracy. The design proves more safety in dealing with devices and in executing their duties.

Keywords Data governance · Cloud Computing governance · Big Data governance · IoT governance · IIoT governance

M. A. Al-Hamami (✉)
Applied Science University, Eker, Kingdom of Bahrain
e-mail: prof.alaa.alhamami@hiuc.edu.iq

A. A. Al-Hamami
Bedfordshire University, Luton, UK

1 Introduction

Networks have developed in a very large way, and the same applies to communication media, as the other has evolved to allow the transfer of Big Data of multimedia (voice, image, text, video, messages, and others) at high speed while providing the advantage of maintaining its security, integrity, and safety [1, 2]. To satisfy the correctness and safety of Big Data, it needs a governance to organize the search, flow, analyze, and store the data. Nowadays, most service providers adopt Cloud Computing (CC) technology. Moving to Cloud creates new risks and challenges like data location, storage, and processing, which are usually unknown to the user. Big Data provides the utility (Hadoop MapReduce) to manipulate huge data in a secure and correct way according to its governance. Cloud network permits users to access the data storage and using network at anytime from anywhere. IoT introduced the hyperconnectivity concept that makes organizations and individuals communicate from remote locations effortlessly. IoT governance is rely on the governance of Cloud and Big Data. Surely, IoT governance has more specifications and constraints than Cloud and Big Data and it is essential for the IoT.

Industrial Internet of Things (IIoT) is a major element for industrial systems future. IIoT industrial providers have already put security concerns in a high priority. In the recent years, IIoT security has been hot trend in academic research [3].

1.1 Research Aims and Objectives

Networks have developed in a very large way, and the same applies to communication media, as the other has evolved to allow the transfer of Big Data of multimedia (voice, image, text, video, messages, and others) at high speed while providing the advantage of maintaining its security, integrity, and safety [2]. To satisfy the correctness and safety of Big Data, it needs a governance to organize the search, flow, analyze, and store the data. Nowadays, most service providers adopt Cloud Computing (CC) technology. Moving to the Cloud creates new risks and challenges like data location, storage, and processing, which are usually unknown to the user. Big Data provides the utility (Hadoop MapReduce) to manipulate huge data in a secure and correct way according to its governance. Cloud network permits users to access the data storage and use network at anytime from anywhere. IoT introduced the hyperconnectivity concept that makes organizations and individuals communicate from remote locations effortlessly. IoT governance relies on the governance of Cloud and Big Data. Surely, IoT governance has more specifications and constraints than Cloud and Big Data and it is essential for the IoT. The Industrial Internet of Things (IIoT) is a major element for industrial systems future. IIoT industrial providers have already put security concerns in a high priority. In recent years, IIoT security has been hot trend in academic research [4].

1.2 Statement of the Problem and Research Questions

In 2023, the number of IOT devices will become approximately 65 billion devices around the world. These devices accept and generate a lot of information. This research proposes to develop governance for IIoT to solve several problems such as storage limitation, secure communication, and authentication problem and solve one device problem without affecting other devices. During the implementation of the proposed IIoT governance, the following questions emerged [5, 6]:

- Q1. How you treat the storage limitations of the IIoT devices?
- Q2. How can you secure the communications in transferring data?
- Q3. How can you authenticate the IIoT devices?
- Q4. How can you assure the data correctness?
- Q5. Can you make sure for the availability of the Internet protocol number?

1.3 Research Motivation

There are several factors behind the motivation of the IIoT data governance and these are the following:

- There is a high demand on the IIoT devices and huge usage for the increased devices. This will generate a new requirements and challenges.
- Existence of huge numbers for the Internet protocol (IP) offered by the IPv6 which allow a huge number of devices existence.
- The connections between IIoT devices are random and unorganized which will generate a new threat in security and unsafely.
- There is a need for smart cities future development that demands new constraints in the IIoT governance.
- Development of new smart devices that need more constraints to be within the governance.
- The IIoT governance is still immature and needs a lot of attention.

2 Literature Survey

Data governance (DG) “is the process of managing the availability, usability, integrity and security of the data in enterprise systems, based on internal data standards and policies that also control data usage” [1]. When we talk about Big Data governance, we are talking about the management of enormous amounts of data held by an organization and how to use that data in making decisions for the company. Consequently, the purpose of this research is to examine the current Big Data governance frameworks, identify their flaws, and provide a recommendation for a better one. The authors Ali et al. [1] proposed conceptual Big Data governance framework which

consists of eight components. These components are: identify organization structure, identify relevant stakeholders, identify the scope of Big Data, set the policies and standards, optimize and compute, measure and monitor quality, store the data, and communicate and manage the data.

In Ali et al. [7], not only the corporate data integrity and data quality are at risk but also IT professionals are facing problems in the transition from existing datasets to Big Data because of this lack of a Big Data governance framework. The existing data governance faces challenges in the light of Big Data cases. Cloud Computing with its massive scalability, elasticity, multi-tenancy, self-provisioning, and relatively low entity cost via the web is a system of powerful computing capability. As a result, its adoption may be very appealing. It has been suggested by Karkokova and Feuerlicht [8] to establish and continuously enhance CC governance activities using a lifecycle model. The Industrial Internet of Things (IIoT) has risen to prominence as a vital development technology for enhancing the smartness of our professional lives. Large-scale IIoT networks and applications, heterogeneous devices and networks, a vast number of connected devices, dynamic device states and measurements, and the resultant enormous data production are major IIoT features [9, 10].

3 Literature Survey

The goal of this research is to develop a better Industrial Internet of Things data governance framework that associated with Cloud governance, Big Data governance, and Internet of Things governance. As a result, a mix research methodology that used qualitative and quantitative was implemented. The research methodology includes four methods: quantitative, qualitative, mixed methods and data science methodology.

In this research, the research strategy consists of five stages and these are the following:

1. Identifying the problems that face data security and prevent safety in IIoT.
2. Using two research approaches deductive and inductive.
3. Collect data should using methods in order to ensure that the data is collected correctly and presented in a high methodology. The second part was semi-structured interview; this was done through a visit to one of the largest and most modern factories in the region, which specializes in food products and has obtained many quality certificates, in addition to using Internet of Things equipment in quality control of products and production control, where some questions were asked to a group of employees.
4. Understanding and preparing the collected data to be analyzed for identifying and linking the themes and ideas in order to move to the next stage.
5. Proposed method for data governance in Industrial Internet of Thing by integrating the relevant particles and taking the advantage of Big Data, Cloud Computing, and Internet of Thing Governances.

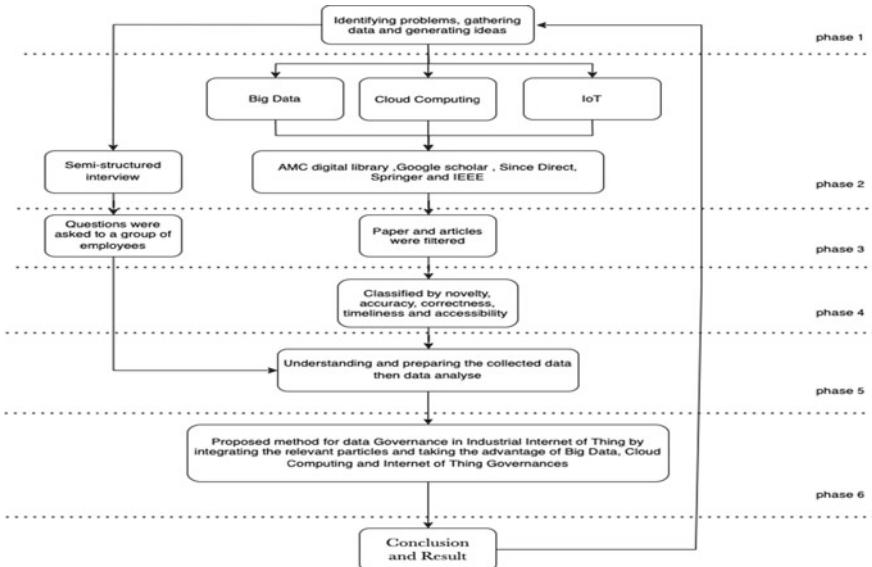


Fig. 1 Phases of the research design

Figure 1 shows the research design phases.

Reports such as Gartner [11] clarified that there is an urgent need for more research and development in this area of governance. Currently, there is no concrete IIoT governance framework, which is available to establish a data governance environment, particularly from a regulatory perspective. In this section, a governance security and safety of Industrial Internet of Things (IIoT) will be developed. This development will depend on the governance of Cloud Computing (CC) and governance of Big Data. Governance of the Cloud will organize the flow of data and communications between the IIoT devices. The interconnection for the IoT is into two sides either between devices only or between human and devices. The governance of Big Data helps in receiving, storing, analyzing, and transferring data. The governance of IoT depends on those two governances (Cloud and Big Data). The governance of IIoT is subclass of the governance of IoT because IIoT inherits some security challenges from IoT. In addition to the requirements needed by the IIoT from the governance of the Cloud, Big Data, and IoT, we have the following limitations for the IIoT such as Resource Limitations, Big Data, Authorization and Access Control, Secure Communication, System Resilience, and Complex Systems.

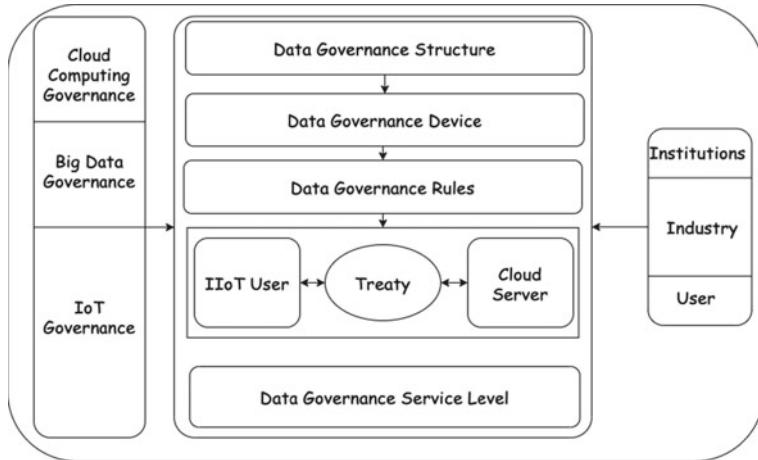


Fig. 2 Structure of data governance framework

4 The Proposed Framework Design for IIoT Data Governance

The IIoT ecosystem includes a variety of types of entities, each of whom has a collective interest in strengthening safety and security. These primary stakeholder groups include the device manufacturers, data center or Cloud service providers, middleware vendors, software vendors, and IT service providers. The proposed governance can offer a global, multidisciplinary view of risk, helping you adapt to changing business requirements, so you feel comfortable in a more connected environment replacing your uncertainty with confidence that you can handle a cyber-attack. Figure 2 shows the structure of data governance framework [7, 12].

5 Method for Designing IIoT Data Governance Framework

To design a framework for data governance in industrial organizations using IIoT, this method proposed six stages that must be followed to achieve an effective design [13, 14].

Stage 1: Create a data governance framework to impose and define roles and responsibilities among data governance teams. This will support IIoT users in ensuring that the necessary roles and responsibilities for data governance are managed across the organization at the effective organizational level. A common three-tiered data governance organization describes a series of senior managers, an intermediate management group, the data governance department, and data governance working committee.

Stage 2: Data governance for IIoT should present threats, vulnerabilities, and possibilities for the industrial organization; also the committee should develop a data governance maturity model at this time.

Stage 3: Establish the data governance rules for IIoT. It refers to primary activities that data governance committees must consider while designing data governance. The data governance rules include various activities such as policies, principles, processes, incident management plans, roles and responsibilities, communication, and a change management plan.

Stage 4: Set up a Treaty for IIoT data governance. It is important for IIoT user to evaluate and inform the IIoT infrastructure provider for their requirements in general and more specifically, for data governance before transferring to industrial IoT environment.

Stage 5: Develop a service level for data governance. The data governance requirements must be met by an acceptable service level and the treaty between the IIoT user and the Cloud server.

Stage 6: Adding and merging the related part of the Big Data, Cloud Computing, and the subclass of IoT governances to the IIoT data governance to increase the security of information, encrypt communication, and increase storage capacity.

Figure 3 shows the six stages.

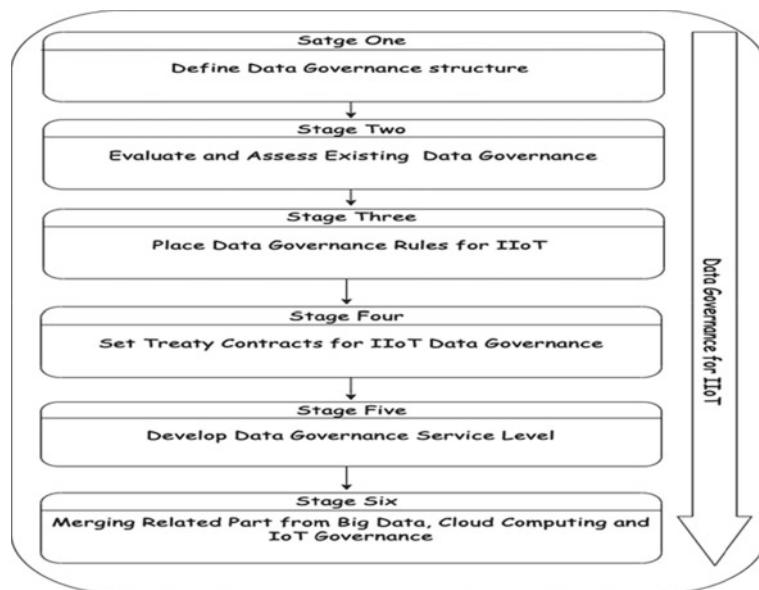


Fig. 3 Six stages of the IIoT data governance framework

6 Results

The responsibility of the Big Data is to store large volume of data, analyzes it, clean it, and prepares it to be useful in taking decision. Big Data will originate a very large volume of data and must be store for a long period; this target will be executed with the help of Cloud Computing by using data center. Internet of Thing (IoT) data governance framework is responsible for the data governance of the IIoT. Internet of Thing is the major class for the IIoT, and for that, IIoT data governance includes the rules and constraints that concern the devices in addition to some rules concerning the devices only. The final data governance framework of the IIoT will contain all the new rules, constraints, specifications, and anything else concerning the IIoT devices. Even this data governance will contain some instructions and rules to train and advice the unexpert users in using the IIoT devices. We believe that the proposed data governance framework for the IIoT will satisfy the requirements of the users and the providers. The governance will answer the research questions as the following:

- Q1. This question concerns the storage limitations on the IIoT devices. Now, the number of the connected devices in IoT will increase rapidly and the evolution of multimedia will collect a huge volume of data. This volume of data needs storage, and due to the limitations of IoT devices, Big Data will handle this problem in storing, analyzing, and transferring the data. These data may be stored for a long time, and in this case, the data center will be used.
- Q2. Securing the communication Cloud will take this responsibility in transferring data between these devices. To secure the communications between the devices, we use method of Encryption, Data Encryption Standard (DES) for example, or we can use another method from the Public Key encryption family.
- Q3. Due to the large number of the connected devices, there is weakness in authentication problem. They suggest using one of the authentication methods such as two-factor method.
- Q4. Problem of data correctness can be handled by the Big Data. It is the responsibility of Big Data to handle this problem and make sure about cleaning of data.
- Q5. IPv6 is the new version that can provide a huge number of IP numbers so whatever the size of the IIoT is big and the number of the connected devices is large, IPv6 can handle this problem. In addition to that, IPv6 provides security (IPsec) for the data.

7 Conclusions and Future Works

The design of Industrial Internet of Thing data governance has been evaluated for its efficiency. We developed the data governance framework successfully after the gaps were covered. These solutions are in successful securing the communication

by using Encryption algorithm and at the same time by using the two factors, three factors, or multifactor authentication was made the authentication process so good. The proposed data governance solved the problem of the limit storage for the IIoT devices and this is by using the Big Data techniques (Hadoop and MapReduce) in reducing the storage space and analyzing data. The IIoT data governance framework has proved its successful work in protecting data and securing the devices in the network. We notice that there are some limitations or drawback in the network due to the weakness of some instructions or rules. It is possible to upgrade the data governance by extending some rules or by expanding some facilities. The following suggestions are to improve the data governance framework for the IIoT:

1. Due to the large increase in data sources, it is getting harder to ensure the security of this enormous volume of data, necessitating the adoption of more security measures.
2. By using artificial intelligence, we can develop a smart data governance to be suitable with the development of the smart home, smart city, and smart world.
3. We can use adaptive security to give the IIoT devices more freedom to dialog with other devices. This is very useful in securing the communication by using different advanced Encryption techniques.
4. Security chain power is in its weakest ring. It is very important to keep other devices working properly when one of them failed.
5. It is very important to think about more advanced design for the data governance framework due to the explosion usage of the devices in an intelligent way. Now most of the devices are independent and can decide for their decisions.

References

1. Al-Badi A, Tarhini A, Islam Khan A (2018) Exploring big data governance frameworks. In: The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018)
2. Hassan WH (2019) Current research on Internet of Things (IoT) security: a survey. Computer Net 148:283–294
3. Karale A (2021) The challenges of IoT addressing security, ethics, privacy, and laws. Internet of Things, journal homepage: www.elsevier.com/locate/iot
4. Raghuvarsh A, Singh UK Internet of Things for smart cities—security issues and challenges. Science direct. Elsevier Ltd. <http://www.elsevier.com/locate/matpr>. Available Online, Accessed on 12 Oct 2021
5. Cheryl B-K, Ng B-K, Wong C-Y (2021) Governing the progress of internet-of-things: ambivalence in the quest of technology exploitation and user rights protection. Technol Soc 64:101–463
6. Andersen DL, Ashbrook CSA, Karlborg NB (2020) Significance of big data analytics and the internet of things (IoT) aspects in industrial development, governance and sustainability. Int J Intell Netw 1:107–111
7. Moghadama RS, Colomo-Palacios R (2018) Information security governance in big data environments: a systematic mapping. In: CENTERES-International Conference on Project MANagement/HCist—International Conference on Health and Social Care Informations systems and technologies, CENTERES/ ProjMAN/HCist

8. Karkoškova S, Feuerlicht G (2016) Cloud computing governance reference model. Springer, Cham, pp 193–203. https://doi.org/10.1007/978-3-319-45321-7_14
9. Lee S, Bae M, Kim H. Future of IOT networks: a survey. *Appl Sci* 7(10):1–25. <https://doi.org/10.3390/app7101072>
10. Carr M, Lesniewska F (2020) Internet of Things, cybersecurity and governing wicked problems: learning from climate change governance. *Int Relat* 34(3):391–412
11. Gartner A (2016) Data risks in the Internet of Things demand extensive information governance
12. Sampath Kumar P, Vijayasree J, Saikumar K, Ayad Alkhafaji M, Obaid AJ and Ali Zearah S (2023) A novel advanced algorithm in automation of food and health monitoring system using IOT. In IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, pp 1–6.<https://doi.org/10.1109/I2CT57861.2023.10126255>
13. Wu M, Li Q, Bilal M, Xu X, Zhang J, Hou J (2021) Multi-label active learning from crowds for secure IIoT, 1570-8705/© 2021 Published by Elsevier B.V.
14. Jacobs N, Edwards P, Cottrill CD, Salt K, Governance and accountability in Internet of Things (IoT) networks. *The Oxford Handbook of Digital Technology and Society*
15. Hassani HL, Bahnsse A, Martin E, Roland C, Bouattane O, Diouri MEM (2021) Vulnerability and security risk assessment in a IIoT environment in compliance with standard IEC 62443. In: The 18th International Conference on Mobile Systems and Pervasive Computing (MobiSPC) August 9–12, 2021, Leuven, Belgium
16. Basukie J, Wang Y, Li S (2020) Big data governance and algorithmic management in sharing economy platforms: a case of ridesharing in emerging markets, 0040-1625/ © 2020 Elsevier Inc.
17. Sun L, Zhang H, Fang C (2021) Data security governance in the era of big data: status, challenges, and prospects, 2666-7649/© 2021 Xi'an Jiaotong University. Publishing services by Elsevier B.V.
18. Mahajan HK (2020) Quantitative research: a successful investigation in natural and social sciences. *J Econ Dev Environ People* 9(4):50–79

IOT-Based Water Level Management System



N. C. A. Boovarahan, S. Lakshmi, K. Umapathy, T. Dinesh Kumar,
M. A. Archana, K. Saraswathi, S. Omkumar, and Ahmed Hussein Alkhayyat

Abstract In countries like India, shortage of water is a serious concern, especially in the southern states such as Tamil Nadu, Kerala and Andhra Pradesh. This shortage of water becomes entangled when there is a loss of water during transmission process. Hence, there must be innovative methodologies for water management and automation for commercial buildings in order to sort these concerns. The proposed system in this paper will execute water management in the form of planning, distributing and maintaining resources of water with IoT. The autonomous water controller includes a microcontroller (ESP8266), relay-controlled motor, a waterproof ultrasonic sensor and a float switch. The water level control system is implemented with a web dashboard arrangement for appropriate display. The presented system will be an important approach for managing water resources both for residential and commercial purposes effectively.

Keywords Internet of Things (IoT) · Water · Management · Microcontroller · Sensor · Motor

1 Introduction

The oceans comprise nearly 95% of the water on earth which is totally not suitable for consumption by human beings. The remaining 1%—fresh water are found in lakes, rivers and streams, good for human consumption, while around 4% is frozen in polar ice caps. The daily consumption of water in India is around 135 l as per statistical

N. C. A. Boovarahan · K. Umapathy (✉) · T. Dinesh Kumar · M. A. Archana · K. Saraswathi · S. Omkumar

SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India
e-mail: umapathykannan@gmail.com

S. Lakshmi
Thirumalai Engineering College, Kanchipuram, Tamil Nadu, India

A. H. Alkhayyat
Scientific Research Centre of the Islamic University, The Islamic University, Najaf, Iraq

reports. This consumption will increase to 40% in the coming days. Hence, it is highly important to safeguard our existing water supplies. In addition to drinking and cleaning purposes, water may be required for commercial and multi-story structures. The water supplies of India are subjected to a lot of stress from a variety of problems. Reusing water, decreasing the consumption of water and storage of water are the various objectives toward sustainability initiatives of water at various places. The paper gives a clear picture how to manage sustainable water in an educational institution and their implementation into architecture of the institution. The system employs a website for the display of water level and a button for controlling the pump for water management and prevention of waste. The autonomous water controller includes a microcontroller (ESP8266), relay-controlled motor, a waterproof ultrasonic sensor and a float switch. The above arrangement will participate in operating the switch between ON and OFF. The relay is controlled by ESP8266 module which also operates the local web server displaying the dashboard. As a consequence, the liquid levels are monitored precisely and the respective tanks are refilled to required levels. By this arrangement, the system will be able to cut down waster and enhance productivity thereby saving money.

2 Literature Survey

Exposure to all processes of farming can be obtained if IoT sensors and devices are installed in the fields appropriately. This will give information on status of crops, quality of soil, conditions of weather, etc. [1, 2]. If the water level reaches the threshold level, a message will be sent to specified mobile phone immediately. The system is seemed to be more effective to reduce the probability of deficit in water supply [3, 4]. The three nodes of operation such as sink, sensor and monitoring are constructed with modules like PH, temperature, etc. The performance of system was evaluated with respect to specific water areas [5, 6]. The system will send alert messages when there is a decline in water quality in pond. The data will be accessed with parameters such as temperature, pH value and dissolved level of oxygen [5, 7]. The development of wireless sensor networks can be exploited in water environments to monitor quality of water [1, 6]. Wireless radio frequency receivers are employed for the design of water management system with the help of a microcontroller [4, 8]. A sensor-based system was proposed for purpose of water management in addition to design of embedded systems. The practical difficulties of implementation are discussed [1, 9–11]. Experimental results show that accuracy and cost of sensor employed in the prototype system are very much comparable to that of any sensor available in the market [11–13]. The system will be able to sense physiochemical parameters and can display those values in proper format after appropriate processing [2, 3, 14, 15]. The immediate requirement for a water quality monitoring system has made more dangerous the present waste water discharge and treatment in case of industries [8, 16]. The wireless sensor network-based water monitoring system is recommended for treatment by which maintenance of water level can be improved

and burden of farmers can be reduced [10, 13–15]. This system is meant to observe the quality of water on the basis of real-time operation. Arduino microcontroller is used along with a set of sensors for the purpose of water management [9, 17–19].

3 Proposed System

The purpose of system is to control water flow once highest level is reached and permit water flow once lowest level is achieved. Thus, the supply of water is regulated with a monitoring system. When the water level achieves highest level of tank, motor will be switched off automatically. The water monitoring system comprises a microcontroller (ESP 32), a relay and ultrasonic sensors. The relay and ultrasonic sensors are connected to the output of microcontroller. When the motor is switched ON, the system will recognize the sensors thereby sensing water level in tank. When tank is complete, motor will be switched OFF automatically. This arrangement can be installed at the top of buildings and apartments in order to maintain the level of water in overhead tanks. The controller will observe level of water in tank and switch ON the motor for pumping water from a water source to tank. The working of motor depends on actual requirements of the building. It may pump water over long distances in a vertical manner or through a group of pipes integrated with valves. The installed motor must be able to meet the requirement of flow rate and pressure in order to provide adequate supply of water to specific locations. Figure 1 shows block diagram of water management system. The water level can be monitored by the controller using water level sensor. The level of water can be monitored and transmitted to a local server or monitoring system in wireless mode. By this approach, the water level can be tracked on real-time basis.

Ultrasonic sensor is interfaced with ESP 32 module using trigger and echo pins. This arrangement activates ESP 32 to interface with ultrasonic sensor and distance measurements can be retrieved. ESP 32 can be enabled to perform as a web server and do communication with clients. To implement this, it must be flashed with Arduino core ESP 32. Figure 2 shows the schematic of water tank monitoring system. Figure 3 shows prototype of water tank monitoring system.

4 Results and Discussions

Figure 4 shows inside view of overhead tank with ultrasonic sensor. The ESP 32 module gets the value of distance from ultrasonic sensor by using a function called `getDistance()`. Once value of distance is received, ESP 32 prepares an HTML page which includes the distance information. This web page is sent to web browser of client. When the user opens this web page in his web browser, it will be displayed. This permits the user to see the distance reading received from ultrasonic sensor by means of web page. The webpage of the system includes a water level animation. This

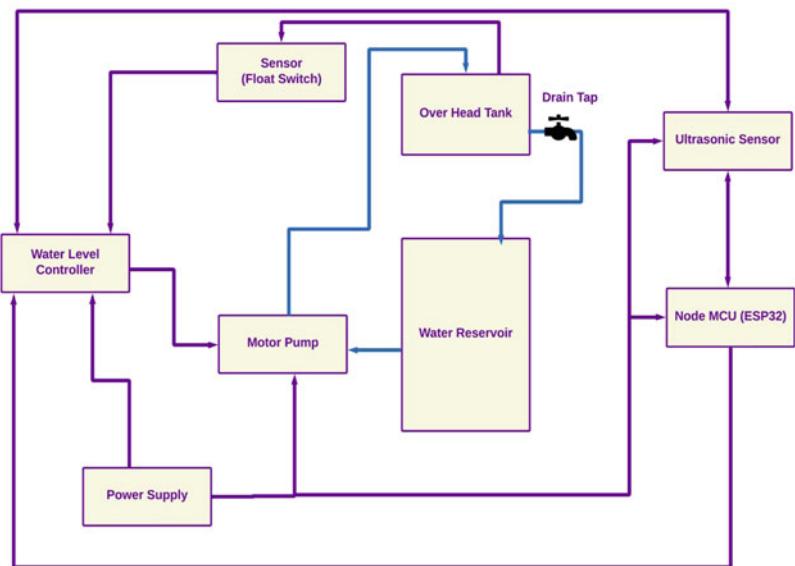


Fig. 1 Block diagram of water management system

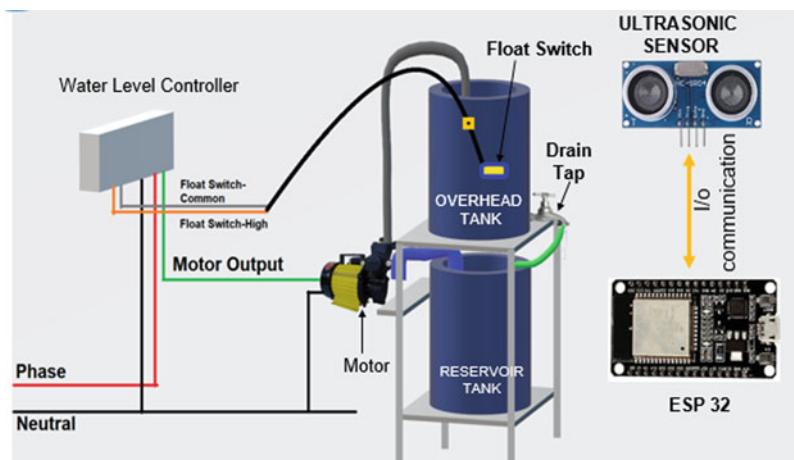


Fig. 2 Schematic of water management system

type of animation indicates the current level of water with respect to distance reading sensed by ultrasonic sensor. In order to obtain this, the web page gets the distance value from HTML page sent by the microcontroller. The water level animation is updated appropriately and reflects the water level on real-time basis. Thus, ESP 32 controller acts an intermediate agent between sensor and web browser of the client. This integrated arrangement permits users to track the level of water by means of a



Fig. 3 Prototype of water tank monitoring system

web interface at the remote. Figures 5, 6 and 7 indicate the dashboard outputs for sensing low level, moderate level and top level of water inside the tank, respectively.



Fig. 4 Inside view of overhead tank with ultrasonic sensor

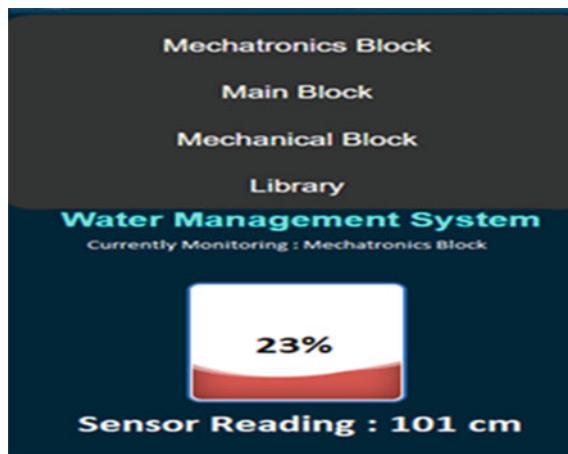


Fig. 5 Dashboard output for low water level

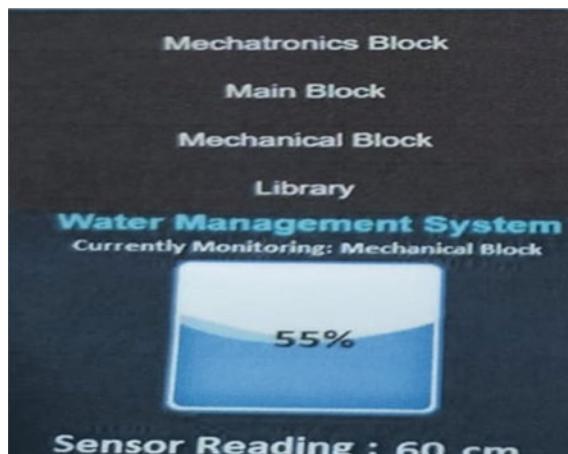


Fig. 6 Dashboard output for moderate water level

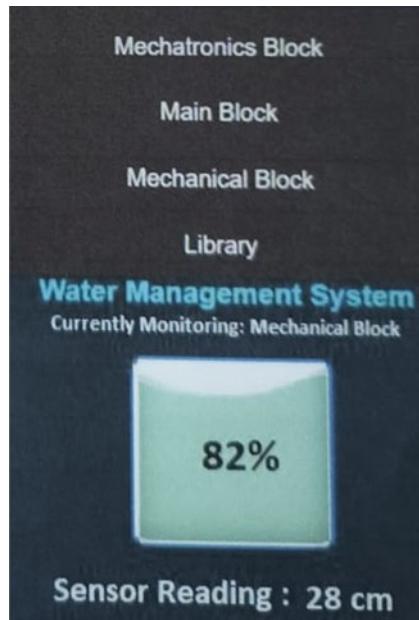


Fig. 7 Dashboard output for high water level

5 Conclusion

The design and development of a water level management system are presented with a web dashboard facility for commercial buildings and residential apartments in the present scenario in view of water scarcity. By implementing the advanced technologies such as microcontroller, sensor, relay and web dashboard, the system provides an efficient methodology to monitor and regulate the consumption of water. The motor pump will be operated ON and OFF whenever the tank reaches dry and overflow conditions, respectively. The model was constructed and expected results were obtained accordingly. This methodology is an invaluable solution for conservation and resource management of water. Thus, the system provides utilization of water resources, prevents wastage of water and promotes practices of water consumption both in residential and commercial outlets. The water management system indicates satisfactory performance with its technology and it is durable and economical. The web page involved in the system for viewing the sensor reading can be developed as mobile application in the future so that online viewing can be done easily on smart phones.

References

1. Loizou C, Koutroulis E., Zalikas D, Lontas G (2016) A low-cost capacitive sensor for water level monitoring in large-scale storage tanks. In: IEEE International Conference on Industrial Technology (ICIT), pp 125–165
2. Raja V, Sivabalakrishnan D, Uthariaraj R (2019) An intelligent decision support system for precision agriculture using IoT and machine learning. In: International Conference on Computer Communication and Informatics (ICCCI), pp 1–6
3. Gama-Moreno LA, Corrales A, Ramirez A, Carrillo-Ruiz, Jaramillo F (2019) Smart water quality monitoring system using IoT and cloud computing. In: IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), pp 1–5
4. Johari A, AbdWahab MH, Abdul Latif NS, Erdi Ayob M, Izwan Ayob M, Afif Ayob M, Haji Mohd MN (2011) Tank water level monitoring system using GSM network. *Int J Comput Sci Inf Technol* 2:1114–1115
5. Haron, NS, Mahamad MKB, Aziz IA, Mehat M (2008) A system architecture for water quality monitoring system using wired sensors. In: International Symposium on Information Technology, Kuala Lumpur, Malaysia, vol 4, pp 1–7
6. Jin N, Ma R, Lv Y, Lou X, Wei Q (2012) A novel design of water environment monitoring system based on WSN. In: International Conference on Computer Design and Applications, vol 2, San Jose, CA, USA, pp 2–593
7. Peng P (2007) Study on key technology of remote real-time monitoring system for wetland water environment based on wireless sensor network. *Tech J Sens Actuators*, 187–190
8. Shankari M, Jyothi ME, Naveen I, Herle H (2013) Wireless automatic water level control using radio frequency communication. *Int J Comput Sci Inf Technol* 2:1320–1324
9. Kaur A, Sandhu K (2018) Real-time water quality monitoring system using IoT. In: Third International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), pp 1–6
10. Kumar R, Mittal A, Bansal V (2018) Smart water quality monitoring and control system. In: 8th International Conference on Cloud Computing, Data Science and Engineering (Confluence), pp 654–658
11. Kedia N (2015) Water quality monitoring for rural areas—a sensor cloud based economical project. In: International Conference on Next Generation Computing Technologies, pp 50–54
12. Clote NA, Malekian R, Nair L (2016) Design of smart sensors for real-time water quality monitoring. *IEEE Access* 4(9): 3975–3990
13. Xu Y, Liu F (2017) Application of wireless sensor network in water quality monitoring. In: IEEE International Conference on Computational Science and Engineering (CSE), Guangzhou, China, pp 21–24
14. Qian D, Xia X, Zeng G, Li Y (2017) Wireless sensor network-based water quality monitoring system for the protection of urban rivers. *Water Sci Technol* 75(4):822–831
15. Singh R, Gehlot A, Thakur AK, Swain M, Akram SV (2020) Wireless sensor network with power management system for water level regulation in paddy fields. *Int J Innovative Technol Explor Eng (IJITEE)* 9:1243–1246
16. Shahin G, Elamvazuthi I, Taib A (2017) A review of wireless sensors and networks applications in agriculture. *Comput Stan Interfaces* 53:322–331
17. Dinesh Kumar T, Archana MA, Umapathy K, Gayathri G, Bharathvaja V, Anandhi B (2023) RFID based smart electronic locking system for electric cycles. In: IEEE Xplore, Fourth IEEE International Conference on Electronics and Sustainable Communication Systems, 1CESC 2023 Proceedings, Coimbatore, pp 76–81
18. Umapathy K, Omkumar S, Muthukumaran D, Chandramohan S, Sivakumar M (2023) Autonomous health care robot. *Lect Notes Netw Syst* 617:227–233
19. Umapathy K, Omkumar S, Muthukumaran D, Chandramohan S, Sivakumar M (2023) Thingspeak based garbage monitoring and collecting system. *Lect Notes Netw Syst* 617:235–242

A Review on Privacy Preservation in Cloud Computing and Recent Trends



Srutipragyan Swain, Prasant Kumar Patnaik, and Banchhanidhi Dash

Abstract Cloud computing is a distributed computing architectural model that offers on-demand computing services via the Internet. It includes servers, computing resources, applications, data storage, development tools, etc. However, despite the different services and benefits offered by the cloud, security and privacy are the primary hindrances for organizations for opting cloud. The real benefit of the cloud can be only enjoyed if the sensitive data kept in the database can be secured from unauthorized access. Privacy preservation in cloud computing is nothing but hiding sensitive data, where the data is stored in the software and databases scattered around the Internet. Various methods have been put forward by researchers to preserve privacy in cloud computing over the past few years.

Keywords Privacy preservation · Sensitive attribute · QID · Anonymization

1 Introduction

Cloud computing otherwise known as utility computing by which the utility of IT increases by manifold in terms of infrastructure, platform, applications, or general services without making any significant investment [1]. Recent time has witnessed increased invention in cloud computing technology because of its reduced cost, time, maintenance, and easy and fast access to data from any corner of the world [2]. Applications of cloud computing [3] are best-suited in such fields where the user wants to access different applications on the go using multiple devices from

S. Swain (✉) · P. K. Patnaik · B. Dash

School of Computer Engineering, KIIT Deemed to Be University, Bhubaneswar, India

e-mail: sruti56@gmail.com

P. K. Patnaik

e-mail: patnaikprasantfcs@kiit.ac.in

B. Dash

e-mail: banchhanidhi.dash@kiit.ac.in

different locations, the user prefers minimum hardware and at the same time can utilize maximum storage space. Such fields are security applications, disaster relief, crowd computing, etc. Among different challenges/issues in cloud computing such as multi-tenancy support, identity management of cloud users, security of applications, and attaining control over the lifecycle of outsourced data [4], security is the main issue as it risks the data placed for processes or references in the external repository [5]. The most critical aspect of cloud computing is the protection of sensitive data. One of the major obstacles lies in ensuring confidentiality when opting to outsource data over the cloud.

1.1 Security and Privacy

Security of data includes confidentiality, integrity, and availability along with data privacy. Figure 1 shows the privacy requirements in a cloud environment.

Data Integrity: It ensures consistency and accuracy over the lifecycle of data while outsourcing data over the cloud.

Confidentiality of Data: It pertains to protecting data from illegitimate access which preserves the privacy of data.

Data Availability: Though integrity and confidentiality are maintained data must be available when required, otherwise it is useless unless it is used by organizations.

Data Privacy: It deals with the aspect of protecting and handling personal and confidential data where a third party is involved in the cloud environment.

Privacy preservation in cloud computing can be categorized into three broad areas data anonymization-based algorithms, access control-based algorithms, and encryption algorithms, based on the characteristics and methodology adapted.

2 Anonymization-Based Techniques

Data anonymization is one of the widely adopted techniques to achieve privacy preservation. Here the original data is converted into another form and only those data are shared with the public. This enables to hiding of private and sensitive data of an individual [6]. A dataset contains different attributes. They are explicit attributes,

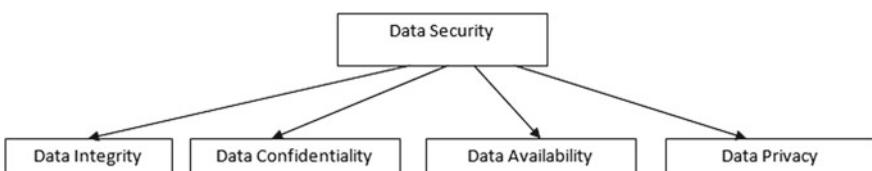


Fig. 1 Security and privacy requirements

sensitive attributes, and non-sensitive attributes [7]. Explicit attributes do not take part in anonymization. Only sensitive attributes take part in the anonymization process. Some non-sensitive attributes can disclose privacy if it is linked to other external data sources. These types of non-sensitive attributes are termed quasi-identifiers [8]. So, the challenge is to transform the quasi-identifier's values to hide sensitive information. In the anonymization technique, we need to anonymize QID and sensitive attributes as they are prone to identity disclosure indirectly [9]. So only QID and sensitive attributes are involved in this process to anonymize data for categorical attributes like sex, education, etc. The generalization technique is used and for numerical attributes like the zip code discretization technique is used [10]. Privacy attacks on shared data can be of different types. "It may be record linkage attack, table linkage attack, attribute linkage attack, and probabilistic attack" [6]. A record linkage attack is performed by linking published records with background knowledge processed by the adversary. To prevent this attack K-anonymity privacy model is used, which is proposed by Samarati and Sweeney [11, 12]. The size of the generalization must be equal to zero or at least k for any quasi-identifier (QID). By this, the probability of linking a target to a record through QID can be at most 1/k. An attribute linkage attack is performed by linking sensitive attributes with the published table after analysis. To prevent this l-diversity privacy principle was proposed [13]. This theory requires each quasi-identifier group to contain l distinct sensitive values. Consider a scenario where we will apply the l-diversity principle to hospital patient records. Table 1 is an example of patient health data. Here, Name is an explicit identifier, Disease is a sensitive attribute, and {sex, zip code, education} are quasi-identifiers [14]. Table 2 represents a distinct 2-diverse release since here each quasi-identifier (QID) group contains at least 2 distinct sensitive values and hence makes the attack difficult. For example, the QID group < any_sex,448**, graduate > contains < HIV and hepatitis >. Similarly, < Any_sex,448**,Any_Education > contains < HIV and Bronchitis >. The drawback of this technique is it can only deal with categorical attributes and fails with numerical attributes. To deal with this situation emerges other privacy models such as (k,e)-anonymity [14], which requires that the maximum sensitive values minus minimum sensitive values in a QID group must be at least e. Variance control [15] requires that the variance of sensitive values must not be less than a threshold value. (E, m)-anonymity [16] requires for any value of a sensitive attribute in a quasi-identifier group, a maximum of 1/m of records can have sensitive values similar to the value and E indicates the similarity. t-closeness [17] is used to prevent data distribution skewness attacks. It applies to both numerical and categorical attributes. This requires that the distribution of sensitive values should be proximate to the distribution of the entire dataset. But this cannot sustain proximity attacks. Another model known as the general dissimilarity proximity privacy model was proposed [18] which deals with this problem. Anonymization increases privacy but reduces data utility. Hence, we should only prefer anonymization which provides more privacy at less cost of data utility. The HAC model [19] uses partial homographic encryption along with K-anonymity. It mitigates the double-edged risk of privacy preservation along with the data utility services of storing data in the cloud to maintain the data utility feature, this model has proposed to apply homomorphic encryption only to

Table 1 Patient health data

No.	Name	Sex	Zip code	Education	Disease
1.	Annie	Female	44,815	Master	HIV
2.	Boopathy	Male	44,807	Doctorate	Hepatitis
3.	Eswar	Male	44,803	12 th	Bronchitis
4.	Esther	Female	44,810	Doctorate	HIV
5.	Janardhanan	Male	44,807	11 th	Flu
6.	Jeeva	Male	44,811	Bachelor	HIV

Table 2 Distinct 2-diverse release

No.	Sex	Zip code	Education	Disease
1.	AnySex	448*	Graduate	HIV
2.	AnySex	448*	Graduate	Hepatitis
3.	AnySex	448*	AnyEducation	HIV
4.	AnySex	448*	AnyEducation	Bronchitis
5.	Male	448*	AnyEducation	Flu
6.	Male	448*	AnyEducation	HIV

* for anonymization which is one of the principle of privacy preservation

sensitive attributes. Quasi-identifier index-based model [9] preserves privacy and maintains data utility of incremental and distributed data in the cloud. If anonymity is lost then generalization [20] will be done on the datasets and if datasets are over-generalized then specialization [21] operation will be done, otherwise data utility will be degraded.

Here we have taken the example of a student record (Table 1) and anonymization is applied, and the output is shown in Fig. 1. For anonymization we have used the anonympy tool and for implementation, PYTHON is used (Table 3).

2.1 Background

For data anonymization, numerous anonymization methods like generalization, PCA masking, and perturbation are used for numeric data. For categorical data, methods like tokenization and partial email masking are used. For date time data, synthetic date and perturbation both can be used.

Table 3 Student record before anonymization

First_name	Age	Birth date	City	Web	Postal	std_email	Dept	Course
Alisha	23	4/12/89	Cuttack	www.ali sha.com	754,123	asahoo@rediffmail.com	comp_sci	MCA
Sambit	21	3/11/06	Kendrapada	www.ope nho use.com	753,001	sambitkumar@gmmail.com	mech	BTech
Manoj	25	2/14/20	Balasore	www.msg roc ery.com	753,007	manoj1.das@gmail.com	elect	BTech
Naitik	23	1/6/16	Bhadrak	www.sruti.com	653,002	ndas@gmail.com	maths_dept	MSc
Sujata	33	2/5/88	Jagatsinghpur	www.dol lar1.com	862,354	sujata2.ray@gmail.com	maths_dept	PhD

2.2 Calling of Generic Function for Data Anonymization

```
>>>from anonympy.pandas import dfAnonymizer
>>>from anonympy.pandas.utils_pandas import available_methods
>>>anonym = dfAnonymizer(df)
```

For privacy preservation and securing of data in the cloud, different techniques, like data anonymization, encryption, and a combination of different cryptographic approach have been proposed. The combination of RSA and DES algorithms [22] makes the privacy of user data more robust. The encryption technique is to be applied only for sensitive data to achieve privacy and data utility (Fig. 2).

06-03-11	Kendrapada	www.openhouse.com	753001	sambitkumar@gmmail.com	mech	bttech
20-02-14	Balasore	www.msgrocery.com	753007	manoj1.das@gmail.com	elect	bttech
16-01-06	Bhadrak	www.sruti.com	653002	ndas@gmail.com	maths_dept	msc
88-02-05	Jagatsinghpur	www.dollar1.com	862354	sujata2.ray@gmail.com	maths_dept	phd
soFlow >= 2.0, or Flax have been found. Models won't be available and only tokenizers, configuration and file/data utilities						
birthdate	city	web	postal	std_email	dept	
1988-11-12	Cuttack	c4715a6290	754128	a*****o@rediffmail.com	comp_sci	
2006-02-07	Kendrapada	5085bb70c1	752999	s*****rg@gmmail.com	mech	
2019-10-12	Balasore	b109f00e8b	753015	m*****sgmail.com	elect	
2015-03-12	Bhadrak	0431bb49fa	653002	n*****sgmail.com	maths_dept	
1987-05-07	Jagatsinghpur	e2c2010229	862346	s*****y@gmail.com	maths_dept	

Fig. 2 Student record after anonymization

3 Access Control-Based Techniques

Access control mechanism prevents data from illegal access. Selective encryption models [23] have selective access to the encrypted data. This technique uses keys for encryption with varying authorization levels. It maintains a key ring having multiple keys with different levels of authorization. Here the challenge is key management among users. A graph-based authorization policy and its equivalent policy are used to apply read privilege over the data in the cloud. Here a heuristic approach is proposed to obtain an approximate solution. For both read and write privileges [24] two-level access policies, namely coarse-grained and fine-grained are employed for this purpose. This approach has more overhead in key management and policy update operations. In the attribute-based encryption (ABE) model [25] data records are encrypted using access policies based on attributes. ABE can be classified into “Key Policy-Attribute Based encryption (KP-ABE)” [26] and “Cipher text Policy-Attribute Based Encryption (CP-ABE)” [27]. “Access control-based Cloud single sign-on architecture (CSSOA) model” [28] was implemented. In SSO architecture one user is confirmed once. Once the user is logged in, authenticated subsequent request of the user from accessing the service is dealt with by the application till the user remains authenticated. Here we have taken the example of collusion resistance cipher text attribute-based encryption. For reference purposes, we use the student data given in Table 2. For instance, the institute publishes the results of students in the mathematics department. Let the access policy for accessing the result file be student must belong to the mathematics department and have course Ph.D. or MSc (Fig. 3).

((“maths_dept” AND “student”) AND (“MSc” OR “PhD”))

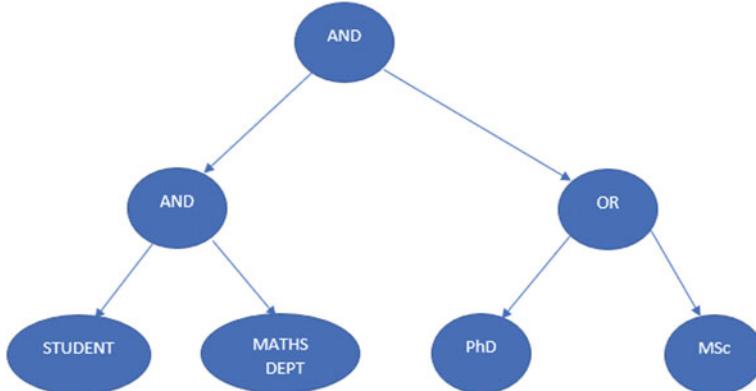


Fig. 3 Access policy structure [The access policy is implemented with the help of the CPABE toolkit]

3.1 Background

Setup module: The Pub_key and Mast_key is generated with the setup module with the help of a bilinear group G_0 .

$$\text{PK} = G_0, g, h = g^\beta$$

$$M = (\beta, g^\alpha) \quad (1)$$

Encryption:

The encrypted message is generated in this step. It selects $s \in Z_p$ and assigns $q = s$. Here, s is a random number. The polynomial q is selected for S and the ciphertext is computed as

$$CT = (C' = Me(g, g)^{\alpha s}, C = h^s) \quad (2)$$

Keygen:

The private key SK can be generated as

$$SK = D = g^{(\alpha+r)/\beta}, \forall j \in S : D_j = g^r H(j)^r, D'_j = g^r \quad (3)$$

where r and r_j are random numbers and $r \in Z_p$ and $r_j \in Z_p$.

Decryption:

$$\frac{C'}{\frac{e(C,D)}{A}} = \frac{C'}{e\left(\frac{h^s, g^{\frac{\alpha+y}{\beta}}}{(g,g)^{ys}}\right)} = M \quad (4)$$

The Ciphertext C , Private Key $S K$, and a policy attribute string S are used in the decryption algorithm. It is recursive in nature.

The Cpabe Tool Kit:

The CPABE tool kit uses the pairing-based cryptography (PBC) library. CPABE package is available under GPL on the web.

cpabe-setup:

This step generates public_key and master_key.

cpabe-keygen:

For the master key, it generates a private key for the attribute set.

cpabe-enc:

It is used to encrypt a file under a certain policy language specified in the access tree with the help of a public key.

cpabe-dec:

It is used to decrypt a file using the generated private key.

The master_key and public_key is generated in the cpabe-setup phase. With the help of the master key, the private key is generated. In the student database in Table 1

```
[srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % cpabe-keygen -o manoj_priv_key pub_key master_key elect btech
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % cpabe-dec pub_key manoj_priv_key result.pdf.cpabe
can't read file: result.pdf.cpabe
[srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % ls
AUTHORS                               missing
[Alisha_priv_key] cpabe-enc             mkinstalldirs
COPYING                                cpabe-enc.more-man   naitik_priv_key
INSTALL                                cpabe-keygen        policy_lang.c
[Makefile]                             cpabe-keygen.1      policy_lang.h
[Makefile.in]                           cpabe-keygen.more-man policy_lang.o
NEWS                                   cpabe-setup        policy_lang.y
README                                cpabe-setup.1     pub_key
acinclude.m4                            cpabe-setup.more-man result.pdf
aclocal.m4                            cpabe.h            sachin_priv_key
common.c                               dec.c              sara_priv_key
common.h                               dec.o              setup.c
common.o                               enc.c              setup.o
config.log                            install-sh       sruti_priv_key
config.status                           kevin_priv_key  student.xlsx
configure                             keygen.c          sujata_priv_key
configure.ac                           keygen.o          test-lang
cpabe-dec                            manoj_priv_key  test-lang.c
cpabe-dec.1                           master_key
cpabe-dec.more-man
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % cpabe-enc pub_key result.pdf
(mathsm_dept) and (1 of (msc,phd))
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % cpabe-keygen -o manoj_priv_key pub_key master_key elect btech
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % cpabe-dec pub_key manoj_priv_key result.pdf.cpabe
cannot decrypt, attributes in key do not satisfy policy
[srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3 % ]
```

Fig. 4 Simulation output

for the two students Manoj and Sujata two private keys Manoj_priv_key and Sujata_priv_key are generated using the master key. Suppose later someone wants to encrypt the sensitive document result.pdf file. For this, it requires a public key and cpabe-enc to encrypt under a specified policy. To access the sensitive document the student must belong to the maths department and the course must be either MSc or PhD as per the policy. The output in Fig. 4 shows that Manoj_priv_key does not satisfy the policy as he belongs to the electrical department and has a course Btech. So, we get the output as denial of decryption of file as attributes in the key do not satisfy the policy. Similarly, the result.pdf file is successfully decrypted in the case of Sujata_priv_key as it satisfies the policy.

```
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3% cpabe-keygen
-o Naitik_priv_keypub_keymaster_keymaths_deptmsc
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3% cpabe-keygen
-o Sujata_priv_keypub_keymaster_keymaths_deptphd
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3% cpabe-keygen
-o Manoj_priv_keypub_keymaster_key elect btech
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3% cpabe-enc
pub_key result.pdf
(mathsm_dept) and (1 of (msc,phd))
srutipragyanswain@Srutipragyans-MacBook-Air cpabe-0.11 3% cpabe-dec
pub_keyManoj_priv_keyresult.pdf.cpabe
Cannot decrypt, attributes in the key do not satisfy policy.
cpabe-dec pub_keySujata_priv_keyresult.pdf.cpabe
Successfully Runs.
```

4 Hybrid Technique

To achieve privacy preservation, it may be required to use a combination of more methods such as data anonymization, data fragmentation, or data mining [30] to develop a privacy preservation technique that can address the problems associated with a single privacy-preserving technique.

5 Comparison of Different Privacy Preservation Techniques

Table 4 gives the comparison between different privacy preservation techniques.

6 Recent Trends in Privacy Preservation

Recent years have seen an increase in the research of many areas relating to cloud computing. Much research has been done to solve one of the major challenges of preserving the privacy of outsourced data over the cloud. In this paper, we have collected different research work related to privacy preservation in the cloud and categorized it into three categories. From the study, it was found that various algorithms and combinations of more than one technique are used for privacy preservation in the cloud environment. Considering the factors like data utility and privacy the hybrid approach is gaining popularity and better results compared with the existing methods. Figure 5 indicates the research done in various fields of privacy preservation in cloud environments and has been categorized into different groups. Anonymization with data fragmentation or data mining, K-anonymity with encryption is gaining popularity among researchers in the current era.

7 Conclusion

Cloud computing is becoming increasingly popular as the data is available to the user on the go anywhere in the world. However, when data is outsourced over the cloud, it generates the risk of a data breach. Protecting the user's personal data over the cloud platform is the most researchable topic in the current era. Different privacy preservation algorithms are proposed. The algorithms have been classified into anonymization-based algorithms, access control-based, and hybrid algorithms. Further research will be done to find a method that is more feasible in terms of privacy preservation and data utility in a cloud computing environment.

Table 4 Comparison of different privacy preservation techniques

Author	Method	Pros	Cons
Mohammad Karim Sohrabi et al.	SSO Architecture	User once authenticated, subsequent requests can be served by the system's application in an integrated manner	<ol style="list-style-type: none"> 1. Misuse of passwords or disclosure can cause security flaws and could affect multiple applications and sources 2. Requires high SSO IAM availability as the limitations of a single point of failure exist
Huda Osman et al.	Anonymization	Prevent identity and attribute disclosures of data before outsourcing to the cloud	<ol style="list-style-type: none"> 1. More information loss because of top-level generalization
Abhisek R. Ladole et al.	Hybrid	Both data utility and privacy can be maintained	
Xuyun Zhang et al.	QID, Anonymization	It preserves privacy over incremental and distributed datasets over cloud	Privacy preservation for data and computation-intensive applications on the cloud involving huge volume of incremental datasets is still a challenging issue
Jayashree Agarkhed et al.	Encryption	The efficiency and consistency of cloud servers can be maintained by combining of RSA and AES algorithms	Insufficient to manage huge datasets
Manish H. Gurkhede et al.	Encryption	Both digital rights of the content and the privacy of the user are preserved without depending on a Trusted Third Party (TTP)	Information may be lost during the process of reconstruction
Supachai Tangwongsan et al.	Access Control	Secured data retrieval from the cloud storage	
Ting Wang et al.	K-anonymity	It addresses general proximity breaches, provides effective protection against linking attack	
Ninghui Li et al.	t-closeness	Prevents skewness attack, similarity attack	This approach does not solve the purpose in the case of multiple sensitive attributes

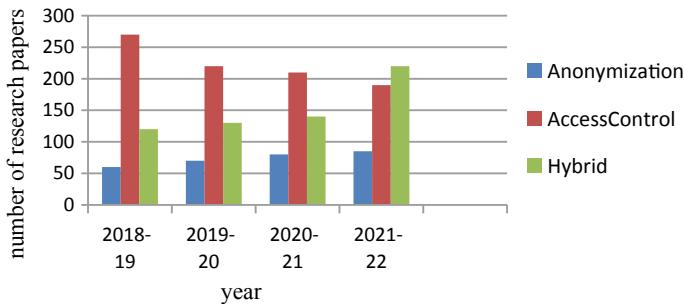


Fig. 5 Graph shows the number of research papers in the field of cloud computing from 2019 to 2022 according to different categories of privacy preservation techniques

References

- Knorr E, Gruman G (2008) What cloud computing really means. *InfoWorld* 7(20–20):1–17
- Le DN, Pal S, Pattnaik PK (2022) Auditing concept in cloud computing. In: *Cloud computing solutions: architecture, data storage, implementation and security*, pp 165–180
- Ali M, Miraz MH (2013) Cloud computing applications. In: *Proceedings of the International Conference on Cloud Computing and eGovernance*, vol 1
- Zhang X, Liu C, Nepal S, Pandey S, Chen J (2012) A privacy leakage upper bound constraint-based approach for cost-effective privacy preserving of intermediate data sets in cloud. *IEEE Trans Parallel Distrib Syst* 24(6):1192–1202
- Lin HY, Tzeng WG (2011) A secure erasure code-based cloud storage system with secure data forwarding. *IEEE Trans Parallel Distrib Syst* 23(6):995–1003
- Fung BC, Wang K, Chen R, Yu PS (2010) Privacy-preserving data publishing: a survey of recent developments. *ACM Comput Surveys (Csur)* 42(4):1–53
- Andoni A, Indyk P (2008) Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Commun ACM* 51(1):117–122
- Lee KH, Lee YJ, Choi H, Chung YD, Moon B (2012) Parallel data processing with MapReduce: a survey. *AcM SIGMoD Record* 40(4):11–20
- Fung BC, Wang K, Philip SY (2007) Anonymizing classification data for privacy preservation. *IEEE Trans Knowl Data Eng* 19(5):711–725
- Samarati P, Di Vimercati SDC (2010) Data protection in outsourcing scenarios: Issues and directions. In: *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, pp 1–14, April
- Sweeney L (2002) K-anonymity: a model for protecting privacy. *Int J Uncertain Fuzziness Knowl Based Syst* 10(05):557–570
- Gotz M, Machanavajjhala A, Wang G, Xiao X, Gehrke J (2011) Publishing search logs—a comparative study of privacy guarantees. *IEEE Trans Knowl Data Eng* 24(3):520–532
- Zhang Q, Koudas N, Srivastava D, Yu T (2007) Aggregate query answering on anonymized tables. In: *2007 IEEE 23rd international conference on data engineering*. IEEE, pp 116–125, April
- LeFevre K, DeWitt DJ, Ramakrishnan R (2008) Workload-aware anonymization techniques for large-scale datasets. *ACM Trans Database Syst (TODS)* 33(3):1–47
- Li J, Tao Y, Xiao X (2008) Preservation of proximity privacy in publishing numerical sensitive data. In: *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pp 473–486, June
- Li N, Li T, Venkatasubramanian S (2006) t-closeness: privacy beyond k-anonymity and l-diversity. In: *2007 IEEE 23rd international conference on data engineering*. IEEE, pp 106–115, April

17. Wang T, Meng S, Bamba B, Liu L, Pu C (2009) A general proximity privacy principle. In: 2009 IEEE 25th International Conference on Data Engineering. IEEE, pp 1279–1282, March
18. Osman H, Siraj MM, Maarof MA (2021) HAC: model for privacy-preserving outsourced data over cloud. In: 2021 3rd International Cyber Resilience Conference (CRC). IEEE, pp 1–4, January
19. Zhang X, Liu C, Nepal S, Chen J (2013) An efficient quasi-identifier index-based approach for privacy preservation over incremental data sets on cloud. *J Comput Syst Sci* 79(5):542–555
20. Wang J, Zhao Y, Jiang S, Le J (2010) Providing privacy preserving in cloud computing. In: 3rd International Conference on Human System Interaction. IEEE, pp 472–475, May
21. Agarkhed J, Ashalatha R (2017) A privacy preservation scheme in cloud environment. In: 2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). IEEE, pp 549–552, February
22. Vimercati SDCD, Foresti S, Jajodia S, Paraboschi S, Samarati P (2010) Encryption policies for regulating access to outsourced data. *ACM Trans Database Syst (TODS)* 35(2):1–46
23. Raykova M, Zhao H, Bellouin SM (2012) Privacy enhanced access control for outsourced data sharing. In: International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, pp 223–238, February
24. Sahai A, Waters B (2005) Fuzzy identity-based encryption. In: Annual international conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, pp 457–473, May
25. Goyal V, Pandey O, Sahai A, Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on Computer and communications security, pp 89–98, October
26. Bethencourt J, Sahai A, Waters B (2007) Ciphertext-policy attribute-based encryption. In: 2007 IEEE symposium on security and privacy (SP’07). IEEE, pp 321–334, May
27. Sohrabi MK, Ghods V, Mohammadian N (2017) Privacy of cloud data using a secure SSO architecture. In: 2017 Computing Conference. IEEE, pp 224–229, July
28. Ladole AR, Chhajed KK, Shelke FM (2018) A survey on privacy preserving techniques in cloud environments. In: 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT). IEEE, pp 1–5, March

EEECT-IOT-HWSN: The Energy Efficient-Based Enhanced Clustering Technique Using IOT-Based Heterogeneous Wireless Sensor Networks



Mustafa Dh. Hassib, Mohammed Joudah Zaiter,
and Wasan Hashim Al Masoody

Abstract IoT refers to the interconnection of electronic devices, machines, and physical objects in our environment. Heterogeneous wireless sensor networks (HWSNs) are among the promising wireless technologies that play an important role in monitoring remote areas. The clustering algorithm reduces energy consumption by using a key technique. It increases the network's scalability and lifetime. Wireless sensor networks with heterogeneous characteristics should be designed with energy-efficient clustering protocols. In this paper, design a novel EEECT-IOT-HWSN technique for the three-tier heterogeneous networks. The EEECT-IOT-HWSN technique has used the modified threshold formula for the cluster head selection based on the combination of the energy and distance of the SNs. The performance of the proposed model shows the higher residual energy, less dead SNs, and higher network lifetime when compared with the ADV-LEACH1 (HETRO), and ADV-LEACH1 (HOMO) technique.

1 Introduction

In IoT-based heterogeneous wireless sensor networks (HWSNs), IoTs are used for a wide range of tasks, including military operations, surveillance, environmental monitoring, and animal tracking. The environments they operate in are challenging,

M. Dh. Hassib (✉)

Department of Communications Engineering, University of Technology, Baghdad, Iraq
e-mail: Mustafa.d.Hassib@uotechnology.edu.iq

M. J. Zaiter

Electrical Engineering Technical College Middle Technical University, Baghdad, Iraq
e-mail: mjzaiter@mtu.edu.iq

W. H. Al Masoody

Electrical Engineering Department, College of Engineering, Babylon, Iraq
e-mail: eng.wasan.hashim.lec@uobabylon.edu.iq

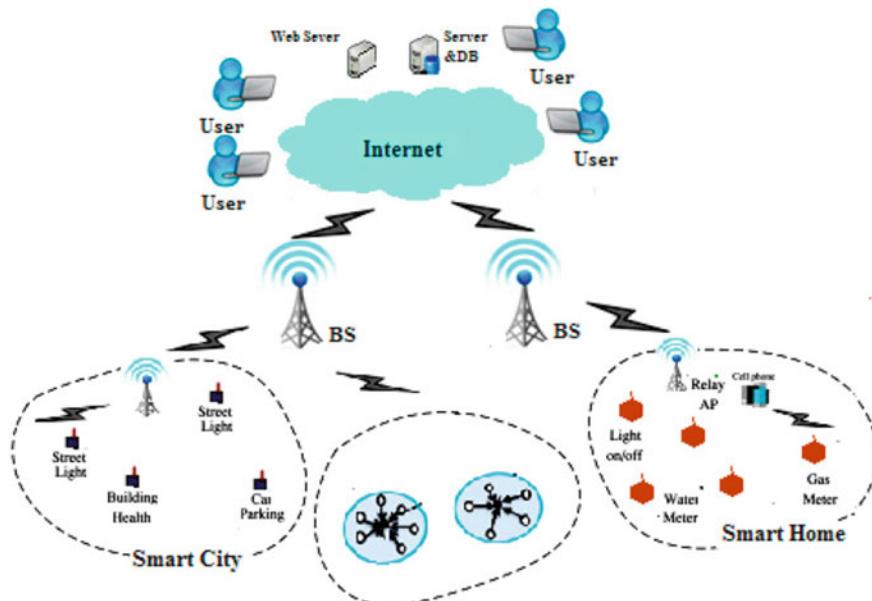


Fig. 1 IoT-based heterogeneous wireless sensor network architecture

including deep seas, arctic regions, and dangerous battle zones. Several IoT-based WSN applications have been discussed, including smart agriculture, smart cities, and intelligent water applications [1–3]. The IoT aims to achieve ubiquitous connectivity [4]. Smartphones and sensor nodes (SN) are some examples of IoT devices with heterogeneous energy consumption, costs, and Internet accessibility. To provide high-quality services to users at anytime and anywhere, various sensor nodes (SNs) are interconnected with different networking capacities [5]. The WSN is different from the VANET [6–8] network due to its mobility [9]. Heterogeneous WSN architecture with IoTs is depicted in Fig. 1.

For high-impact services to be provided to users connected to various applications, scenarios, and requirements, high scale and standardized services must be created as IoT sensor nodes that are united, comprehensive, and flawless. The development of IoT applications is a continuous process, however numerous issues remain to be solved, including costs, power, self-organization, low-latency, computing, and distributed intelligence. In spite of the fact that IoT provides new opportunities to end users and the industry, it still lacks an efficient architecture and system set of criteria that link the real as well as virtual worlds together [10]. There are several key challenges to be addressed. (1) Architecture challenge: Different sensors (e.g., physical sensors, chemical sensors, biometric sensors, and cameras) are emerging, and smart interconnected devices and intensive far-ranging technology are being used in IoT applications. A wireless, ad hoc, automatic connection is made between these interconnected devices. Decentralization, mobility, and complexity are increased by

it. (2) Technical challenge: IoT is a complex system, which requires a range of technologies to be developed, regardless of the application area. A smart IoT environment faces heterogeneous challenges as a result. (3) Hardware challenge: With IoT, smart devices are integrated into smart systems. The deployment of IoT applications and services must be swift in order to maximize the potential of communication between the devices. Hence, researchers develop cost-effective, compact, and highly functional hardware like wireless tractable systems. Detailed analysis of the latest research findings follows in Sect. 2. Section 3 presents the proposed methodology design, energy model, and network design model. Section 4 describes the results and discussion of the proposed model with comparative analysis. Section 5 presents the conclusion and last is the references.

2 Literature Review

The combined infrastructure is now being used to conduct integrated projects in some industries in addition to creating an IoT-based HWSN integration system. Integrated environments are highly profitable for networking industries, which is why their demand for these environments is growing rapidly [10, 11]. The DCHSM method is proposed as a way to extend the lifetime of IoT systems by dynamically selecting the CH. The large-scale sensing area is distributed into smaller clusters with the help of voronoi diagrams to ensure extreme coverage. A two-phase selection process is then followed by CH. CH selection involves two stages: a first one considers perceived probability, and a second one considers estimates of survival time. In [12] IoT-enabled cluster-routing protocol for WSNs is proposed by the authors. Cluster-based routing protocols reduce data transmission to gateway nodes by accumulating and decreasing data transfers within the cluster. Energy consumption is reduced as a result. LEACH [13] is a proactive routing algorithm based on clustering. When sensor nodes (SNs) in LEACH are grouped, they use less energy. To spread the load uniformly over all SNs in a wireless sensor network, hundreds of thousands of sensor nodes are randomly distributed. These sensors nodes continually sense data and deliver it to their corresponding cluster heads (CHs), receiving, aggregating, and forwarding data packets to the base station (BS) or sink. All of the SNs in the LEACH network are homogenous, and each node has finite battery power. Clusters are built to distribute workload across all nodes and to increase their lifetime. In this network, each sensor node is configured to become a CH in turn [13]. Cluster heads (CHs) are chosen at random by each node and only become CHs once every round. After all other nodes have had a chance to become cluster heads (CH), the same node will become CH again. A probabilistic technique is used to choose cluster heads (CHs) [13]. There is a threshold value calculated by the equation below for determining whether a generated value exceeds the threshold value generated by each node, ranging from 0 to 1 [13] that node becomes a CH. The MSEP [21] protocol is an upgrade and augmentation of the LEACH [13] protocol, which employs a clustering-based routing method based on sensor node heterogeneity in networks. Compared

with the normal nodes, advanced nodes in the network have a higher probability of becoming CHs due to their high energy. As well as having a lower energy level than advanced nodes, normal nodes also consume less energy. In WSNs, SEP selects a CH using a distributed approach. CH selection probabilities are weighted according to initial energy for each node in WSN, as it is a heterogeneity-aware protocol. Nodes with different levels of heterogeneity are used in the SEP protocol. ESEP [14, 21] is improvement and enhancement of SEP technique. According to their energy levels, ESEP method considers three types of sensor nodes: normal, advance, and intermediate. A self-configuring WSN is the objective of ESEP, which extends the network lifespan and stability. Data is continuously transmitted from each sensor in a network to the CH, whereas the CH aggregates the data to reduce data redundancy. SEP nodes with additional energy are called advance nodes in ESEP. Some nodes in normal nodes have a little additional energy, while some nodes in intermediate nodes have some extra energy but not as much as advance nodes. For each kind of node in ESEP, CHs are chosen using a probability-based technique.

Author [15] proposes an energy routing protocol based on the optimal cluster head (CH) selection. As a result of this protocol, the network's lifespan is prolonged. The multifaceted nature of operations still results in delays. As a result, the sensor node with the greatest residual energy is selected without taking into account various factors, such as the SN's distance the BS. Author [16], a random timer algorithm was proposed to construct clusters without requiring global information. Cluster heads and sensor nodes consume a large amount of energy in this algorithm. Additionally, there is a problem with residual energy and cluster head count. Authors in [17] the LEACH-B protocol was proposed. LEACH is used to select CH for the first time. The remaining energy of the node determines the number of CHs starting from the second selection. As a result, CHs per round are stable and near-optimal. Based on current simulation, they have found a balance between network lifetime and network energy consumption compared with the LEACH protocol. In [18] using a LEACHMAC cluster head selection algorithm, authors improved FND and LND times compared with LEACH [24–26].

3 Proposed Methodology

Using a LEACHMAC cluster head selection algorithm, authors improved FND and LND times compared with LEACH. Three types of sensor nodes (SNs) are available in the EEECT-IOT-HWSN protocol: basic, advanced, and super. Randomly distributed SNs are present in all three types of networks. Super SNs have the highest initial energy of all SNs. In contrast, advanced SNs have the lowest initial energy of all SNs. N sensor nodes with m_s percent are super SNs that have β times more initial energy than advanced sensor nodes. The m_a are advanced nodes with more energy at the outset than normal nodes, and $(1 - m_a - m_s)\%$ are nodes that have residual energy at the outset. Normal SNs have E_{in} as their initial energy; advanced and super SNs have $E_{in} * (1 + \alpha)$ and $E_{in} * (1 + \beta)$ initial energies. To calculate CH

selection probabilities, the initial energy of SNs, residual energy of SNs, and distance from BS of each SN are evaluated. CHs are decided by the values of the proposed probability formula. The decision is based on a threshold probability. IOT-HWSNs have total energy $E_{\text{total}} = E_{\text{normal}}^t + E_{\text{advanced}}^t + E_{\text{super}}^t$, as described above. E_{normal} , E_{advanced} , E_{super} t is the sum of all types of SNs energy, in Eqs. (1–3):

$$E_{\text{normal}}^t = N \times E_{\text{in}}(1 - m_a - m_s) \quad (1)$$

$$E_{\text{advanced}}^t = N \times m_a \times E_{\text{in}}(1 + \alpha) \quad (2)$$

$$E_{\text{super}}^t = N \times m_s \times E_{\text{in}}(1 + \beta) \quad (3)$$

The latest design and developed clustering approach of ADV-LEACH1 [22], EEECA-THWSN [23], and TEEECH [19] is used to design the novel proposed EEECT-IOT-HWSN approach. The proposed EEECT-IOT-HWSN technique is similar to ADV-LEACH1 because it uses initial energy, remaining energy, distance from the BS, average distance and maximum distance from the BS, parameters to estimate the distance between the BS, and the EEECT-IOT-HWSN, Kumar et al. [22]. In Eqs. (4–6) calculate normal (P_{normal}), advanced (P_{advanced}), and super (P_{super}), probabilities for cluster head selections:

$$P_{\text{nrm}} = \frac{P}{1 + m_a \alpha + m_s \beta} \quad (4)$$

$$P_{\text{Adv}} = \frac{P * (1 + \alpha)}{1 + m_a \alpha + m_s \beta} \quad (5)$$

$$P_{\text{Sup}} = \frac{P * (1 + \beta)}{1 + m_a \alpha + m_s \beta} \quad (6)$$

The proposed EEECT-IOT-HWSN technique modified the threshold formula of the cluster head selection. Data transmission size and SN distance determine energy consumption. For the new cluster head threshold formula, round energy (E_{round}) and initial energy (E_{initial}) factors are combined along with round distance from the bulk source (D_{round}), maximum distance from BS D_{max} and average distance from the BS D_{avg} of the SNs are used to design the new cluster head threshold formula. CHs must have a lower distance between the SNs and the BS in order to lose less energy. Therefore, less CH will die as a result of the above statement. SNs thresholds are calculated according to the modified threshold formula outlined in Eqs. (7–9).

For normal SNs:

$$T(n)_{\text{nrm}} = \begin{cases} \frac{P_{\text{nrm}} * \left(u \left(\frac{E_{\text{round}}}{E_{\text{initial}}} \right) + v \left(\frac{D_{\text{round}}}{D_{\text{max}}} \right) + \left(\frac{1}{D_{\text{avg}}} \right) \right)}{1 - P_{\text{nrm}} * \left(\text{rmod} \left(\frac{1}{P_{\text{nrm}}} \right) \right)}, & n \in G \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

For advanced SNs:

$$T(n)_{\text{adv}} = \begin{cases} \frac{P_{\text{adv}} * \left(u \left(\frac{E_{\text{round}}}{E_{\text{initial}}} \right) + v \left(\frac{D_{\text{round}}}{D_{\max}} \right) + \left(\frac{1}{D_{\text{avg}}} \right) \right)}{1 - P_{\text{adv}} * (\text{rmod}(\frac{1}{P_{\text{adv}}}))}, & n \in G \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

For super SNs:

$$T(n)_{\text{sup}} = \begin{cases} \frac{P_{\text{sup}} * \left(u \left(\frac{E_{\text{round}}}{E_{\text{initial}}} \right) + v \left(\frac{D_{\text{round}}}{D_{\max}} \right) + \left(\frac{1}{D_{\text{avg}}} \right) \right)}{1 - P_{\text{sup}} * (\text{rmod}(\frac{1}{P_{\text{sup}}}))}, & n \in G \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

u, v is the ratio factor. The value of u lies between 0 to 1 and $v = 1 - u$.

Sensors in a WSN consume energy while transmitting and receiving data [20]. Energy consumption is the highest for data transmission. Because sensing and processing have no effect on routing, we solely examine energy usage for communication reasons in this article. Using the energy (radio) model presented, k bits of data are transmitted with the following energy consumption:

$$\begin{aligned} E_{TX}(k, d) &= E_{\text{TX-elec}}(k) + E_{\text{TX-amp}}(k, d) \\ &= \begin{cases} E_{\text{elec}} * k + \epsilon_f * k * d^2, & d < d_0 \\ E_{\text{elec}} * k + \epsilon_m * k * d^4, & d > d_0 \end{cases} \end{aligned} \quad (10)$$

where $d_0 = \sqrt{\frac{\epsilon_f}{\epsilon_m}}$ power loss models are determined by threshold distances.

When calculating energy consumption based on the free space model (d^2 power loss); distance between transmitter and receiver less than a threshold (d_0), is considered; otherwise, multipath fading is considered (d^4 power loss).

k bits of data require the following amount of energy to be received:

$$E_{\text{Rx}}(k) = E_{\text{elec}} * k \quad (11)$$

There are rounds in the WSN. The sink detects and reports events in each round.

4 Result Analysis and Discussion

Our novel EECT-IOT-HWSN clustering approach based on modified threshold has the advantage of having few parameters, such as ADV-LEACH1 approach for homogeneous and heterogeneous that provide a desirable performance for IOT-HWSNs, including low energy consumption, uniform energy distribution, and low delay. Of the total number of SNs, 20% are advance SNs, 30% are super SNs, and 50% are normal SNs. There are two BS positions (50,50) and (50,150). As given in Table 1,

Table 1 Initial parameter

Parameters	Value
No. of round ®	2500
P	0.2
E_{TX} or E_{RX}	50nj/bit
ϵ_f	10pj/bit/m ²
E_{elec}	5nj/bit/message
ϵ_m	0.0013pj/bit/4
No. of sensor (N)	200
Packet size	4000 bits

the initial parameters used to simulate EEECT-IOT-HWSN are those specified in the proposal. We compare EEECT-IOT-HWSN with ADV-LEACH1 (HOMO) and ADV-LEACH1 (HETRO) to prove the validity of the proposed clustering protocol for IOT-HWSNs. MATLAB 2020a on Windows 10 is used to implement the proposed technique.

The residual energy is shown in Fig. 2 as a function of the number of rounds. These SNs are dead nodes since their residual energy is zero. As per the Fig. 2 the proposed EEECT-IOT-HWSN technique shows the higher residual energy as likened to the ADV-LEACH1 approach for homogeneous and heterogeneous. Cluster heads are more likely to be formed by SNs with a high residual energy. Homogeneous networks have similar types and energies to SNs. There is a lower chance that ADV-LEACH1 (HOMO) will become the cluster head because it has less energy. In comparison with the ADV-LEACH1 approach with homogeneous and heterogeneous techniques, the proposed EEECT-IOT-HWSN technique shows better performance.

EEECT-IOT-HWSN analyzes the number of rounds to determine whether SNs are alive or dead. Based on the number of rounds, Fig. 3 demonstrates the number of dead/alive SNs. Compared with ADV-LEACH1 (HETRO), the performance of ADV-LEACH1 (HOMO) decreases significantly after 1500 rounds. With EEECT-IOT-HWSN, the number of SNs is higher than the ADV-LEACH1 approach with homogeneous and heterogeneous networks. A chart of dead SNs according to the number of rounds can be found in Fig. 3 (right side). The ADV-LEACH1 (HOMO) shows the all SNs are dead after 2000 rounds. The proposed EEECT-IOT-HWSN and ADV-LEACH1 (HETRO) technique shows the node after 2000 rounds. ADV-LEACH1 approach with homogeneous and heterogeneous is similar in performance with the proposed technique.

CH selection significantly influences EEECT-IOT-HWSN energy efficiency. Modified CH selection thresholds and increased rounds are used in the proposed enhancement. In terms of network lifetime, IOT-HWSNs are stable. The first dead node (FND) and last dead node (LND) of the proposed technique are shown in Fig. 4 according to the different positions of the BS. The proposed EEECT-IOT-HWSN technique outperforms over ADV-LEACH1 approach with homogeneous and heterogeneous in terms of the FND metrics. A dead SN analysis is shown in

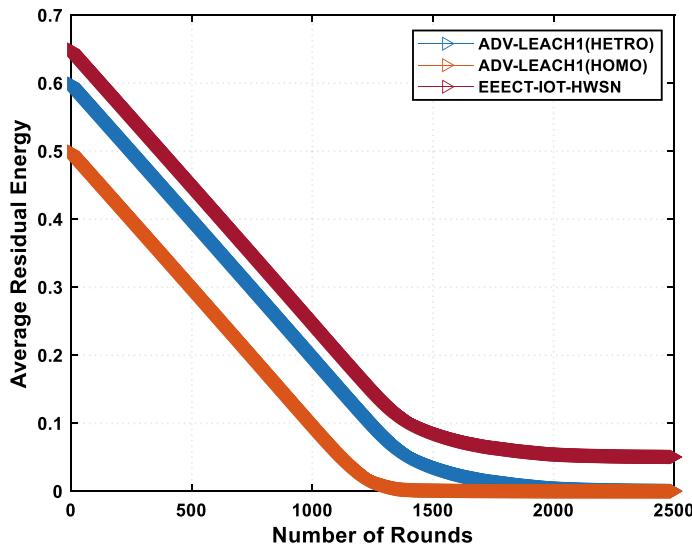


Fig. 2 Average residual energy versus number of rounds

Fig. 4 for BS located inside the network or outside the network, based on the position of the BS. The quantity of dead SNs analysis is designed at 2500 rounds. The proposed EEECT-IOT-HWSN technique shows the 190 and 197 dead SNs for BS center and outside positioned, respectively, as mentioned in the Table 2. In terms of network lifetime analysis, EEECT-IOT-HWSN technique has more SNs alive after 2500 rounds compared with ADV-LEACH1 approach with homogeneous and heterogeneous.

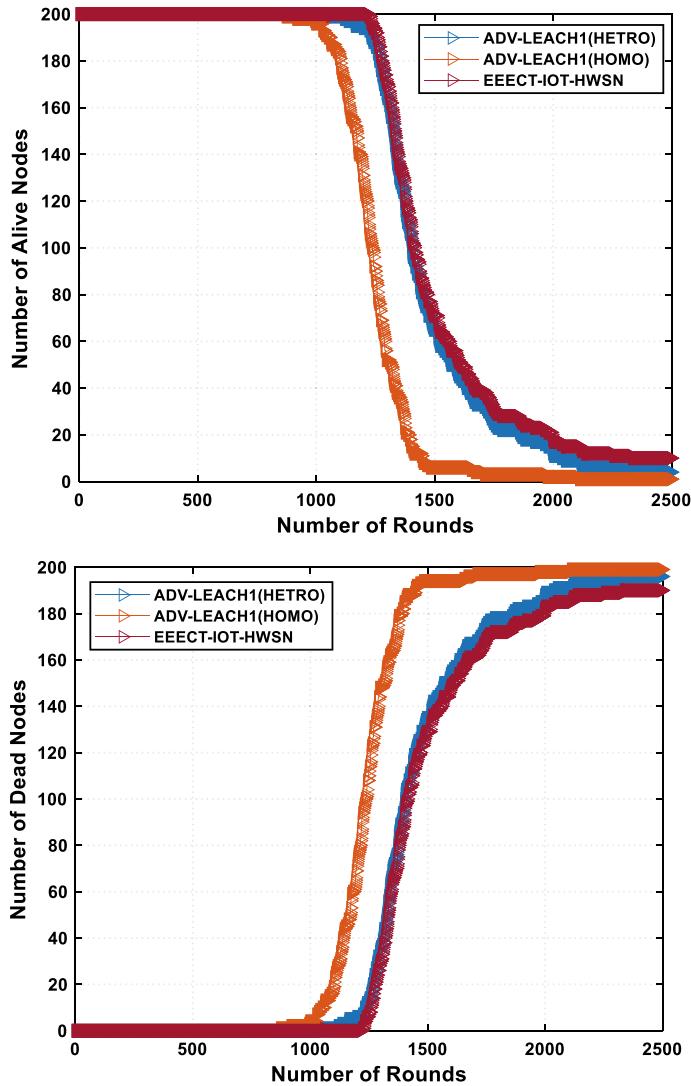


Fig. 3 Performance comparison of number of alive/dead SNs (right) versus number of nodes

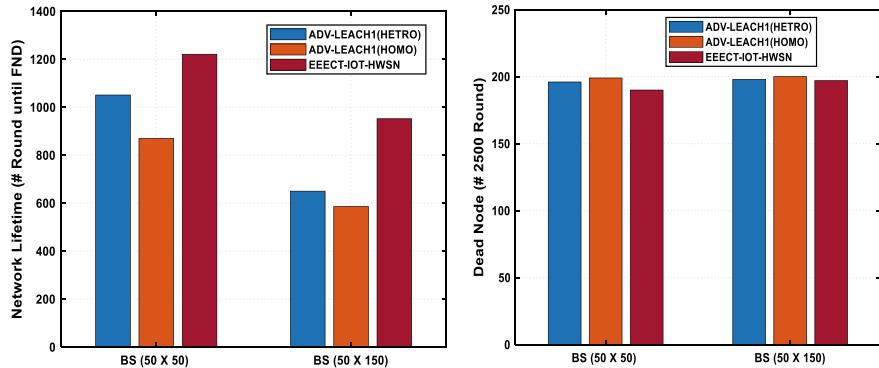


Fig. 4 Performance of the network lifetime (FND) and dead SNs analysis at 2500 rounds for different positions of the BS

Table 2 Comparative analysis

Algorithms	BS position (50 X 50)		BS position (50 X 150)	
	Network lifetime (#FND)	Dead node (# 2500 rounds)	Network lifetime (#FND)	Dead node (# 2500 rounds)
ADV-LEACH1(HETRO)	1051	196	650	198
ADV-LEACH1(HOMO)	870	199	586	200
EEECT-IOT-HWSN	1221	190	952	197

5 Conclusion

A heterogeneous wireless sensor network (HWSN) based on Internet of Things (IoT) technology has emerged as a key technology for developing a range of human-centric applications. The most crucial resource in an IoT-based HWSN is energy, just as in a WSN. In this paper, design a novel EEECT-IOT-HWSN technique for the three-tier heterogeneous networks. The cluster head selection process in the EEECT-IOT-HWSN approach uses a modified threshold calculation based on the energy and distance of the SNs combined. In the monitor region, all three kinds of SNs are placed at random. Validation of the EEECT-IOT-HWSN model's performance is done on the basis of many factors, including residual energy, alive/dead SNs, and network longevity. The simulation results of the proposed model are compared with the latest ADV-LEACH1 (HETRO) and ADV-LEACH1 (HOMO) technique. The performance of the proposed model shows the higher residual energy, less dead SNs, and higher network lifetime when compared with the above said technique.

References

1. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. *Comput Netw* 54(15):2787–2805
2. Muhammad Z, Saxena N, Qureshi IM, Ahn CW (2017) Hybrid artificial bee colony algorithm for an energy efficient internet of things based on wireless sensor network. *IETE Tech Rev* 34(sup1):39–51
3. Yetgin H, Cheung KTK, El-Hajjar M, Hanzo LH (2017) A survey of network lifetime maximization techniques in wireless sensor networks. *IEEE Commun Surv Tutorials* 19(2):828–854
4. Dezfooli B, Amirtharaj I, Li C-CC (2018) EMPIOT: an energy measurement platform for wireless IoT devices. *J Netw Comput Appl* 121:135–148
5. Jamil F, Khan FZ (2020) Multi-criteria-based mobile hotspot selection in IoT-based highly dense network. *Wireless Pers Commun* 112:1689–1704
6. Hussain N, Rani P (2020) Comparative studied based on attack resilient and efficient protocol with intrusion detection system based on deep neural network for vehicular system security. In: *Distributed artificial intelligence*. CRC Press, pp 217–236
7. Hussain N, Rani P, Chouhan, H., & Gaur, U. S. (2022). Cyber security and privacy of connected and automated vehicles (CAVs)-based federated learning: Challenges, opportunities, and open issues. *Federated Learning for IoT Applications*, pp 169–183
8. Rani P, Hussain N, Khan RAH, Sharma Y, Shukla PK (2021) Vehicular intelligence system: time-based vehicle next location prediction in software-defined internet of vehicles (SDN-IOV) for the smart cities. In: *Intelligence of Things: AI-IoT Based Critical-Applications and Innovations*, pp 35–54
9. Devarapalli V, Wakikawa R, Petrescu A, Thubert P (2005) Network mobility (NEMO) basic support protocol
10. Reddy V, Gayathri P (2019) Integration of internet of things with wireless sensor network. *Int J Electr Comput Eng* 9(1):439
11. John A, Rajput A, Babu KV (2017) Dynamic cluster head selection in wireless sensor network for Internet of Things applications. In: *2017 International Conference on Innovations in Electrical, Electronics, Instrumentation and Media Technology (ICEEIMT)*, pp 45–48
12. Mishra J, Bagga J, Choubey S, Choubey A, Gupta K (2021) Performance evaluation of cluster-based routing protocol used in wireless internet-of-things sensor networks. In: *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, pp 1–10
13. Wang, N., & Zhu, H. (2012). An energy efficient algrithm based on leach protocol. *2012 International Conference on Computer Science and Electronics Engineering*, 2, 339–342.
14. Pal R, Sindhu R, Sharma AK (2013) SEP-E (RCH): enhanced stable election protocol based on redundant cluster head selection for HWSNs. In: *Quality, reliability, security and robustness in heterogeneous networks: 9th International Conference, QShine 2013, Greader Noida, India, January 11–12, 2013, Revised Selected Papers* 9, pp 104–114
15. Arumugam GS, Ponnuchamy T (2015) EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN. *EURASIP J Wirel Commun Netw* 2015(1):1–9
16. Junping H, Yuhui J, Liang D (2008) A time-based cluster-head selection algorithm for LEACH. In: *2008 IEEE Symposium on Computers and Communications*, pp 1172–1176
17. Tong M, Tang M (2010) LEACH-B: an improved LEACH protocol for wireless sensor network. In: *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pp 1–4
18. Batra PK, Kant K (2016) LEACH-MAC: a new cluster head selection algorithm for Wireless Sensor Networks. *Wireless Netw* 22:49–60
19. Kumar N, Rani P, Kumar V, Verma PK, Koundal D (2023) TEEECH: three-tier extended energy efficient clustering hierarchy protocol for heterogeneous wireless sensor network. *Expert Syst Appl* 216:119448

20. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, vol 2, 10 pp
21. Islam, M. M., Matin, M. A., & Mondol, T. K. (2012). *Extended Stable Election Protocol (SEP) for three-level hierarchical clustered heterogeneous WSN*.
22. Kumar N, Kumar V, Verma PK (2022) A comparative study of the energy-efficient advanced LEACH (ADV-LEACH1) clustering protocols in heterogeneous and homogeneous wireless sensor networks. Cyber Secur Digital Forensics Proc ICCSDF 2021:433–444
23. Kumar N, Rani P, Kumar V, Athawale SV, Koundal D (2022) THWSN: enhanced energy-efficient clustering approach for three-tier heterogeneous wireless sensor networks. IEEE Sens J 22(20):20053–20062
24. Regin R, Obaid AJ, Alenezi A, Arslan F, Gupta AK, Kadhim KH (2021) Node replacement based energy optimization using enhanced salp swarm algorithm (Es2a) in wireless sensor networks. J Eng Sci Technol 16(3):2487–2501
25. Abdulreda A, Obaid A (2022) A landscape view of deepfake techniques and detection methods. Int J Nonlin Anal Appl 13(1):745–755. <https://doi.org/10.22075/ijnaa.2022.5580>
26. Hmeed AR, Hammad JA, Obaid AJ (2023) Enhanced quality of service (QOS) for MANET routing protocol using a distance routing effect algorithm for mobility (DREAM). Int J Intell Syst Appl Eng 11(4s):188–193

IoT-Based Smart System for Fire Detection in Forests



M. A. Archana, T. Dinesh Kumar, K. Umapathy, S. Omkumar,
S. Prabakaran, N. C. A. Boovarahan, C. Parthasarathy,
and Ahmed Hussein Alkhayyat

Abstract A number of fire accidents in forests occur around the globe every year which amount to catastrophes beyond all sorts of comprehensions. Behind this, many houses and lot of trees pose a serious threat to forests grown in an ambient and healthy environment. This paper enunciates an innovative methodology for detection of fires in forests using the concept of Internet of Things (IoT). IoT has gained a lot of attention due to latest advancements in technology. This work is capable of detecting forest fires in early manner by implementing a smart system integrated with temperature, moisture and humidity sensors. The system employs IoT for transmission of data to the authorized person over internet thereby increasing the efficiency of the fire detection process. The detection is intimated by means of an alarm in that specific location. The data is validated using a sensor threshold value thereby achieving maximum reliability for prevention.

Keywords Sensor · Fire detection · Temperature · Humidity · Internet of Things · Microcontroller

M. A. Archana · T. Dinesh Kumar · K. Umapathy (✉) · S. Omkumar · N. C. A. Boovarahan
SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India
e-mail: umapathykannan@gmail.com

S. Prabakaran
P.T. Lee Chengalvaraya Naicker College of Engineering and Technology, Kanchipuram, India

C. Parthasarathy
Vel Tech Multi Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, Chennai, Tamil Nadu, India

A. H. Alkhayyat
Scientific Research Centre of the Islamic University, The Islamic University, Najaf, Iraq

1 Introduction

The forests which spread over the world play a key role in preservation of environment and balancing of ecological system. The parameters such as development of plants and availability of soil nutrient play an indispensable role toward ecological role of fires in forests. The pattern, shape and composition of ecosystems are completely influenced by fires. The fire in forests can be detected practically after it has exploited a particular area; thus, the process of control and suppression is more difficult. Due to fire in forests, unbearable losses will be caused to both environment and ecology. The long-term after effects of forest fires include change in local climatic conditions, global warming and destruction of specific species [1–4].

These forest fires generally occur in remote places which are packed with lot of trees, dry wood and leaves that form an ideal source of fuel. The above elements form an integrated composite material which is very much suitable for ignition of fire initially and an ideal fuel for fire later on. The forest fires will be a continuous threat to human life, infrastructure and ecological systems. Fire is the key issue of environment which causes more damages both economically and ecologically to the human society. The ignition of fire may come from usual activities of day-to-day life such as smoking, high temperature due to hot summer day, concentration of sun light using a broken glass[5]. The wood caught with fire becomes the vital reason for destroying a forested area which can create threat and risk to the people living near the forests with hundreds of kilometers. After ignition, fire grows wider and larger becoming uncontrollable. This may cause great damage to the landscape for a longer period based on the parameters—weather and terrain. Due to this, so many hectares of forests will be demolished [6, 7].

2 Related Works

Timely detection of forest fires can reduce time, effort, threat and cost for fire-fighting. The major reasons for forest fires are temperature of the atmosphere and humidity of the environment. Hence, the environment must be analyzed both under normal and fire situations as a part of an effective detection system. Hence, it is very much essential to detect fire by employing early systems of detection. There are various methods available for fire detection. One typical example is Tower surveillance system which uses the satellite images for detection. But the drawback with this system is requirement of large number of trained personnel for detection due to complexity in forest infrastructure. Hence, a smart detection system is very much required for early detection and prevention of fire spread thereby protecting land and ecosystems [8, 9].

IoT is a technique by which information is sensed from various objects and transmitted to the other end over the internet. The transmission of sensitive information about forest fires to the central station is highly important. This is a challenging task

for the researchers and scientist [10, 11]. The conventional method of fire protection uses either mechanical or human observers for tracking the environment. Some of above methods include forecasting of fire weather, moisture towers, detection of optical smokes, lightening detectors, etc. In case of fire watch tower method, human beings are used as watch towers to monitor a particular location consistently and give alerts to the concerned if fire occurs. The accuracy of human observations depends upon various parameters such as time, climate, location and fatigue of operator [12–14].

The fire detection by satellite-based systems is not that much feasible due to poor pixel density in pictures taken. Real-time detection of fire is not reliable because detection is done after fire spreading appropriately. Moreover, the cost of these detection systems is very high. The typical methods used for fire detection depend on testing of air transparency, temperature and particle sampling. Smoke sensors and alarm devices are integrated with fire detection systems [15–17]. In existing systems, there is no room for intelligent detection of fire with appropriate accuracy and reliability.

3 Materials and Methods

The proposed work is to develop technique of data transfer which can be used to observe the trees in a forested area in an effective and controlled manner. The detection of fire can be done precisely and the respective fire units can be notified appropriately at the earliest time. The controller used here is Node MCU which includes programmable pins, Wi-Fi facility and powered by a USB port. It is an open-source system in which editing, altering and building can be done easily. It employs TCP/IP protocol for transition of data. It uses a flash memory called SPIFFS [18–21].

The sensors are very much essential for measurement of temperature with fire detection system. A temperature sensor is used to evaluate conditions of environment, data sensing, controlling temperature and checking the occurrence of heating. The output of the sensor is used to switch off the detector alarm. Moreover, the humidity and ultrasonic sensors are employed to monitor the conditions of the forest in a precise manner. The purpose of ultrasonic sensor is to evaluate the distance of the forest trees accurately. The smart fire detection system is shown in Fig. 1.

Humidity sensor (DHT11) is used for measurement of humidity in the forest environment and gives an alert message whenever there is an increase in the humidity factor which ignites a forest fire possibly. A soil moisture sensor is employed to check soil level of ground in forests as the percentage of dryness may amount to probable fire in forests. The third sensor used for the system is Light-Dependent Resistant (LDR) sensor. Due to its light sensing property, the fire detector changes whenever more intensity of light is observed. This will increase the resistance thereby recording sufficient values of voltages.

Figure 2 depicts the flowchart of smart fire detection system. The temperature sensor collects information regarding temperature prevailing in and around the trees

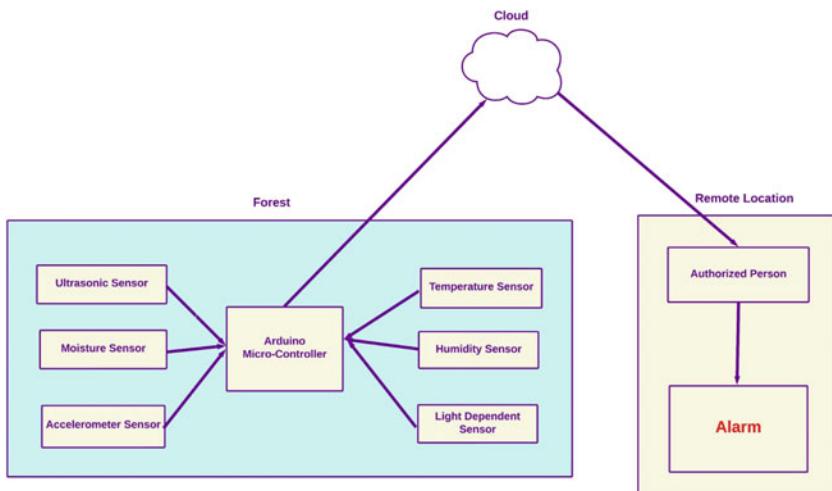


Fig. 1 Block diagram of smart fire detection system

situated in forest areas. The values of temperature are checked regularly by this sensor which is in turn transmitted to the microcontroller. This controller informs fire alarm system and sends information to cloud API with respect to moisture in the soil, temperature and humidity.

Using IoT, the server receives information from the cloud and forwards to the concerned personnel by means of email notifications. Thus, IoT is used for collecting and processing the data from the real world. In the future, each and every physical gadget is going to be as source of data which will be a source of economic value [22–25]. This system provides sufficient accuracy and compactness by employing less expensive and small size sensors.

The selection of sensors depends upon various parameters such as accuracy, resolution and response time. The dual-state button algorithm is implemented for executing ON/OFF operations.

4 Results and Discussion

As per schematic diagram shown in Fig. 1, hardware connections are given appropriately and the program is uploaded into the controller using Arduino IDE 1.8 version software. Figure 3 shows hardware connections of system which is used to detect forest fire. Figure 4 shows the system model during ignition of fire, and on detection of fire, it sends the detected information to the authorized person by means of cloud API mail, message and turns on the Alarm as shown in Fig. 5.

The prospective values of temperature, humidity and soil moisture are then received in an email by the authorized person for further action.

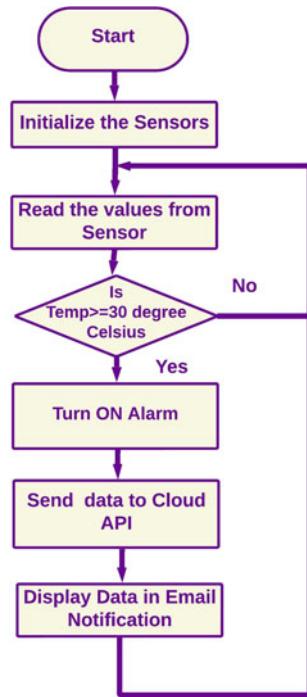


Fig. 2 Flowchart for smart fire detection system

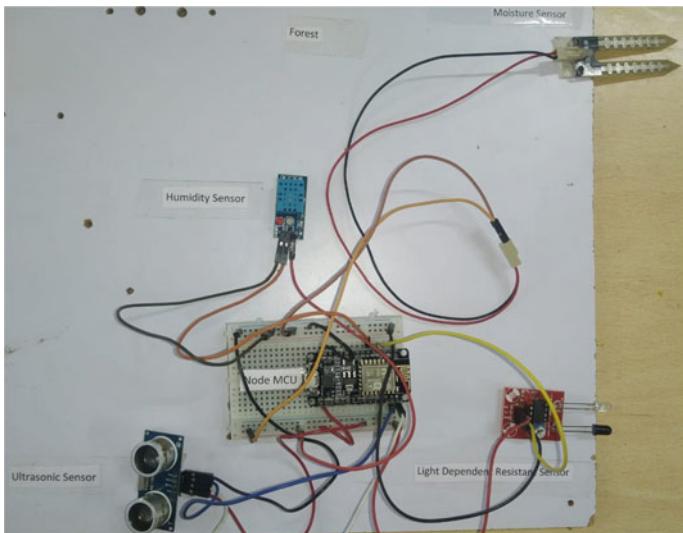


Fig. 3 Hardware connections of the proposed system

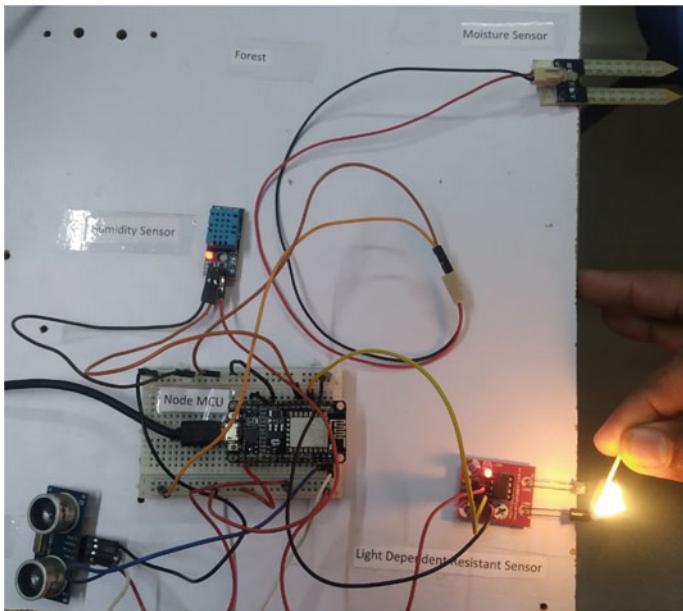
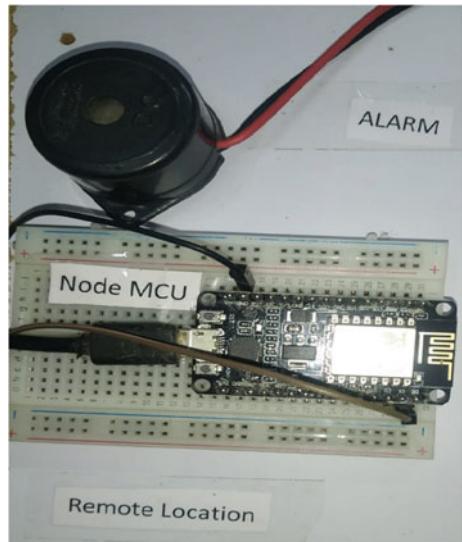


Fig. 4 System model during detection of fire

Fig. 5 Remote location turning ON the alarm after receiving email notification



5 Conclusion

This paper enunciates an early-stage forest fire detection system integrated with concept of IoT. Detection and localization of fire in accurate and timely manner are the key objectives of fire detection. This proposal was to develop a reliable smart fire detection system which can be implemented in forests in order to prevent fire accidents. Additionally, this system focuses on analyzing temperature, humidity and content of soil moisture in forest area. This data was forwarded to the concerned by means of email notifications for daily reporting. By implementing the proposed systems, number of forest fires will be reduced at the global level and environment will be protected with detection and forecasting of fires in time. In the future, system model can be integrated with respective sensors using least amount of energy and the concept of wireless sensor networks can be implemented, especially for large forest areas. To enhance the detection process further, artificial intelligence technique can be implemented with trustable machine learning algorithm.

References

1. Varela N, Díaz-Martinez Jorge L, Ospino A, Zelaya NAL (2020) Wireless sensor network for forest fire detection. *Proced Comput Sci* 175:435–440
2. Giglio L, Kendall J, Justice CO (1999) Evaluation of global fire detection algorithms using simulated AVHRR data. *Int J Remote Sens* 20(10):1947–1955
3. Gonzalez JR, Palahi M, Trasobares A, Pukkala T (2006) A fire probability model for forest stands in Catalonia (north-east Spain). *Ann For Sci* 63(2):169–176
4. Solobera J (2010) Detecting forest fires using wireless sensor networks with Wasp mote. *Libelium World*
5. Nakau K, Fukuda M, Kushida K, Hayasaka H, Kimura K, Tani H (2006) Forest fire detection based on MODIS satellite imagery, and comparison of NOAA satellite imagery with fire fighters information. *IARC/JAXA Terr Team Workshop*, 18–23
6. Sarwar B, Sarwar Bajwa I, Jamil N, Ramzan S, Sarwar N (2019) An intelligent fire warning application using IoT and an adaptive neuro-fuzzy inference system. *Sensors* 19(14):3150
7. Alkhatab AAA (2014) A review on forest fire detection techniques. *Int J Distribut Sens Netw* 10(3):597368
8. Dauda MS, Toro US (2020) Arduino based fire detection and control system. *Int J Eng App Sci Technol (IJEAST)* 4(11):447–453
9. Shukla A, Kumar S, Singh H (2020) Fault tolerance based load balancing approach for web resources in cloud environment. *Int Arab J Inf Technol* 17(2):225–232
10. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. *Comput Netw Int J Comput Telecommun Netw* 54(15):2787–2805
11. Vining J, Merrick MS (2008) The influence of proximity to a national forest on emotions and fire-management decisions. *Environ Manage* 41(2):155–167
12. Kuhrt E, Knollenberg J, Mertens V (2001) An automatic early warning system for forest fires. *Ann Burns Fire Disasters* 14(3):151–154
13. Yu L, Wang N, Meng X (2005) Real-time forest fire detection with wireless sensor networks. *Proc Int Conf Wireless Commun Netw Mob Comput Maui, HI*, 2:1214–1217
14. Mangayarkarasi T, Umapathy K, Sivagami A, and Subith D (2021) An IoT based safe assembly point alert system. *J Phys Conf Ser* 1964(7):1–4

15. Toreyin BU, Dedeoglu Y, Cetin AE (2005) Flame detection in video using hidden Markov models. In: IEEE International Conference on Image Processing, Genova, Italy, pp 1230–1233
16. Li Z, Nadon S, Cihlar J (2000) Satellite detection of Canadian boreal forest fires: development and application of the algorithm. *Int J Remote Sens* 21(16):3057–3069
17. Lynham TJ, Dull CW, Singh A (2022) Requirements for space-based observations in fire management: a report by the Wildland Fire Hazard Team, Committee on Earth Observation Satellites (CEOS) Disaster Management Support Group (DMSG). In: IEEE International Geo-science and Remote Sensing Symposium, vol 2, pp 762–764
18. Dinesh Kumar T, Archana MA (2022) Fundamentals of Internet of Things and its applications. Alpha International Publication
19. Sifakis, NI, Iossifidis C, Kontoes C, Keramitsoglou I (2011) Wildfire detection and tracking over Greece using MSG-SEVIRI satellite data. *Remote Sens* 3(3):524–538
20. Sahin YG, Ince T (2009) Early forest fire detection using radio-acoustic sounding system. *Sensors* 9(3):1485–1498
21. Singh H, Shukla A, Kumar S (2021) IoT based forest fire detection system in cloud paradigm. *IOP Conf Ser Mater Sci Eng* 1022(1):012068
22. Umapathy K, Sai Swaroop V, Viswam P, Balaswami Sairaja T (2020) Counterfeit bank note detecting system. *Int J Sci Technol Res (IJSTR)* 9:1033–1035
23. Dinesh Kumar T, Archana MA, Umapathy K, Gayathri G, Bharathvaja V, Anandhi B (2023) RFID based smart electronic locking system for electric cycles. In: IEEE Xplore, Fourth IEEE International Conference on Electronics and Sustainable Communication Systems, 1CESC 2023 Proceedings, Coimbatore, pp 76–81
24. Umapathy K, Omkumar S, Muthukumaran D, Chandramohan S, Sivakumar M (2023) Autonomous health care robot. *Lect Notes Netw Syst* 617:227–233
25. Umapathy K, Omkumar S, Muthukumaran D, Chandramohan S, Sivakumar M (2023) Thingspeak based garbage monitoring and collecting system. *Lect Notes Netw Syst* 617:235–242

Machine Learning Approach to Lung Cancer Survivability Analysis



Srichandana Abbineni , K. Eswara Rao , Rella Usha Rani ,
P. Ila Chandana Kumari, and S. Swarajya Lakshmi

Abstract The majority of people in the current atmospheric conditions are affected by lung cancer disease. The analysis of respiratory illness offers a captivating and dynamic research space with far-reaching implications for human health. A diagnostics like this can only assist in reducing the likelihood of obtaining human life in jeopardy by initial detection of metastatic disease to address this problem. Lung cancer is the leading cause of cancer death worldwide, so different algorithms have been used to forecast the prognosis of cancer patients. Because of this, patients with lung cancer are living longer on average. When making predictions, the logistic regression assessment method is more accurate than that of other methods. This report examines two additional different approaches to machine learning for forecasting a lung participant's life expectancy, including linear discriminant analysis (LDA), random forest (RF), and artificial neural networks (ANN). In order to increase success rates,

The original version of the chapter has been revised: The author's "K. Eswara Rao and Rella Usha Rani" affiliations has been updated. A correction to this chapter can be found at
https://doi.org/10.1007/978-981-99-9562-2_67

S. Abbineni ()

Department of CSE (DS), CVR College of Engineering, Hyderabad, India
e-mail: chandu.abb@gmail.com

R. U. Rani

Department of CSE (AI&ML), CVR College of Engineering, Hyderabad, India
e-mail: teaching.usha@gmail.com

K. E. Rao

Department of CSE, Aditya Institute of Technology and Management, Tekkali, India
e-mail: eswarkoppala@gmail.com

P. I. C. Kumari

CSE Department, Hyderabad Institute of Technology and Management, Hyderabad, India
e-mail: ilachandana@gmail.com

S. S. Lakshmi

Department of CSE, KMIT, Hyderabad, India
e-mail: swarajya15.s@gmail.com

various algorithms were tested. The primary goal of this is to evaluate the accuracy of classification methodologies to develop a melanoma statistical method and a resilience analysis. The correctness, accuracy, recall, and selectivity of the numerous models' performances are assessed and compared. In this enquiry, linear discriminant analysis will perform the best among the three algorithms.

Keywords Lung cancer disease prediction · Machine learning (ML) · Survivability · Linear discriminant

1 Introduction

The main cause of lung cancer is passive smoking. The healthy tissue [1] is harmed by cigarette that enters the lungs. Lung cancer in people who smoke can be brought on by thoron irradiance, resale fumes, air pollution, or other factors. Another cause of lung disease is heredity. Breast cancer (malignant growth) can be prevented in the initial phases, despite the fact that it is difficult to make a diagnosis. One of the threatening cancer types that is frequently found is lung disease. There are as of now more than a million new cases of cancer reported. In addition, grades are given to cancer according to its level. The number of people with lung cancer is rising rapidly, and by the disease, nearly everyone has been affected negatively. In addition, lymph node swelling, nervous system, and jaundice are the issues additionally. During the diagnosis of lung cancer, patients face many challenges. Therefore, mechanization in this area may facilitate the pathologist's work while also accelerating the process. Cancer is caused by a variety of other variables in addition to heredity. The increase in lung cancer cases is largely due to the way people live today.

The World Health Organization has identified disease as the leading cause for increased death rate around the world. Hence, the lung cancer is the most studied and given a way to consider it as a diagnosis disease. As a result, increasing awareness and forecasting the initiation of lung disease in its beginning phases can help people take the appropriate preventive measures, lowering the number of individuals killed by lung cancer. This refers to the uncontrolled proliferation of malignant growth inside one or even both respiratory systems, most commonly in respiratory cells. Mutated lymphocytes do not develop into good health respiratory system, degrade rapidly, or bulk up. The work nature of lungs is to send the blood with oxyzen some times expands into a tissue form. "Mild tissue" is defined as vasculature that remains in one location but does not visible to expand. The lymph is distributed to lymph vessels, that also cause lymph nodes in the lung tissue and shoulders to be released. Many disease detection models like "Blood pressure Classification" stands as best model for tunning parameters [2, 3]. Lung cancer frequently spreads to the center of the chest because nearby lymph endpoints in the respiratory system are located there.

Lung cancer is typically classified into two types: non-small squamous cells and small cells. Based on their number of clusters, these would be assigned to individual

kinds of cancer. Breast cancer can be classified into four levels: Level I, Level II, Level III, and Level IV. The level is determined by the size of the tumor, the extent of cancer spread to the lymph nodes, and whether the cancer has spread to other parts of the body. Inflammatory breast cancer is a rare type of breast cancer that is classified as Level IV. It is characterized by redness, swelling, and warmth in the breast, as well as the presence of cancer cells in the lymph nodes.

A most lethal risky stroma tends to spread all through body via the interstitial fluid or lymph system. The term “cancer” refers to a cancer that has progressed well beyond its original location to other tissues within the body. Tertiary cancerous cells can propagate anywhere in the body and finally enters into lungs where as lung parenchyma cancer cell will starts its movement flow from lungs to other parts of body. There are many different types of cancer, and not all of them are treated in the same way. Individuals with respiratory illnesses including chronic bronchitis and a history of chest issues have a higher risk to develop cancer. Smoking cigarettes, etc. are the most common risks for lung disease in Indian men; even so, cigarette smoke is less prevalent in Indian women, indicating there are additional variables that contribute to lung cancer. Improved knowledge of risk variables can aid in the prevention of cancer. The key to improving life expectancies is early diagnosis utilizing machine learning, and if we would use this to start making the specific diagnostic process much more efficient for radiographers, it will be a significant step toward to the objectives of improving early diagnosis.

Numerous methods are employed to increase the life expectancy of cancer patients, including consistent physician join, tracking the expansion of lung tumors, and providing people with rehabilitation services. Annually, the incidence rate of lung cancer in men and women in the USA has reduced. Depending on the level of disease, there are clinical strategies and efforts to fund clients likelihood of living. Collection categorization is a critical component of both operational digital business apps and traditional machine learning problems. Machine learning technique is used to determine whether a particular set of traits pertains to an individual with cancer or not. Machine learning is often used in data classification, prediction, and even cluster analysis. It is essentially the schooling of a prototype, which is used to complete a task. This model was trained stuff new in machine learning techniques. As the overall survival of lung cancers rises, many methods are being proposed to estimate their preservation. Among such methodologies and algorithms, the LDA outperformed the others.

2 Literature Survey

Zahid and colleagues [1] developed a linear diagnostic model to estimate malignancy risks for nonrecurring cases and disease recurrence periods. On a dataset of 569 patients, the model was evaluated using a cross-validation technique, providing an accuracy of 97.5%. A linear regression approach was used to create the model. The model’s input variables were patient age, gender, tumor size, and lymph node

status. The patient's risk of (3) is a potential method for anticipating malignant hazards in nonrecurring situations and the recurring time period of malignant recurrence. According to the findings of this investigation, the linear diagnostic model was proposed by Zahid et al. (202 illnesses). The model is accurate and can be used to identify patients who are at high risk of malignant recurrence.

Lee et al. [4] created a machine learning model to predict malignant recurrence in patients with breast cancer. The model was trained on a dataset of 1,010 patients and had an accuracy of 92.4% in predicting malignant recurrence. The model was built using a random forest technique with input factors such as patient age, tumor size, lymph node status, and hormone receptor status.

Gupta et al. [5] also created a machine learning model to predict malignant recurrence in patients with breast cancer. The model was trained on a dataset of 1,5758 patients and predicted malignant recurrence with a 93.3% accuracy. The model employed a similar set of clinical characteristics as input variables as the model created by Lee et al. [4].

Wang et al. [6] used a deep learning method to construct a machine learning model to predict malignant recurrence in breast cancer patients. The algorithm was trained on a dataset of 2,149 patients and predicted malignant recurrence with a 94.2% accuracy. As input variables, the model employed a mix of clinical characteristics and gene expression data.

Ryu and colleagues [7] A machine learning (ML) approach was applied to increase the survival rate of patients with Spinal Ependymoma. Yutong Xie et al. [8] employed a multi-view information-based tool synergy (MV-KBC) depth algorithm to identify malignant lumps from healthy chest radiography utilizing chest computed tomography (CT) data in this evaluation. Nine KBC threads were used to train the model. The LIDC-IDRI statistical model was used in the evaluation, which was compared to five recent theories of categorization.

Gray-level co-event matrices (GLCM) on artificial neural systems have been suggested for cancer diagnosis by Anifah et al. [9] The tumor image archive collection, which consists of CT scans, is utilized to collect pulmonary information. The feature extraction of images, edge detection, and cancer progression diagnosis utilizing a three-layer neural network backpropagation. The results showed that the framework can more reliably discriminate between healthy lung tissue and lung cancer. Bray et al. [10] present and discuss the different types of cancer that are most common in different parts of the world. Lynch et al. [11, 12] automated training and testing classifiers have been used to forecast the lifespan of lung cancers.

Ranjan Baitharu et al. [13, 14] Women and men are both at risk of dying from pulmonary cancer which is caused by unchecked cell cycle. In the process of Knowledge Discovery in Data (KDD), clustering is a crucial step. There are numerous possible benefits for it. The learning sample has a significant impact on how well classifications work. As a result, categorization systems better in terms of forecast or diagnostic quality, need less computational power to develop models, since they pick up knowledge more quickly and are easier to comprehend. Using information on lung cancer in various settings, a quantitative examination of data categorization quality is

offered. Comparison of algorithms is based on prediction accuracy numerical indice comparison.

Krishnaiah et al. [15, 16] offered a strategy for almost detecting and correctly diagnosing the illness, aiding the physician in preserving the service user. The chance of someone developing emphysema can be predicted utilizing common bowel cancer signals such age, race, whistling, chest tightness, and pain in the neck, chest, or arms.

Cruz et al. [17] Techniques for machine learning can be used to significantly boost the accuracy of detecting disease risk, recurring, and death, according to the better quality and tested studies. It is also clear that automation is assisting in bettering the underlying grasp of human cancer and recurrence at a more deep level.

Sujitha et al. [18] are using a classifier to categorize nodules as invasive carcinoma, as well as the amount of cancer.

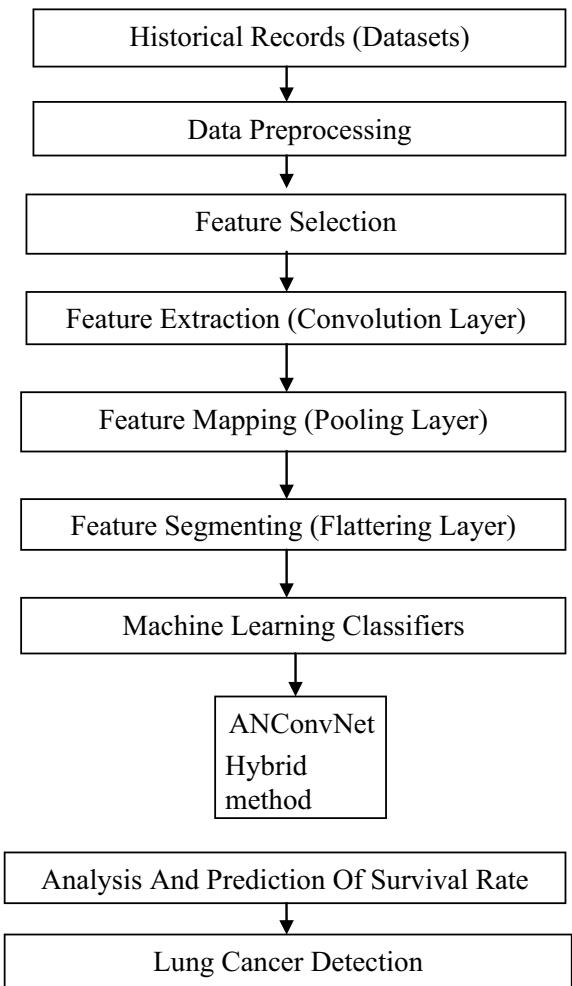
Dr. S. Senthil et al. [19] According to this point of view, lung carcinoma is caused by the spread of malignant tumors in the pleural space, and it is important to anticipate and detect emphysema ahead of time by utilizing optimum conditions neuromorphic attributes. Initially, the alveolar repository is collected and supplied into the framework. The characteristics of the pictures supplied as input are obtained using an optimizer, and then an artificial neural descriptor is utilized to characterize the specified frequency of the input images as cancer cells or being growths. The job removal is a component of pattern recognition algorithms that are used on feedback data to collect critical qualities that are more succinct, repaired, and receive carcinoma information to interpret the patient's symptoms.

3 Lung Cancer Survivability Analysis and Prediction Model

Inside the evaluation function, the device version has three phases. Those are information series, information training, and information evaluation. Each section includes the sections. The presented model for lung cancer survivability analysis and prediction is shown in Fig. 1.

The Dataset utilised in this model implementation is from licensed data repository and the input belongs to numerical continuous data. Information preparation involves cleansing missing records after facts have been accrued from legal repositories. Aside from the real-world data, some capabilities or cases are missing due to a variety of reasons, including inability to load the data, inadequate patient follow-up, or unexpected patient death. In order to limit the amount of inaccurate forecasts or statistical classes, the proper selection for missing values must be made. The function selection approach comprises extracting the preferred functions and discarding the remainder, which is a useful strategy to avoid excessive complications in future calculations. Patients with lung cancer provided data for the study. An input dataset is made up of six parameters. The system will be able to forecast lung cancer as a

Fig. 1 Presented model for lung cancer survivability analysis and prediction



consequence of algorithms below the input parameters. The estimate's precision is based on the Kaggle `lung_cancer_examples` dataset.

The applied technique is the approach of correlation matrix so that you can select out the functions which make a contribution to the output function to increase the accuracy and teach quicker. Because of a well-defined relation, the results from complement framework should increase from -1 to 1 but not 0 . In this manner, it is far less difficult to decide the essential functions related to their relationship esteems with each other function. The feature extraction includes extracting the prominent distinguished features along with their normalised values for supporting similarity index identification. The process of partitioning an information and also used to discover gadgets and limitations viz. lines, curves, and many others in a photograph

is known as segmentation. Total pixels in an area or an object share a similar attribute. Threshold is one of the easiest techniques.

Machine learning classification algorithms are employed to anticipate lung cancer at its earliest steps in order to preserve lives and boost resilience. The main goal of this proposed technique is to significantly boost phase's precision by utilizing machine learning techniques like hybrid model (ANN + CNN) coined as ANConvNet.

Discriminatory practices assessment differs from material characterization in that it is not an enough that; rather, it necessitates the difference of study variables (also known as conditional variables). Woods at irregular intervals can lead to fluctuations in the tuning. It is both a supervised learning algorithm and a versatile framework used for regression and classification analyses, resembling outfitts. The prototype is also well-known for its estimations that involve the creation of choice foliage. It is also a powerful procedure that generates N tree structure. Each clustering algorithm acquired is constructed from such an arbitrary subset of the learning set and characteristics. The aggregated vote totals of the N judgment trees formed ascertain the production class. Supervised learning, ANNs are trained on labeled data. They learn to map input to output by adjusting weights through trial and error. Unsupervised learning ANNs are trained on unlabeled data. They find patterns by adjusting weights and clustering data.

Artificial neural networks is an approach that works by incorporating foundational library functions such as keras (a library for neural systems) and machine learning as the server side, which helps make nerve cells simpler and faster. By using common devices, learning algorithms (ANNs and CNNs) are made to act like a network of linked neurons. It includes numerous levels that are organized in sequence, but every level contains numerous input signals, or subunits, formats, and perpetuates it to one or so more hidden layers where it studies just the conclusion and generates a conclusion for the activation function. Pooling layer and downsampling are the two factors used for the intake, hiding, and convolutional, respectively. Calculating the gap between the real and planned, the mistake is then backpropagated till it may be as low as possible. The star's procedure is selecting the inputs and results that characterize its architecture, initializing the values with the "Adam" algorithm and filters at arbitrary, and then exposing the net to a test dataset. Testing the program and continuing with the error correction are the next steps. Recalculate the values of nodes and the source block if the halt condition yields a negative outcome, which will cause the margins of error to flow down to hidden neurons later.

If the halt requirements are met, the fit is different to a testing dataset, in which the model is expected to approximate the outcome and assess the outcome. The assessment and prediction of life expectancy is ascertained after trying to apply these classification models. Utilizing machine learning algorithms, these achievement analyses detect lung disease and thus boost the rate of survival. ANConvNet models can accurately predict the exact level and type of cancer by analyzing medical images.

4 Examine the Outcomes

The result analysis of the presented model for lung cancer survivability analysis and prediction using ML approach is shown by using the following definitions for true positive (TP), true negative (TN), false negative (FN), and false positive (FP), the performance of the described model is assessed:

True Positive: TP is the weighted sum of genuinely good, correctly classified high clinical cases.

True Negative: TN is the weighted sum of genuinely adverse, correctly classified low r-squared cases. False positives are examples of re-right outcomes that are mistakenly classified as such and are not genuinely good. **False Negative:** FN is the maximum sum of false negatives or occurrences of doubts that are not based on fact.

Accuracy: Indications of correct identifications are described as a proportion of all instances

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100 \quad (1)$$

Precision: Specificity reveals the percentage of datasets that a model claims to be important and was in fact meaningful. In precise, this indicates that classifiers only give pertinent examples, and it is written as:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \times 100 \quad (2)$$

Recall: In a dataset, recall indicates that all instances are relevant. The classification model identifies the relevant instances in the recall and is expressed as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (3)$$

Specificity: The ratio between true negatives and actual negatives ($\text{FP} + \text{TN}$) is expressed as:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (4)$$

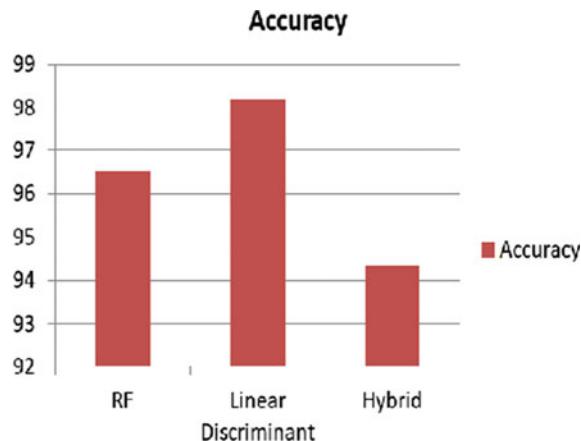
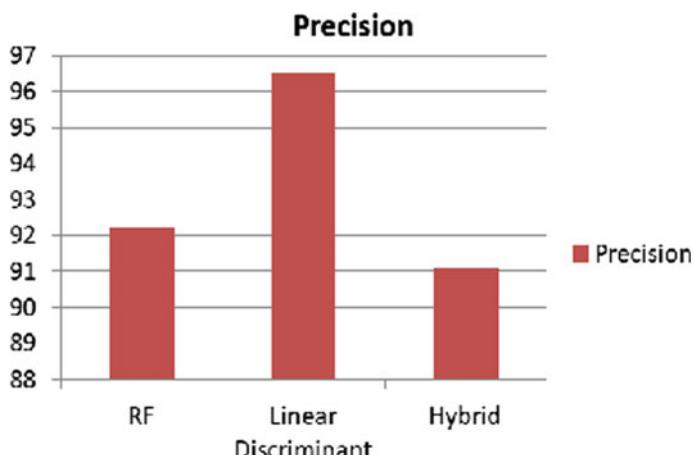
Table 1 provides the performance measure analysis of the presented model for lung cancer survivability analysis and prediction using ML approach.

According to Table 1, linear discriminant analysis has good precision, recall, accuracy, and specificity when applied for lung cancer survivability analysis and prediction (Figs. 2 and 3).

The accompanying graph illustrates that linear discriminant analysis has superior accuracy and precision in this comparison (Figs. 4 and 5).

Table 1 Performance measure analysis

Performance metrics	RF (random forest)	Linear discriminant	Hybrid ANConvNet
Accuracy (%)	96.52	98.20	94.34
Precision (%)	92.25	96.53	91.10
Recall (%)	94.11	97.48	92.33
Specificity (%)	93.74	98.85	92.54

**Fig. 2** Accuracy rating**Fig. 3** Precision comparison

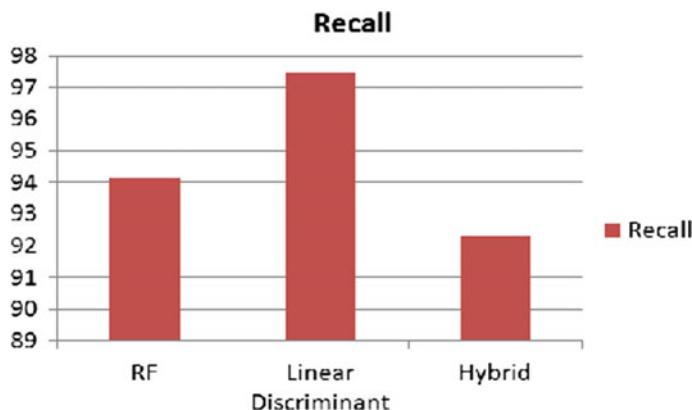
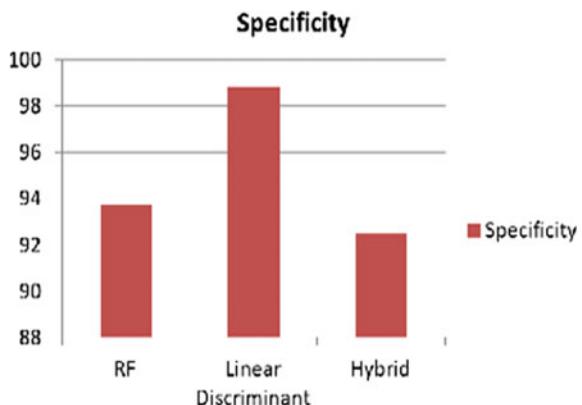


Fig. 4 Recall performance

Fig. 5 Specificity comparison of hybrid ANConvNet with RF, linear discriminant



In this comparison, the above graph shows that linear discriminant analysis has higher recall and specificity.

5 Conclusion

In the past, a physician would need to do a number of tests to determine whether one patient had lung disease or not. However, this was a lengthy procedure. A patient may occasionally be required to undergo pointless examinations or further tests in order to diagnose cancer. There must be a testing process that alerts the patient and the doctor to the possibility of lung cancer in order to reduce duration and pointless examinations. These days, algorithms are crucial for the classification and prediction of medical data. An experiment is being conducted to find the model

that provides the most accurate results from the lung cancer survivor dataset. The investigation's goal is to use a variety of neural network-based methods to identify early level lung cancer in a person. Numerous strategies for predicting and classifying were proposed as the life expectancy of lung cancers has increased recently, but the efficiency they offered was not adequate. For this comparison study, techniques like RF, hybrid, and LD analysis were applied. Statistical comparisons are made between algorithms' prediction abilities. For each classification on the lung sample, various findings are shown in the effectiveness chart. Accordingly, the various techniques determined metrics for various factors like efficiency, expertise, retrieval, and rigor. In this experiment, the linear discriminant analysis performed better than the hybrid and RF. There is a greater scope of the dataset parameter improvisation and discriminant analysis through support vector machines.

References

1. Zahid U, Ashraf I, Khan MA et al (2022) BrainNet: optimal deep learning feature fusion for brain tumor classification. *Comput Intell Neurosci* 2022:1–13
2. Saroja P, Udayaraju P, Sureesha B (2019) A survey on large scale bio-medical data implementation methods. *Int J Pharm Res* 1(11):649–656
3. Udayaraju P, Bharat Siva Varma P, Jeevana Sujitha M (2018) A survey of methods for genome functional analysis in comparative genomics. *Int J Eng Technol (UAE)* 7(12):681–688
4. Lee J, Kim H, Kim H, Lee S, Park E (2019) Development of a machine learning model to predict malignant recurrence in breast cancer patients. *BMC Cancer* 19(1):1010
5. Gupta R, Chaturvedi S, Singh V (2020) Machine learning-based prediction model for malignant recurrence in breast cancer patients. *Sci Rep* 10(1):15758
6. Wang Y, Xu J, Wang Y, Zhang Y, Li X (2021) Prediction of malignant recurrence in breast cancer patients using machine learning. *Sci Rep* 11(1):2149
7. Ryu SM, Lee S-H, Kim E-S, Eoh W (2019) Predicting survival of patients with spinal ependymoma using machine learning algorithms with the SEER database. Citation: *World Neurosurg* (2019)
8. Yutong Xie, et al (2018) Knowledge-based collaborative deep learning for benign malignant lung nodule classification on chest CT, IEEE
9. Anifah L, Haryanto, Harimurti R, et al (2017) Cancer lung detection on CT Scan image using ANN backpropagation based gray level co occurrence matrix feature. 978-1-5386-3172-0/17/. IEEE
10. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics: GLOBOCAN evaluates the incidence and death rates for 36 cancers in 185 countries worldwide. CA: Cancer J Clin 68(6):394–424. <https://doi.org/10.3322/caac.21492>
11. Lynch CM, Abdollahi B, Fuqua JD et al (2017) Prediction of lung cancer patient survival via supervised machine learning classification techniques. *Int J Med Infom* 108:1–8
12. Karhan Z, Tunç T (2016) Lung cancer detection and classification with classification algorithms. *IOSR J Comput Eng (IOSR-JCE)* 18(6):71–77. e-ISSN: 2278-0661, p-ISSN: 22788727, Ver. III (Nov-Dec)
13. Ranjan Baitharu T, Kumar Pani S (2015) Comparative study of data mining classification techniques using lung cancer data. *Int J Comput Trends Technol (IJCTT)* 22(2), April.
14. Vinitha D, Gupta D, Khare S (2015) Exploration of machine learning techniques for cardiovascular disease. *Appl Med Inf* 36: 23–32
15. Kaur S (2015) Comparative study review on lung cancer detection using neural network and clustering algorithm. *Int J Adv Res Electron Commun Eng (IJAECCE)* 4(2), February

16. Krishnaiah V, Narsimha G, Subhash Chandra N (2013) Diagnosis of lung cancer prediction system using data mining classification techniques. (IJCSIT) Int J Comput Sci Inf Technol 4(1)
17. Cruz JA, Wishart SD, Applications of machine learning in cancer prediction and prognosis, PMID: 19458758
18. Sujitha R, Seenivasagam V, Classification of lung cancer level with machine learning over big data healthcare framework, <https://doi.org/10.1007/s12652-020-02071-2>
19. Senthil S, Ayshwarya B, Lung cancer prediction using feed forward back propagation neural networks with optimal features. Int J Appl Eng Res 13(1):318–325
20. Nimala S, Rani RU, Rao PS (2023) High blood pressure classification using meta-heuristic based data-centric hybrid machine learning model. Commun Comput Inf Sci, 169–188. 1798 CCIS

Application of Analytical Network Processing (ANP) Method in Ranking Cybersecurity Metrics



Seema Gupta Bhol, Jnyana Ranjan Mohanty, and Prasant Kumar Patnaik

Abstract Since the advent of the Internet and digital technology, every organization uses digital tools to conduct its daily business. Finding new threats and vulnerabilities and calculating their influence on an organization's operations are the key goals of cybersecurity. To have strong cybersecurity in place, it is important to have some kind of mechanism to measure it. Metrics are used to measure something that has no direct method/instrument to measure. Cybersecurity metrics refer to the quantitative and qualitative measurements used to assess the effectiveness, efficiency, and overall health of an organization's cybersecurity efforts. By leveraging multi-criteria decision-making (MCDM) techniques in cybersecurity, organizations can improve their decision-making processes, optimize resource allocation, and enhance their overall cybersecurity posture to better defend against evolving cyber threats. Therefore, the purpose of this study is to model cybersecurity metrics evaluation by developing a decision network using analytical network process (ANP). The identification of various criteria is very important. Security metrics are organized into five primary classes divided down into 15 sub-classes, and these 15 sub-classes are further subdivided into 29 sub-classes.

Keywords Cybersecurity metrics · Multi-criteria decision-making (MCDM) · Analytic network process (ANP)

1 Introduction

Cybersecurity is of paramount importance in today's digital world due to the increasing reliance on technology and the Internet. It encompasses a range of practices, technologies, and processes designed to protect computer systems, networks, and data from unauthorized access, cyberattacks, and damage. A metric is a quantifiable measure used to assess, evaluate, or track a particular aspect of a system, process,

S. G. Bhol (✉) · J. R. Mohanty · P. K. Patnaik
KIIT University, Bhubaneshwar, India
e-mail: seemaguptabhol@gmail.com

or performance. Metrics provide objective data and help in understanding the effectiveness, efficiency, and overall performance of various activities [1]. Metrics are commonly used in different fields, including business, engineering, technology, and science, to monitor progress, identify areas for improvement, and make informed decisions. Cybersecurity metrics are essential for assessing the effectiveness of an organization's security posture, identifying potential weaknesses, and making data-driven decisions to improve security [2]. These metrics help measure various aspects of cybersecurity performance. However, there is no standard taxonomy for cybersecurity metrics. The present study is based on the taxonomy proposed by Bhol et al. [3]. Multi-criteria decision-making (MCDM) is a decision-making approach that involves evaluating and selecting the best option from a set of alternatives based on multiple criteria or objectives. In the context of cybersecurity, MCDM plays a crucial role in helping organizations make informed decisions to enhance their security posture and effectively allocate resources. ANP stands for analytic network process, and it is a decision-making methodology developed by Saaty [4]. It is an extension of the analytic hierarchy process (AHP) and is used to address more complex decision-making problems that involve interdependencies and interactions between criteria, sub-criteria, and alternatives [5].

2 Literature Review

The integrated approach was applied in making judgments relating to cybersecurity to determine a measure of effectiveness for risk assessment [6]. An information security risk-control assessment model was proposed in the study conducted by Yu-Ping et al. [7]. Xuan-ru suggested the cognitive framework of cyberspace operation support degree to the joint combat system and applied the ANP method for support degree evaluation [8]. To improve security management efficiency, an ANP-based evaluation index system was proposed [9]. Torbacki proposed the integration of DEMATEL and ANP (DANP) and PROMETHEE II methods for cybersecurity assessment in the manufacturing industry [10]. Bhol et al. evaluated cybersecurity metrics using a hybrid approach of AHP and Electre [11].

Cyber defense is very important for a nation's national security concern. The level of cyber power of a country was evaluated by a model based on ANP [12]. ANP-based approach for combined risk assessment for safety and security was proposed, which allows consideration of dependence and conflict among attributes [13]. A susceptibility evaluation is made using the ANP method to deal with incidents, concerned with denial of service attacks on micro grids in power sector [14]. Xiahou proposed resilience theory to the security concerns related to data center physical infrastructures. The ANP method was adopted to set up the evaluation model [15]. Liu proposed a model with ANP to prioritize assets according to value, which might assist the system administrator in deciding whether to increase security [16].

The Internet of Things is another area, where the ANP method has seen lot of applications. For smart cities, billions of devices are connected to IoT, inviting security and confidentiality threats. Huang et al. advocated the application of the ANP method in the analysis of cybersecurity risks for smart cities [17]. The study by Kim–Dong examines the IoT security requirement model, with the help of the ANP method [18]. Abbas suggested monitoring and control of IoT devices through a scalable model for better connectivity and coverage. The model was evaluated using the ANP method [19].

In a study by Hinduja et al., evaluation of the security features of IoT-based equipment was suggested based on ANP and Gray Relational Analysis (GRA) method [20]. Fuzzy ANP was also applied for the best selection of the attributes for Internet of Health devices to boost industry productivity [21]. To attain a high level of sustainable security, Alfakeeh proposed a fuzzy ANP that was used to calculate the influence of the overall sustainable security of health information software systems and their attributes [22]. Yazdani in their study on risk management for cloud computing environment analyzed the collected data with the help of the fuzzy ANP method [23]. The criteria must be prioritized, and the security operations chosen must be flexible in nature.

3 Methodology

The methodology implemented for the evaluation of cybersecurity metrics is discussed here. Initially, pertinent published works to date were reviewed and found that MCDM methods are extensively applied in decision-making situations, involving multiple criteria. The next stage after a literature review is the selection of criteria and alternatives. Previous work by the authors [3] forms the basis for criteria selection. This section is consisting of two subsections. First, cybersecurity metrics taxonomy is explained in brief. Next, the application of the ANP method for the evaluation of cybersecurity metrics is discussed in detail.

3.1 Cybersecurity Metrics

Cybersecurity metrics refer to specific measures and indicators used to evaluate and quantify various aspects of an organization's cybersecurity performance, effectiveness, and risk levels [24]. These metrics provide valuable data-driven insights into the state of an organization's security posture and its ability to defend against cyber threats [25]. Bhol et al. [3] suggested a classification in which cybersecurity was categorized into various classes and sub-classes, as shown in Fig. 1.

Vulnerabilities are flaws or weaknesses in a system or software that can potentially be exploited by hackers. Cybersecurity defense mechanisms are strategies, technologies, and practices implemented to protect computer systems, networks, and

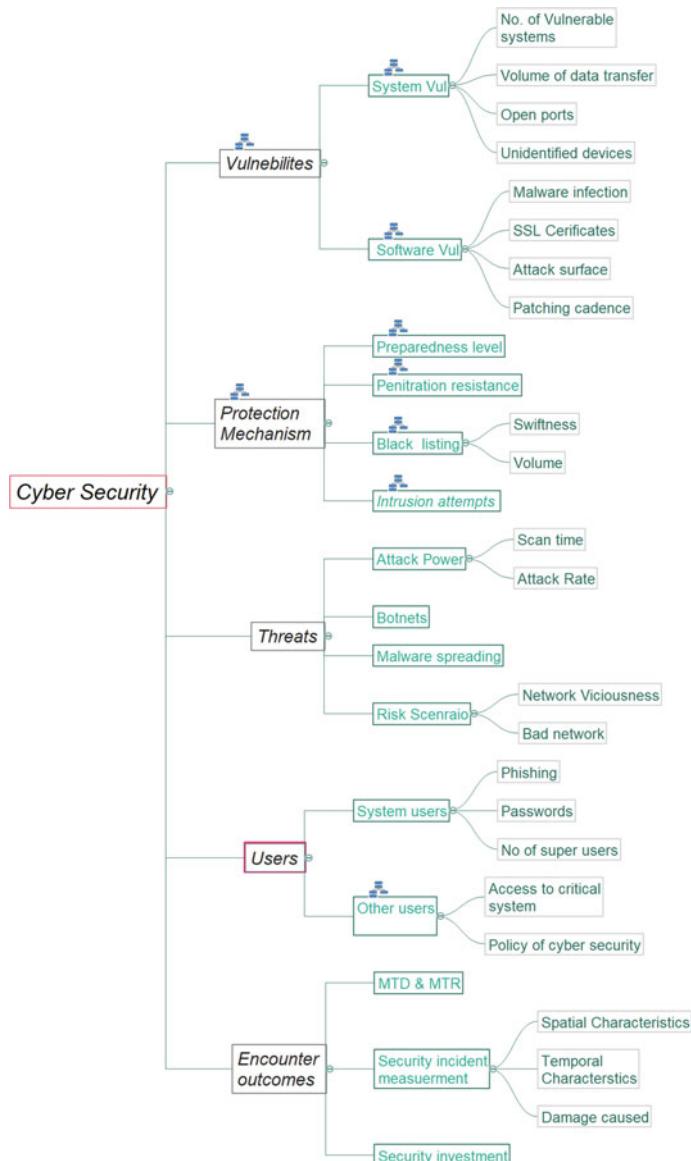


Fig. 1 Classification of criteria and sub-criteria, as proposed by Bhol et al. [3]

data from cyber threats and attacks. Cybersecurity risks refer to potential threats and vulnerabilities that can compromise the confidentiality, integrity, and availability of digital information and systems. The awareness level of employees, users, and

other stakeholders is crucial to cybersecurity. Employees or individuals with authorized access to systems may misuse their privileges to steal data, disrupt operations, or cause harm. Incidents between cyberattacks and defense are defined here as encounters.

3.2 Cybersecurity Metrics Evaluation Using ANP

The ANP method is widely used in various fields to address complex decision-making problems that involve interrelated and interdependent criteria. By considering the interactions between elements, ANP provides a more robust and comprehensive approach to decision-making compared with traditional methods that do not account for these interdependencies [26]. For the evaluation of cybersecurity metrics, 29 sub-criteria are identified under five main criteria clusters. All of the criteria and sub-criteria are given a code letter and these codes are used in the supermatrix, as given in Table 1.

For the evaluation of cybersecurity metrics, three alternatives are selected. Three companies Alpha Computers, Beta Technologies, and Delta Systems are three alternatives. Codes are assigned to alternatives as well, as given in Table 2.

Next, determination of the network control hierarchy was performed. Super Decision software which is designed to aid decision-making processes using various multi-criteria decision-making methods is utilized in this study [27, 28]. Each of the sub-nets contains a sub-criteria cluster and an alternatives cluster. A screenshot from the software Super Decisions (Fig. 2) shows the control structure.

The unweighted supermatrix was computed by making pairwise comparisons, using Saaty's basic scale. ANP technique allows an inconsistency ratio of no more than 10%. Different types of pairwise comparisons were conducted at the node level. Also, it was ensured that the inconsistency ratio is less than 10%. The judgments with a consistency ratio of more than 10% were rejected and fresh judgments were made.

An unweighted supermatrix is the resultant of all pairwise comparisons made, where all the computed priorities from pairwise comparisons are stored. Next, the weighted supermatrix was computed from the unweighted supermatrix. To create a weighted supermatrix, each element in a component of the unweighted supermatrix is multiplied by the weight of the corresponding cluster. Instead of nodes, cluster-wise comparisons are made in this step. All criteria cluster pairs are compared with respect to the goal. Now, criteria priorities are calculated and a consistency check was applied. The normalized priority of criteria clusters with respect to goal is given in Table 3.

Further, each of the criteria clusters and alternatives clusters are compared with respect to each cluster. Cluster priorities with respect to alternatives cluster are given in Table 4. The consistency index is 0.07678 which is less than 10%, hence judgments are accepted.

Table 1 Code assignment to criteria and sub-criteria

Clusters	Nodes	Indicators	Code
Vulnerabilities (B)	System vulnerabilities (B1)	Number of vulnerable systems	B11
		Volume of data transfer	B12
		Open ports	B13
		Unidentified devices	B14
	Software vulnerabilities (B2)	Malware infections	B21
		SSL certificates	B22
		Attack surface	B23
		Patching cadence	B24
Protection mechanism (C)	Preparedness level (C1)	Preparedness level	C1
	Penetration resistance (C2)	Penetration resistance	C2
	Black listing (C3)	Swiftness	C31
		Volume	C32
Users (D)	Intrusion attempts (D1)	Intrusion attempts	D1
	System users (D2)	Phishing	D21
		Password	D22
		No. of super users	D23
	Other users (D3)	Access to critical system	D31
		Policy of cybersecurity	D32
Threats (E)	Attack power (E1)	Scan time	E11
		Attack rate	E12
	Botnets (E2)	Botnets	E2
	Malware spreading (E3)	Malware spreading	E3
	Risk scenario (E4)	Network viciousness	E41
		Bad networks	E42
Encounter outcomes (F)	MTD & MTR (F1)	MTD & MTR	F1
	Security incidents (F2)	Spatial characteristics	F21
		Temporal characteristics	F22
		Damage caused	F23
	Security investments (F3)	Security investments	F3

Table 2 Codes assignment to alternatives

Alternatives	Code
Alpha computers	A1
Beta technologies	A2
Delta systems	A3

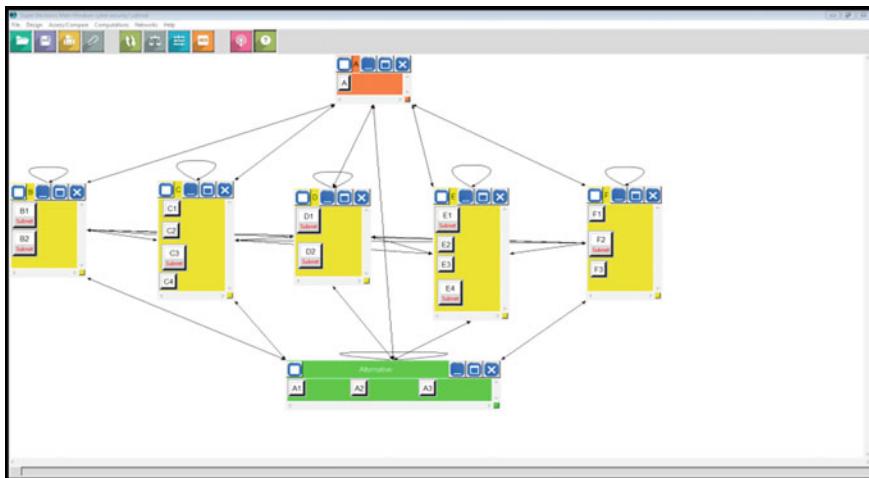


Fig. 2 Network control hierarchy

Table 3 Criteria cluster priorities with respect to goal

With respect to goal	Normalized priority
Vulnerability	0.49
Protection mechanism	0.24
Users	0.14
Risks	0.09
Encounter outcomes	0.04
Consistency ratio = 0.01697	

Table 4 Cluster priorities with respect to alternatives cluster

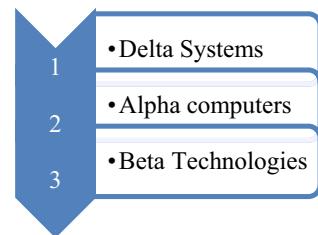
With respect to alternatives	Normalized priority
Vulnerability	0.47
Protection mechanism	0.10
Users	0.23
Risks	0.08
Encounter outcomes	0.12
Consistency ratio = 0.07678	

Next, the limiting supermatrix will be computed using a weighted supermatrix. After computations, overall synthesized priorities for the alternatives are obtained, as given in Table 5.

Table 5 Final priorities of alternatives

Alternatives	Priority
Alpha computers	0.35
Beta technologies	0.27
Delta systems	0.38

Fig. 3 Final ranking of alternatives using ANP



4 Results and Analysis

In the present paper cybersecurity metrics were evaluated using the ANP method. Analytical network model was built with seven clusters: goal cluster, five criteria clusters, and an alternatives cluster. Different types of pairwise comparisons were made between different criteria and clusters. There are five main criteria divided into 15 sub-criteria and these 15 sub-criteria are further classified into 29 sub-criteria (Fig. 3). Three companies are selected Alpha Computers, Beta Technologies, and Delta Systems.

The developed network was evaluated with the help of Super Decisions software. After synthesizing the whole model, alternative's priorities were computed (Table 3). Figure 3 shows the final ranking of these companies. Alpha computers got the highest priority value, i.e., 0.46 followed by Beta Technologies and last is Delta Systems with a 0.25 preference value.

5 Conclusions and Future Research

In the past few years, the demand for Internet connectivity is amplified due to the widespread usage of computers, smart phones, and other electronic devices. However, on the other hand, attacks against software vulnerabilities and data breaches have also grown significantly in size. Cybersecurity metrics should be aligned with the organization's risk management strategy and specific security goals. Decision-making is a complex process, especially in the cybersecurity area, where accuracy is vital due to its impact on vital assets of any business or organization. Employing

multi-criteria decision-making (MCDM) tools like ANP ensures a logical and transparent approach. The proposed study elaborates on the formulation and evaluation of methods for the evaluation of metrics of cybersecurity.

ANP approach is found to be user-friendly that possesses universal applicability, effectively addressing various multiple criteria ranking and selection problems, including evaluation of metrics of cybersecurity. Also, the concept of the consistency index checks if the decision maker is rational in its approach or not. Considering potential bias in human judgments, different criteria preferences may influence the final outcomes. Also, some criteria with qualitative or unknown structures may be challenging to accurately measure. A solution proposed for future research involves using fuzzy numbers to achieve an evaluation matrix, increasing result accuracy.

In conclusion, ANP is proposed to evaluate the metrics of cybersecurity. The intermediate results of the computation process are also presented in the proposed paper. This study focuses on prioritization and selection of various factors for cybersecurity metrics evaluation to build robust secure systems.

References

1. Bendovschi A (2015) Cyber attacks—trends, patterns and security countermeasures. *Procedia Econ Finance* 28:24–31
2. Muiyuro A (2018) Cybersecurity metrics, supporting accurate and timely decision-making. In: *Cybersecurity metrics & dashboards*, pp 1–25
3. Bhol SG, Mohanty JR, Pattnaik PK (2023) Taxonomy of cyber security metrics to measure strength of cyber security. *Mater Today Proc* 80:2274–2279
4. Saaty TL (2005) Theory and applications of the analytic network process: decision making with benefits, opportunities, costs and risks. RWS Publications
5. Saaty TL (2008) The analytic hierarchy and analytic network measurement processes: applications to decisions under risk. *Eur J Pure Appl Math* 1(1):122–196
6. Wilamowski GC, Jason RD, Steven MF (2017) Using analytical hierarchy and analytical network processes to create cyber security metrics. *Defense AR J* 24(2)
7. Yu-Ping Y, Shieh H, Tzeng G (2013) A VIKOR technique based on DEMATEL and ANP for information security risk control assessment. *Inf Sci* 232:482–500
8. Chen X, Shen J, Xing J, Dai Y (2019) Support degree evaluation of cyberspace operation to joint combat system. In: *Chinese Control and Decision Conference (CCDC) 2019*. Nanchang, China, pp 1770–1774
9. Lin S, Ye Y, Han Y, Zhu Y, Li Q, Wu J (2019) Application of analytic network process in power grid development-diagnosis management. In: Abawajy J, Xu Z, Atiquzzaman M, Zhang X (eds) *2021 international conference on applications and techniques in cyber intelligence. ATCI 2021. Advances in intelligent systems and computing*, vol 1398. Springer, Cham
10. Torbacki W (2021) A hybrid MCDM model combining DANP and PROMETHEE II methods for the assessment of cybersecurity in industry 4.0. *Sustainability* 13(16):8833
11. Bhol SG, Mohanty JR, Pattnaik PK (2020) Cyber security metrics evaluation using multi-criteria decision-making approach. In: *Smart intelligent computing and applications: proceedings of the third international conference on smart computing and informatics*, vol 2. Springer Singapore
12. Vuuren V, Jansen J, Leenen L (2018) A model for measuring perceived cyberpower. In: *ICCWS 2018 13th international conference on cyber warfare and security*, p 320. Academic Conferences and Publishing Limited

13. Verma S, Gruber T, Schmittner C, Puschner P (2019) Combined approach for safety and security. In: Romanovsky A, Troubitsyna E, Gashi I, Schoitsch E, Bitsch F (eds) Computer safety, reliability, and security. SAFECOMP 2019. Lecture Notes in Computer Science, vol 11699. Springer, Cham
14. Wang B, Qiuye S, Wang R, Dong C (2022) Vulnerability analysis of secondary control system when microgrid suffering from sequential denial-of-service attacks. *IET Energy Syst Integr* 4(2):192–205
15. Xiahou X, Jialong C, Zhao B, Zixuan Y (2022) Research on safety resilience evaluation model of data center physical infrastructure: an ANP-based approach. *Buildings* 12(11):1911
16. Liu Y, Dejun M (2022) An information asset priority evaluation method with analytic network process. *Inform Serv Use* 1–7
17. Huang C, Shah N (2021) Analyzing and evaluating smart cities for IoT based on use cases using the analytic network process. *Mob Inform Syst* 1–13
18. Dong K, Hee JYC, Jongin L, Bong GL (2016) Developing IoT security requirements for service providers. International Information Institute (Tokyo), vol 19, no 2
19. Abbas AW (2021) Cyber secured framework for control and monitoring of IoT devices in smart logistics. PhD diss., University of Engineering & Technology Peshawar, Pakistan
20. Hinduja A, Pandey M (2020) An ANP-GRA-based evaluation model for security features of IoT systems. In: Choudhury S, Mishra R, Mishra R, Kumar A (eds) Intelligent communication, control and devices. Advances in intelligent systems and computing, vol 989. Springer, Singapore
21. Lin W (2021) Determining the degree of characteristics for internet of healthcare devices using fuzzy ANP. *Scientific Programming*, pp 1–11
22. Alfakeeh AS, Abdulmohsen A, Fawaz JA, Yoosef BA (2022) Sustainable-security assessment through a multi perspective benchmarking framework. *CMC-Comput Mater Continua* 71(3):6011–6037
23. Yazdani AA, Abbas K, Ozgur T, Yazwand P (2023) Evaluation of cloud computing risks using an integrated fuzzy-ANP and FMEA approaches. *Int J Appl Decis Sci* 16(2):131–164
24. Swanson M, Bartol N, Sabato J, Hash J, Graffo L (2003) Security metrics guide for information technology systems. NIST Special Publication vol 800, no 55
25. Black PE, Scarfone K, Souppaya M (2008) Cyber security metrics and measures. In: Handbook of science and technology for homeland security. Wiley, vol 5, 1–10
26. Saaty TL, Cillo B (2008) A dictionary of complex decision using the analytic network process. *Encyclcon* 2(2)
27. Adams B (2011) SuperDecisions limit matrix calculations. Decision Lens Inc.
28. SuperDecisions. Software for Decision-Making. <http://www.superdecisions.com>. Last accessed 2 May 2023

Advanced Real-Time Monitoring System for Marine Net Pens: Integrating Sensors, GPRS, GPS, and IoT with Embedded Systems



Sayantan Panda, R. Narayananamoorthi, and Samiappan Dhanalakshmi

Abstract Aquaculture is rapidly growing to meet the global demand for fish consumption, but it faces several challenges that adversely affect marine environments. To address these issues, economical remote-surveillance system has been developed for marine net pens, utilizing GSM, GPS modules with IoT and embedded systems technology. The goal is to create an actual time, accurate remote-surveillance system that can replace the need for frequent manual inspections by aquaculture farm workers. To overcome this, the project aims to remotely investigate aquaculture parameters such as seawater temperature, humidity, turbidity, chlorophyll, under water pressure, dissolved oxygen (DO), and pH every day, without physical visits to the pens. Ultimately, the system's objective is to enable remote control and observation of sustaining and maritime operations. The key components of the monitoring system include the fish farming enclosure monitoring station and the remote-surveillance centers. This monitoring station is equipped with intelligent water sensors to detect seawater temperature, turbidity, DO, and pH in real time. This data is collected by a remote data-collection terminal and transferred by GPRS network for online processing. The remote-surveillance centers serve as centralized servers and connect to the monitoring stations through the Internet. The seawater data collected by the particular monitoring stations is transmitted to these centers via GPRS and GSM wireless networks with the specific locations by the GPS technology, facilitating continuous monitoring and reporting. One significant challenge faced by Open-Pen Sea Cage Aquaculture is the use of fishmeal sourced from wild fisheries in the feed given to farmed fish. While alternative proteins from land-based

S. Panda · S. Dhanalakshmi (✉)

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, India

e-mail: dhanalas@srmist.edu.in

S. Panda

e-mail: sp0508@srmist.edu.in

R. Narayananamoorthi

Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, India

e-mail: narayanr@srmist.edu.in

crops are included in the feed pellets, carnivorous or omnivorous species still require some fishmeal to obtain essential amino acids. Small, oily fish such as anchovies, sardines, and pilchards, known as “Industrial Fish,” are commonly used for fishmeal. Roughly, 20 million tons of these fish are caught yearly for aquaculture, with much of Australia’s sardine catch, unfit for human consumption, feeding Southern Bluefin Tuna in Port Lincoln. Relying on wild-caught fish for fishmeal raises sustainability and ecosystem impact concerns.

Keywords Marine net pens · GPS · GPRS · GSM · Remote surveillance · IoT · Embedded system

1 Introduction

Mari culture has emerged as quickly growing culinary sector globally, since 1970; the rate of mean yearly growth has been an impressive 8.9% [1]. However, maintaining optimal water quality, particularly with regard to temperature, remains a critical factor influencing fish growth and mitigating the risk of large-scale fish diseases [1]. To address these challenges and foster sustainable aquaculture practices, the future of marine aquaculture is anticipated to incorporate automatic distant monitoring and computerized intensive cultivation as the leading approach. The administration of aquaculture places significant emphasis on monitoring water quality to ensure ideal growth conditions. Nevertheless, the vast potential of aquaculture in resource-rich marine environments has been relatively underexplored in terms of observing and tracking progress in real time, collecting relevant data, storing it securely, and conducting comprehensive analysis on diverse areas of growth parameters. To bridge this gap, our implementation involves deploying an IoT and embedded systems-based real-time controlling system specifically designed for aquaculture marine net pens. This novel system is a comprehensive integration of GPRS communication, embedded systems, and aquaculture cage remote monitoring, effectively supplanting the requirement for conventional manual monitoring.

The aquaculture cage monitoring station’s interaction is seamlessly facilitated by the central server, enabling wireless remote monitoring. Vital seawater parameters such as temperature, humidity, dissolved oxygen, and pH are continuously captured in real time, enhancing the ability to oversee and manage the aquaculture environment promptly.

As part of our research, fish reproduction is methodically managed through a computer system, incorporating insights from relevant studies in the field. This sophisticated system endeavors to optimize aquaculture practices while minimizing any potential environmental damage. As a crucial component of environmental evaluation, many countries have implemented water quality monitoring plans to assess pollution effects in aquatic systems. To address the quantity of analysis required and associated costs, regulatory agencies have devised a general index known as the Water Quality Index (WQI) [2]. Given the importance of sustainable fish farming

and its significant impact on the environment, various studies have focused on understanding and mitigating the negative effects of aquaculture. Improved farming practices, location strategies, and standardized monitoring programs have contributed to reducing adverse environmental impacts [3]. Additionally, simulation models have been devised to estimate water renewal rates, dispersion, and deposition of organic particles from fish farms, as well as the effects of organic material on sediment and benthic infauna in aquatic ecosystems. The integration of mathematical models with monitoring forms an effective management system, aiding in the regulation of the environmental impact of fish farming.

2 Similar Works

Remote-surveillance systems have garnered considerable focus in marine aquaculture, primarily driven by the need to rear migratory fish species like salmon and eel in open-sea aquaculture setups, surpassing the efficient limitations of in-shore fish farming. Notably, various researchers have diligently worked on the development of remote monitoring systems in this domain.

Some of these researchers leveraged sensor network technologies as the foundational framework for Internet of Things (IoT) [4, 5]. Typically, each ground node employs radio frequency communication [6], while acoustic communication is utilized [7, 8]. In a noteworthy research endeavor, a specialized monitoring system was introduced, specifically designed for monitoring sea cages. This suggested system boasts a comprehensive configuration, encompassing a primary server, a remote-surveillance center, a seawater controlling station. Also includes a GSM/GPRS data sender and receiver and sensors. Importantly, our investigation has not revealed any existing system analogous to the uniqueness of our proposal. Given the constraints imposed by limited battery power in ocean sensor networks, we have meticulously optimized the marine water surveillance station by integrating cutting-edge, an inexpensive and energy-efficient wireless data communication technology with low power consumption.

3 Proposed System's Key Components and Architecture

The system focuses on efficient monitoring of open-sea aquaculture cages, addressing experts' needs for remote assessment of critical parameters like temperature, pH, and more. It enables real-time monitoring, reducing physical visits and optimizing fish growth environments. Advanced sensors and data transmission empower informed decision-making. Remote control features revolutionize aquaculture management, enhancing efficiency and sustainability. This approach advances responsible fish farming practices. Monitoring temperature, DO, turbidity, and pH drives species

adaptation over time, influencing aquatic ecosystems. Anomalies prompt swift remedies, aiding optimal aquaculture management. Real-time location data identifies deviations and aids interventions. Notifications facilitate quick decisions.

An appropriate guide is provided according to the situation to assist the owner in making informed choices. The system empowers seamless field interventions by promoting efficient collaboration and coordination among field employees, ensuring effective issue resolution.

To achieve the mentioned objectives, a real-time monitoring system for seawater temperature, dissolved oxygen (DO), pH has been designed, comprising a marine water surveillance station, centralized server, and remote-surveillance center (Fig. 1). The marine water surveillance station gathers data on like humidity, temperature, and dissolved oxygen (DO), pH and transmits it to the central server through wirelessly. The remote monitoring center receives updated data from the central server via the Internet, offering functions for data presentation, storage, and download. Real-time latitude and longitude data of individual sea cages are used to, allowing for the identification of sea cages that may be experiencing deviations monitor and track their locations continuously in factors such as humidity, temperature, dissolved oxygen (DO), pH, and other relevant parameters.

This data helps in assessing the environmental conditions surrounding each cage and enables prompt intervention and management in case of any anomalies. The Water Quality Index (WQI) is a valuable tool used to summarize diverse water quality data into a single numerical value, representing the overall quality for specific uses. The index involves arithmetic weighting of normalized measurements, with different normalizations and weightings depending on the intended water usage. A Water Quality Index called WQI_{min} was calculated with only three parameters: dissolved oxygen (C_{DO}), turbidity (C_{Turb}), and total phosphorus (C_{TotP}). The equation

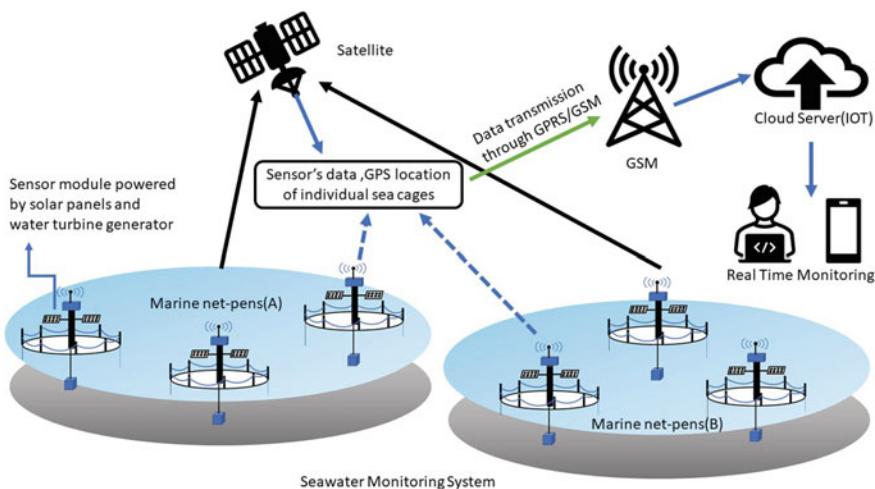


Fig. 1 Proposed sea-cages monitoring system—key components and architecture

for WQI_{min} is as follows (1):

$$\text{WQI}_{\min} = (C_{\text{DO}} + C_{\text{Turb}} + C_{\text{TotP}})/3. \quad (1)$$

Detailed explanations of each component of the remote monitoring system are provided in the subsequent subsections.

4 Hardware Setup and Operational Principles: Sea-Cages Monitoring Station

The marine water surveillance station is installed beneath the fish cages and connects with the on- land monitoring center through a GSM communication module, facilitating data transmission from the sensor module. The station consists of three essential modules: GSM communication, sensor, and GPS. The TTGO T-call module ensures data broadcasting between sensors, the monitoring center, and GPS module. The GPS module enables the identification of exact locations of sea-cages experiencing deviations in temperature [9], humidity, chlorophyll, pH, DO, etc. To describe our developed marine water surveillance station, refer to Fig. 2, which depicts the block diagram showcasing the incorporation of microcontroller, TTGO T-call, sensors, GPS, and power module.

The sensors and GPS module connect to the microcontroller, collecting and storing or transmitting data via the GSM module to the remote center. USART ports interface with the TTGO T-call module. The station is pivotal in managing data from interfaces and wirelessly transmitting it using the ESP32-SIM800L module.

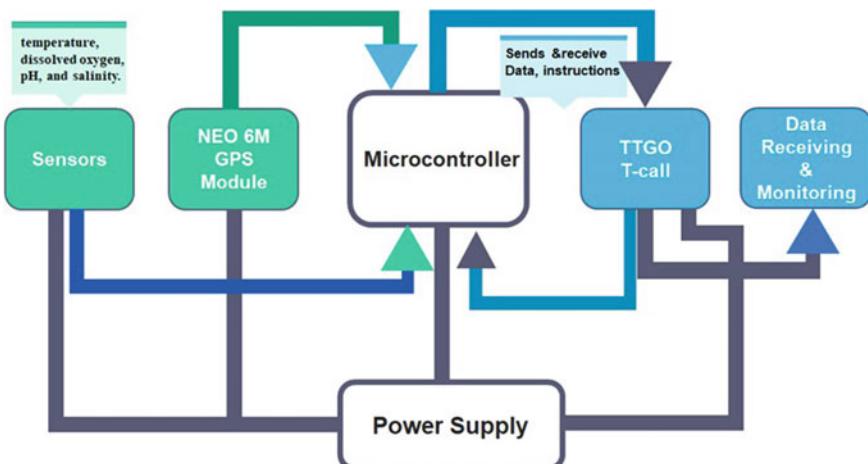


Fig. 2 Block diagram of the proposed marine water surveillance station

The system is powered by a dedicated module that generates and supplies electricity using solar panels and a water turbine generator. After meticulous consideration of project prerequisites encompassing stability, cost, durability, accuracy, and other critical factors, we have opted for the subsequent components to constitute the system.

1. **ATmega328P:** The ATmega328P is a powerful and energy-efficient 8-bit microcontroller with high-performance AVR® architecture [10]. It offers 131 efficient instructions, up to 16 MIPS throughput at 16 MHz, and features 32×8 general-purpose registers. With 32 K bytes of programmable flash memory, 1 K bytes of EEPROM, and 2 K bytes of SRAM, it provides ample storage. Essential peripherals like timers, PWM, ADC, USART, SPI, and I2C enhance its versatility. Six sleep modes and low power use conserve energy. It operates from -40 to $+125$ °C, with an input voltage of 2.7–5.5 V, makes it adaptable to different environments. Compact package options (TQFP, QFN/MLF) offer design flexibility.
2. **TTGO T-Call (ESP32 SIM800L) module:** The TTGO T-Call is an ESP32-based development board with an integrated SIM800L GSM module [11]. It's designed for IoT projects with both Wi-Fi and cellular connectivity. The ESP32 microcontroller provides dual-core processing, Wi-Fi, Bluetooth, and various interfaces for versatile IoT applications. With the SIM800L module, the T-Call connects to cellular networks, supporting SMS, voice calls, and GPRS Internet access (see Fig. 3). This is useful for remote communication and areas without Wi-Fi. The ESP32-SIM800L combo offers a compact platform for flexible IoT projects using both Wi-Fi and cellular connections.

Effective printed circuit board (PCB) layout is crucial for ensuring the optimal performance of electronic products. A well-designed layout plays a significant role in avoiding various issues, such as Time Division Distortion (TDD) noise and SIM card detection problems.

Following the principles mentioned earlier, the recommended layout, depicted in Fig. 3, is essential to achieve reliable and efficient operation.

3. **NEO-6 module:** The NEO-6 series boasts standalone GPS receivers powered by the U-blox 6 engine, ensuring accurate tracking. In a $16 \times 12.2 \times 2.4$ mm package [12], these modules offer connectivity and suit battery-operated devices. The 50-channel U-blox 6 engine achieves rapid Time-To-First-Fix (< 1 s), countering jamming and multipath effects for reliable positioning. Refer to Fig. 4 for the recommended PCB layout.
4. **Sensor module:** The monitoring system comprises a collection of sensors and devices designed to measure and monitor various environmental parameters for water quality assessment and control [13]. These sensors work together to gather critical data that provides valuable insights into the conditions of water bodies, aiding in the prediction of potential issues like algae blooms and nutrient loading. The data collected by these sensors will be transmitted through the ESP32 SIM800L module for remote monitoring and analysis.

Fig. 3 Recommended PCB layout of the TTGO T-Call

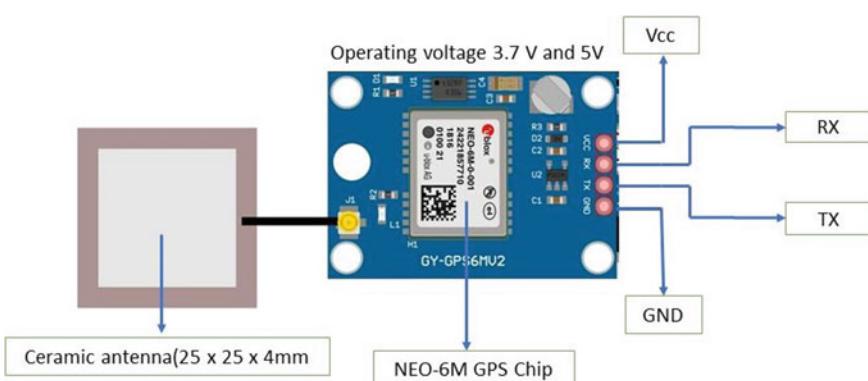
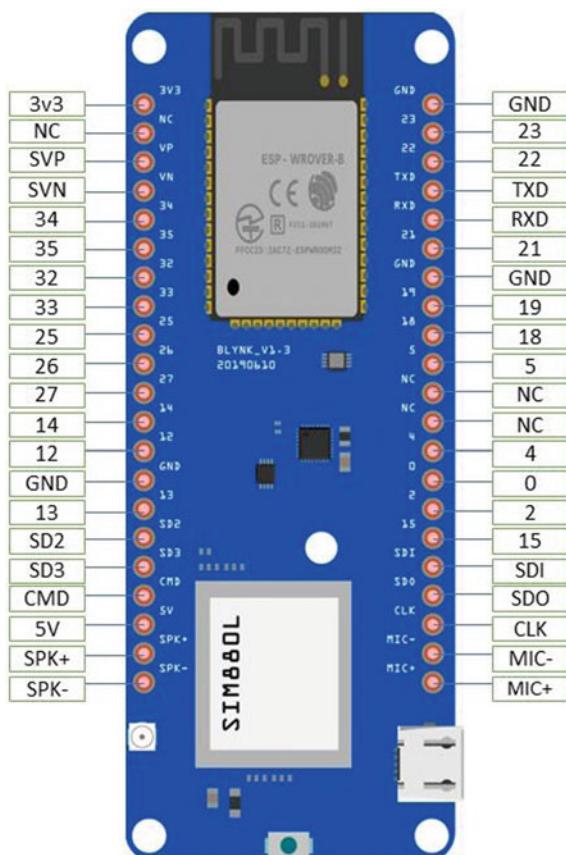


Fig. 4 Pin assignment of NEO-6M

- (i) **DS18B20 Digital Temperature Sensor:** The DS18B20 sensor [14] is used for highly accurate temperature measurements across a broad range, with a precision of typically ± 0.5 °C from – 10 to + 85 °C within a measurement range of – 55 to + 125 °C.
- (ii) **pH Sensor:** The pH sensor [15] is utilized to measure the acidity or alkalinity of a liquid solution, operating within a pH range of 0–14, a temperature range of 0–60 °C, and offering accuracy with zero point at 7 ± 0.5 pH and an alkaline error of 0.2 pH.
- (iii) **DHT11 Humidity and Temperature Sensor:** The DHT11 sensor [16] is employed for measuring temperature and relative humidity levels in the environment, with a temperature range of 0–50 °C (32–122 °F), a humidity range of 20–80% RH, and an accuracy of ± 2 °C for temperature and $\pm 5\%$ for relative humidity.
- (iv) **Turbidity Sensor:** The turbidity sensor [17] measures the concentration of total dissolved solids in the liquid, with a TDS range of 0–990 ppm (mg/L) and an accuracy of $\pm 2\%$.
- (v) **DO Sensor (Dissolved Oxygen Sensor):** The DO sensor [18] is utilized for measuring the concentration of dissolved oxygen in liquid, primarily water, with a measurement range of 0–20.0 ppm, a resolution of 0.1 ppm, and an accuracy ranging from ± 0.2 ppm to ± 1 count.
- (vi) **YSI 6025 Chlorophyll Sensor:** The YSI 6025 Chlorophyll Sensor [19] is designed to estimate phytoplankton concentrations in water bodies by detecting chlorophyll fluorescence in situ. It operates within a chlorophyll range of approximately 0–400 ug/L, an Relative Fluorescence Unit (RFU) range of 0–100 RFU, and provides a detection limit of around 0.1 ug/L. The sensor offers a resolution of 0.1 ug/L for chlorophyll and 0.1% for RFU measurements.

5. **Power Unit:** In Fig. 5, our innovative power module system is revealed, ingeniously crafted to provide uninterrupted power to the marine water surveillance station, utilizing tidal current energy harnessed from the sea stream and efficiently converting it into electrical energy. The system ensures reliable energy storage in a dedicated battery, making the station completely self-powered. Notably, the system's sustainability is further augmented by the incorporation of solar panels, enhancing power generation and ensuring a self-powered and eco-friendly operation, thereby eliminating the need for external power supply management.

The water turbine power generator comprises: The system comprises optimized three-bladed helices for current capture (Fig. 5), a 90° gearbox for power transmission, a waterproof low RPM generator, and energy storage via batteries. A charge controller ensures efficient battery management. This integrated design effectively harnesses sea currents for continuous, sustainable electricity generation.

The process unfolds as the sea current flow sets the helices into motion, their rotation seamlessly transmitted through the gearbox to the low RPM generator housed within the marine water surveillance station. Here, the generator effectively converts

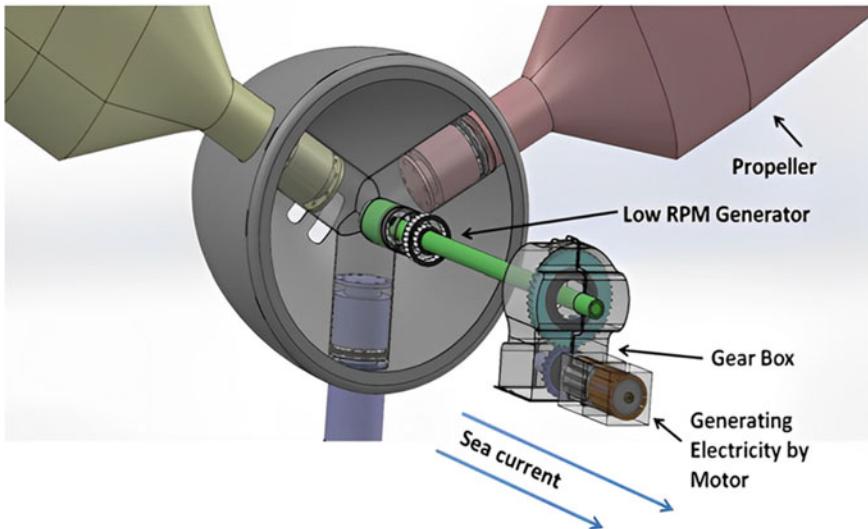


Fig.5 Water turbine power generator

the rotational motion into a substantial electrical power output. To optimize this energy generation, a bevel gear system (Fig. 6) is ingeniously employed, featuring a driver gear with 150 teeth and a driven gear with 15 teeth, creating a 1:10 gear ratio. As a result, the low RPM generator achieves the desired charging speed of 150 rpm when the blades rotate at a relatively slow pace of 15 revolutions per minute. This impressive setup ensures the generator produces an output of 12 and 2 A, thus effectively charging the battery.

This advanced power module caters to the energy requirements of the TTGO T-call module and sensors, handling a peak current of 50–100 mA during network use. It employs an ingenious design with a 12–3.7 V step-down converter (LM2596), effectively powering the TTGO T-call module, sensors, MCU, monitoring station, and the entire system. Additionally, a meticulously integrated TP4056 battery charging module manages voltages (4.12–4.2 V) for efficient battery charging while powering the MCU, TTGO T-call, and sensors. It is crucial to highlight that the same step-down converter and power charging module are ingeniously interconnected to the solar panels, creating a harmonious integration of solar energy [20]. This enables the system to optimally utilize the abundant solar resources, further enhancing the self-powered and eco-friendly attributes of the monitoring station (Fig. 7). In conclusion, this power module system represents a groundbreaking achievement in sustainable power generation for marine water surveillance stations. By synergistically harnessing tidal and solar energy, the system achieves unparalleled self-sufficiency, significantly reducing environmental impact while ensuring a seamless and reliable power supply for the monitoring station.

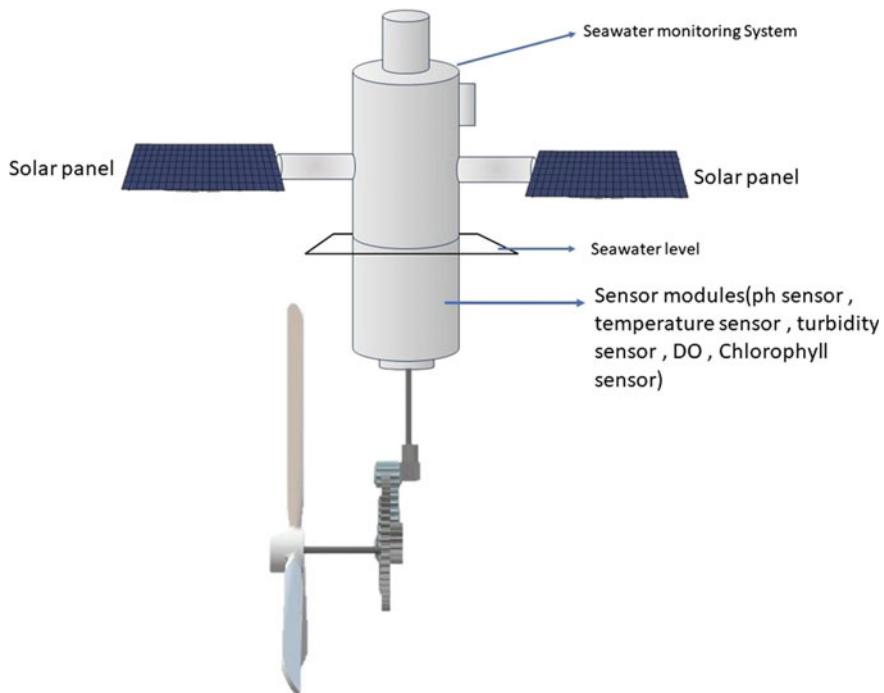


Fig. 6 Comprehensive diagram of the entire power module setup

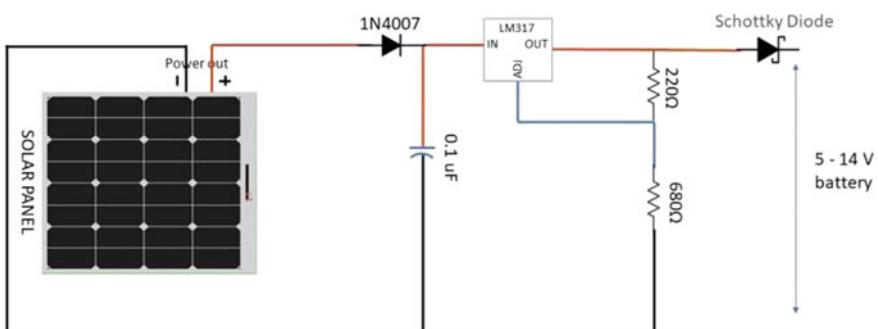


Fig. 7 Battery charging circuit by solar panels

5 Operating Mechanism of the Remote Monitoring Center

The remote monitoring system is designed to address the need for real-time monitoring of sea cages, optimizing fish farming practices while ensuring environmental sustainability. The system utilizes cellular and GPRS communication and MQTT machine to machine network protocols to enable seamless data transmission from

remote locations to the central monitoring center. The remote monitoring system is designed to address the need for real-time monitoring of sea cages, optimizing fish farming practices while ensuring environmental sustainability. The system utilizes cellular and GPRS communication protocols to enable seamless data transmission from remote locations to the central monitoring center.

5.1 System Architecture

By integrating GSM and GPRS modules, we establish reliable communication between sea cages and the central monitoring center. Sensor data, including water quality, temperature, and humidity, is published by monitoring devices. MQTT broker acts as an intermediary, forwarding data to IoT cloud subscribers via secure cellular networks. The center, equipped with IoT tech, serves as a data repository accessible via mobile app or laptop (see Fig. 8). Robust data protocols ensure security, encryption, and authentication. Predetermined data intervals optimize power and bandwidth. The system enables remote monitoring, control, real-time data access, notifications, and parameter adjustments via user-friendly app and laptop interface. Multi-factor authentication and access controls secure data and control features. This IoT platform ensures seamless connectivity and ease of use for users.

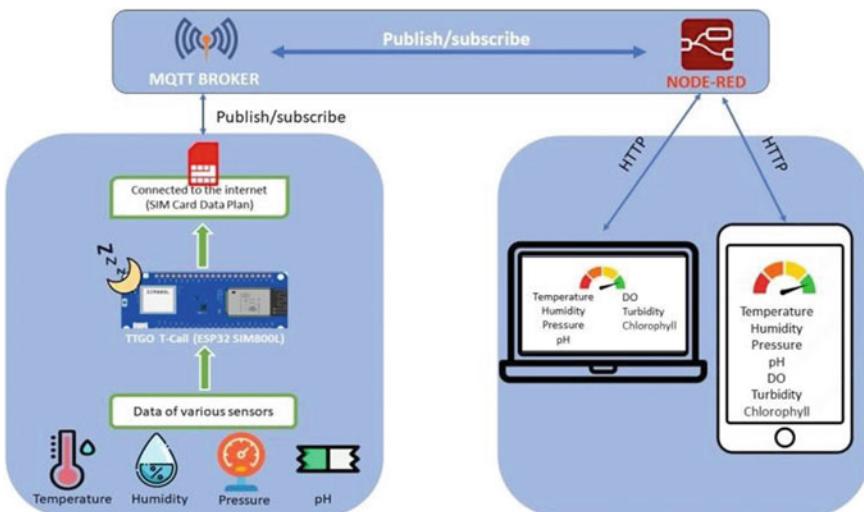


Fig. 8 System architecture of the proposed remote monitoring center

5.2 Digital Framework

We've developed a custom framework for data collection and transmission in our project to meet low power demands. This framework supports wake-up and deep-sleep modes, conserving energy. In TTGO T-Call, “esp_deep_sleep_start()” initiates deep sleep, while wake-up is triggered by timers, interrupts, GSM activity, or UART communication. Commands like “esp_sleep_enable_timer_wakeup()” set timer-based wake-up, “esp_sleep_enable_ext0_wakeup()” enables external interrupt wake-up, and “esp_sleep_enable_uart_wakeup()” allows UART wake-up. These manage power modes, achieving low consumption and timely wake-ups. Our advanced system uses ocean sensor networks for easy access to marine net-pen environmental data. This valuable information is then transmitted efficiently to the central monitoring center through thoughtfully selected data-transmission protocols.

The communication process unfolds as follows:

- (i) Sensor Data Acquisition: Specialized sensor modules in the sea cages continuously measure the environmental parameters. Each sensor generates analog or digital data representing the respective parameter's value.
- (ii) Data Aggregation: The TTGO T-Call gathers data from all the sensor modules, utilizing its GPIO pins or analog-to-digital converters (ADC) to read analog sensor values and acquire digital values from digital sensors.
- (iii) Data Interpretation: Upon receiving the data packets, the TTGO T-Call extracts the information and prepares it for further transmission to the remote monitoring center. For efficiency, it may apply data compression or encoding techniques, ensuring streamlined data transfer over the network.
- (iv) GPRS/GSM Transmission: Leveraging General Packet Radio Service (GPRS) or Global System for Mobile Communications (GSM) communication protocols, the TTGO T-Call transmits the aggregated and formatted data to the remote monitoring center. This enables data transmission over the cellular network to a secure cloud server or IoT platform.

5.3 Efficient Data Transmission to IoT Cloud Using MQTT Protocol

Message Queuing Telemetry Transport (MQTT) is an industry-standard messaging protocol facilitating efficient communication between IoT devices, like sensors, over limited-bandwidth networks. It enables seamless messaging between devices and cloud services, with advantages including efficiency, simplicity, and support for varying levels of Quality of Service (QoS) for reliable message delivery. QoS 0 guarantees message delivery only once, QoS 1 ensures at least once with retransmissions, while QoS 2 provides exact-once delivery via a four-step process, potentially

Table 1 Return codes in MQTT

Return code	Response
0	Connection established
1	Declined connection, invalid protocol version
2	Declined connection, invalid identifier
3	Declined connection, server not reachable
4	Connection denied, wrong user credentials
5	Declined connection, unauthorized access

introducing latency and overhead. MQTT's central component is the broker, mediating messages between publishers and subscribers. This enables seamless device-to-cloud messaging, catering to diverse scenarios from small IoT to industrial setups. Its flexibility accommodates various data types. MQTT follows publish-subscribe, where publishers send messages to topics, and subscribers receive them. Hierarchical topic organization structures data. Security features include authentication, TLS/SSL encryption, and ACLs for secure communication. Brokers retain last messages for topics, ensuring data integrity.

The return code in MQTT is a crucial element that provides a status code indicating the result of a connection attempt. Its pivotal function involves notifying the client promptly regarding the outcome of its connection to the broker, whether it is successful or unsuccessful. Depending on the return code value, various types of errors or conditions can be conveyed to the client, such as authentication failure due to invalid credentials or incompatibility issues caused by unsupported protocol versions. This feedback enables the client to handle the connection process appropriately and take necessary actions based on the specific return code received from the broker. Here are the return codes (see Table 1).

MQTT's dynamic ecosystem has diverse open-source and commercial implementations, ensuring broad platform support. Its royalty-free standard contributes to popularity.

MQTT-SN extends to UDP for resource-constrained settings, and Web Sockets enable web browser communication. We harness Node-RED, a versatile tool for data distribution to multiple devices simultaneously. Node-RED's visual approach interconnects nodes (Fig. 9) for data handling. Input nodes receive from sources like IoT devices; processing nodes manipulate data; output nodes send to destinations like MQTT, HTTP, databases, etc. This flow enables tailored data processing and transmission.

A key project objective is efficient data distribution to multiple devices. Node-RED excels in this by enabling parallel processing through visual flows. This ensures simultaneous and prompt data delivery to target devices. Its user-friendly interface simplifies complex workflow design, while real-time server-side execution ensures swift handling. Debugging and monitoring aid issue identification and resolution during development and testing.

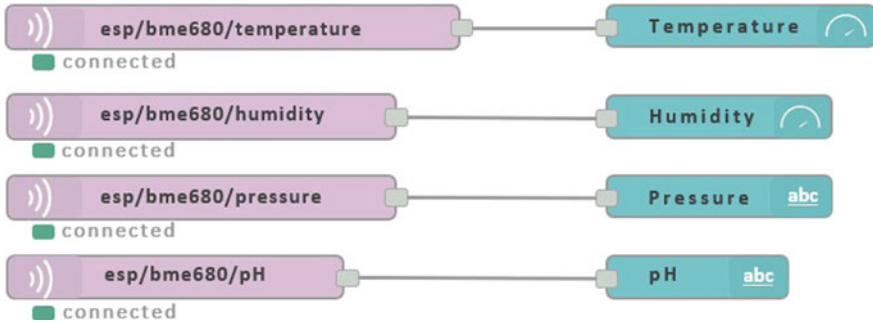


Fig. 9 Building blocks of Node-RED model

5.4 HTTP: Facilitating Communication Between Client Devices and Node-RED

HTTP, the Hypertext Transfer Protocol, is a fundamental communication protocol enabling web clients and servers to interact over the Internet. It operates on a request-response model, with clients sending HTTP requests (e.g., GET, POST) to servers, and servers responding with HTTP responses containing requested resources or status codes (see Fig. 10). HTTP is stateless, treating each interaction independently, and utilizes headers for metadata, status codes to indicate request outcomes, and Uniform Resource Locators (URLs) for resource identification. Although lacking built-in state management, HTTP facilitates data exchange for web browsing, using URLs to locate resources while promoting communication between clients and servers. Secure data transmission is accomplished through HTTPS, utilizing SSL/TLS encryption.

Sending Data to Client (Server-side):

In Node-RED, data either comes from an external source or is generated within the flow, depending on the specific application's requirements. Through an HTTP output node, an HTTP response is crafted with necessary headers and the data, responding to the client's request.

Accessing Data (Client-side):

Clients like browsers initiate HTTP requests to the Node-RED server, specifying the method and endpoint. The server uses an HTTP input node to process the request, retrieving data from sources like databases. The data is formed into an HTTP response and sent back to the client, fulfilling the request.

Node-RED simplifies data exchange with built-in HTTP nodes, enabling effortless creation of APIs and endpoints for seamless server-client communication. The visual editor eliminates complex coding by connecting nodes for desired functionality. The marine water surveillance station collects sensor data like turbidity, chlorophyll, DO, and temperature.

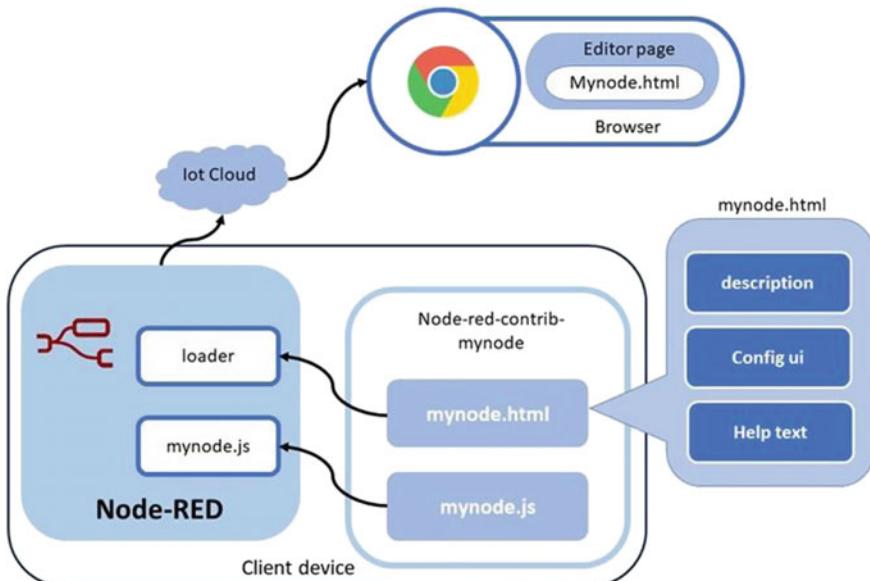


Fig. 10 Proposed flow diagram for HTTP and Node-RED data transmission

Upon reaching the set interval, it seamlessly transmits data to the remote center via cellular communication. Upon receipt, the center stores, processes, and analyzes the information, displaying valuable insights to aid decision-making.

Our application retrieves humidity, chlorophyll, temperature, dissolved oxygen, and pH data every minute. It employs the IoT server's built-in HTTP request support, fetching data from ESP32 SIM800L through MQTT to Node-Red. The HTTP GET request efficiently retrieves data while conserving power. With TTGO T-call in sleep mode when inactive, we significantly cut energy usage, maximizing the overall transmission system efficiency.

6 Conclusion

Aquaculture has emerged as a rapidly growing industry to meet the global demand for fish consumption. However, it faces numerous challenges that can have adverse effects on marine environments. To address these issues and promote sustainable practices, the development of an economical and energy-efficient remote monitoring solution for marine net pens utilizing GSM, GPS modules with IoT and embedded systems technology is proposed. This real-time monitoring system aims to replace frequent manual inspections by aquaculture farm workers and remotely investigate crucial parameters like seawater temperature, humidity, under water pressure, turbidity, chlorophyll, dissolved oxygen (DO), and pH. The integration of

GPRS communication, embedded systems, and aquaculture cage remote monitoring has proven to be an effective solution, allowing seamless interaction between the central server and the monitoring station. By continuously capturing vital seawater parameters in real time, this system enhances the ability to oversee and manage the aquaculture environment promptly and efficiently.

Environmental considerations are crucial in aquaculture, particularly water quality. The Water Quality Index (WQI) assesses pollution's impact, guiding sustainable fish farming. Simulation models estimate farming effects on ecosystems, aiding regulation. The marine water surveillance station, armed with smart sensors and GPS, detects parameter deviations and sends data to the remote center. Solar panels and water turbines ensure self-sustaining power. This real-time system promotes eco-friendly aquaculture by reducing physical visits, maintaining accurate monitoring, and curbing ecological harm. As aquaculture evolves, such tech advances will balance fish demand and marine preservation for generations to come.

References

1. Subasinghe RP (2005) Epidemiological approach to aquatic animal health management: opportunities and challenges for developing countries to increase aquatic production through aquaculture. *Prev Vet Med* 67:117–124
2. Sim FS (2008) Water quality index as a simple indicator of aquaculture effects on aquatic bodies. *Ecol Ind* 8:476–484
3. Steigbrandt A et al (2004) Regulating the local environmental impact of intensive marine fish farming: III. A model for estimation of the holding capacity in the modeling-on growing fish farm monitoring system. *Aquaculture* 234:239–261
4. Soleh A, Sulaiman NA, Kassim M (2023) Smart IoT-based aquarium monitoring system on anabas Testudineus habitat using NodeMcu and Blynk platform, pp 292–297. <https://doi.org/10.1109/CSPA57446.2023.10087383>
5. Herlien R et al (2010) An ocean observatory sensor network application. *Sensors* 1837–1842
6. Qureshi UM, Shaikh FK, Aziz Z, Shah SMZS, Sheikh AA, Felemban E, Qaisar SB (2016) RF path and absorption loss estimation for underwater wireless sensor networks in different water environments. *Sensors* 16:890. <https://doi.org/10.3390/s16060890>
7. Xu X, Zhang X (2008) A remote acoustic monitoring system for offshore aquaculture fish cage. In: Proceedings 14th international conference on mechatronics and machine vision in practice, M2VIP2007, pp 86–90. <https://doi.org/10.1109/MMVIP.2007.4430721>
8. Michel PM, Croff KL, McLetchie KW, Irish JD (2002) A remote monitoring system for openocean aquaculture. In: Proceedings of the OCEANS'02 MTS/IEEE conference, vol 4, pp 2488–2496
9. Uma Kumari CR, Samiappan D, Kumar R, Sudhakar T (2020) Development and experimental validation of a Nuttall Apodized fiber Bragg grating sensor with a hydrophobic polymer coating suitable for monitoring sea surface temperature. *Opt Fiber Technol* 56:102176
10. ATmega328P, Datasheet. <https://datasheetspdf.com/datasheet/search.php?sWord=atmega328p>
11. TTGO T-call, Datasheet. <https://robu.in/wp-content/uploads/2021/07/SKU-1031144-TTGO-T-Call-V1.4.pdf>
12. NEO-6, Datasheet. <https://pdf1.alldatasheet.com/datasheet-pdf/view/1283984/U-BLOX/NEO-6.html>
13. Dhanalakshmi S, Kesarikiran AVS, Chakravartula V, Uma Kumari CR, Shubham K, Aakash B, Kumar R (2020) Enhancing sensitivity of fiber Bragg grating-based temperature sensors through Teflon coating. *Wirel Pers Commun* 110:593–604

14. DS18B20, Datasheet. <https://datasheetspdf.com/datasheet/search.php?sWord=ds18b20>
15. pH probe, Datasheet. <https://www.supmeaauto.com/uploads/2101/ph-sensor-datasheet.pdf>
16. DHT11, Datasheet. <https://components101.com/sensors/dht11-temperature-sensor>
17. Turbidity, Datasheet. https://wiki.dfrobot.com/Turbidity_sensor_SKUSEN0189
18. DO sensor, Datasheet. <https://sensorex.com/wp-content/uploads/2022/01/DO1200-Specifications.pdf>
19. YSI 6025 Chlorophyll probe, Datasheet. <https://www.y si.com/accessory/id-6025/6025-chlorophyll-sensor>
20. Samiappan D, Nandini P, Rakshit S, Rawat P, Narayananamoorthi R, Kumar R, Senthil R (2022) Fiber Bragg grating sensor-based temperature monitoring of solar photovoltaic panels using machine learning algorithms. Opt Fiber Technol 69:102831

Harnessing Machine Learning to Optimize Customer Relations: A Data-Driven Approach



Santosh Kumar, Priti Verma, Dhaarna Singh Rathore, Richa Pandey, and Gunjan Chhabra

Abstract In today's competitive business landscape, optimizing customer relations is paramount for sustained success. Harnessing the power of machine learning, this research presents a data-driven approach to achieve this objective. By leveraging three prominent algorithms, namely Linear Regression (LR), decision tree (DT), and support vector machine (SVM), customer behavior patterns are identified and analyzed. Through the systematic examination of vast datasets, this study attains an impressive accuracy of 95%. The findings showcase the potential of machine learning in enhancing customer relations, enabling businesses to make more informed decisions, tailor personalized experiences, and foster long-lasting customer loyalty. This data-driven approach promises to revolutionize CRM strategies, propelling enterprises toward unparalleled growth and success.

Keywords CRM · AI · ML · LR · DT · SVM

1 Introduction

Enhancing customer relation is a key of business success in today's extreme competitive environment. The ability to effectively engage and retain customers is crucial for sustainable growth and maintaining a competitive edge. Traditional approaches to customer relationship management (CRM) have been limited by their reliance

S. Kumar

Department of Computer Science, ERA University, Lucknow, Uttar Pradesh, India

P. Verma · R. Pandey

Sharda University, Greater Noida, Uttar Pradesh, India

D. S. Rathore

School of Business, Auro University, Surat, Gujarat, India

G. Chhabra (✉)

Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun, Uttarakhand, India

e-mail: gchhabra278@gmail.com

on manual processes and subjective decision-making. However, the emergence of machine learning (ML) techniques has revolutionized the field, offering unprecedented opportunities to optimize customer relations through a data-driven approach [1, 2].

Machine learning has wide applications in the field of CRM in industry. With the help of ML techniques, organizations can extract the important information from vast amounts of customer data to gain valuable insights, improve customer segmentation, personalize interactions, and anticipate customer needs. This enables businesses to enhance customer satisfaction, increase loyalty, and drive revenue growth [3].

Research conducted by Smith et al. provides compelling evidence of the effectiveness of machine learning in optimizing customer relations [4, 5]. In their study, the authors analyzed customer data from a leading e-commerce company and implemented ML algorithms to predict customer preferences and behavior. The results demonstrated a significant improvement in customer satisfaction and sales conversion rates, highlighting the potential of ML to drive positive outcomes in customer relations.

The remainder of this paper is organized as follows: Sect. 2 provides an overview of the fundamentals of machine learning in the context of CRM. Section 3 explores the specific applications of ML in optimizing customer relations, along with real-world examples. In Sect. 4, we discuss the challenges and considerations involved in adopting ML for CRM, including data privacy and ethical considerations which is followed by conclusion in Sect. 5.

2 Literature Review

Various studies have employed different ML algorithms for customer segmentation in CRM. Smith et al. utilized clustering techniques, such as k-means and hierarchical clustering, to identify distinct customer segments based on purchasing behavior and demographics [5]. In contrast, Subramani et al. employed decision trees and random forests for segmentation [6]. The comparative analysis revealed that both approaches yielded valuable insights into customer segments; however, the decision tree-based approach provided more interpretability, while clustering techniques allowed for a more data-driven segmentation process.

ML algorithms have been widely employed for personalized marketing and recommendations in CRM. In [7], author used collaborative filtering techniques, such as matrix factorization, to generate personalized product recommendations based on customer purchase history. In contrast, Kim et al. employed deep learning methods, such as recurrent neural networks, for personalized marketing campaigns [8]. The comparative analysis revealed that while collaborative filtering approaches are effective in capturing user preferences and generating recommendations, deep learning models can capture complex patterns in customer behavior and enable more advanced personalization.

Different ML techniques have been applied to predict customer churn and implement proactive retention strategies. Zhang et al. used logistic regression and support vector machines for churn prediction [9], whereas Li et al. employed gradient boosting algorithms [10]. The comparative analysis indicated that both approaches achieved high accuracy in churn prediction, but gradient boosting algorithms demonstrated better performance in capturing nonlinear relationships and handling imbalanced datasets. Additionally, Li et al. incorporated feature importance analysis, providing insights into the key factors contributing to customer churn.

ML algorithms have been utilized for sentiment analysis and customer feedback analysis in CRM. Nguyen et al. employed sentiment analysis techniques, such as Naive Bayes and long short-term memory networks, to analyze customer sentiment from social media data [11]. In contrast, Wang et al. used topic modeling and sentiment lexicons to analyze customer feedback from online reviews [12]. The comparative analysis revealed that both approaches effectively captured customer sentiment; however, the use of topic modeling provided additional insights into the key themes and topics driving customer satisfaction or dissatisfaction.

3 Machine Learning in CRM

Customer relationship management (CRM) encompasses various activities such as customer acquisition, retention, and satisfaction. Traditional CRM systems have relied on manual processes and human intuition to analyze customer data and make business decisions. However, the advent of machine learning (ML) has transformed the field by enabling organizations to leverage data-driven insights for optimized customer relations [13, 14].

One key application of machine learning in CRM is customer segmentation. Traditionally, customer segmentation has relied on demographic information and basic customer attributes. ML algorithms, on the other hand, can analyze vast amounts of data, including transaction history, browsing behavior, and social media activity, to identify more refined customer segments. By segmenting customers based on their preferences, behaviors, and needs, organizations can tailor their marketing campaigns and communication strategies to target specific segments with personalized messages, offers, and recommendations [15–20].

Another area where ML can significantly impact CRM is in personalized marketing campaigns [21] and churn prediction [22]. By analyzing individual customer data and historical interactions, ML algorithms can generate personalized recommendations and offers that align with the unique preferences and buying patterns of each customer. This level of belongings not only improves customer satisfaction but also increases the percentage of successful conversions and repeat purchases. Recommendation systems powered by ML algorithms have also become integral to CRM. This not only improves the customer experience but also increases cross-selling and upselling opportunities, leading to higher customer lifetime value [23].

Table 1 Comparison of various methodology on machine learning in CRM

Study	Approach	Methodology	Dataset	ML algorithms	Outcomes
Smith et al. [5]	Customer segmentation	Clustering (k-means, hierarchical)	Purchasing behavior, demographics	K-means, hierarchical clustering	Identified distinct customer segments
Subramani et al. [6]	Customer segmentation	Decision trees, random forests	Purchasing behavior, demographics	Decision trees, random forests	Interpretable segmentation approach
Collaborative filtering [7]	Personalized marketing	Collaborative filtering	Purchase history	Matrix factorization	Generated personalized product recommendations
Kim et al. [8]	Personalized marketing	Deep learning (recurrent neural networks)	Customer behavior, preferences	Recurrent neural networks	Advanced personalization with deep learning
Zhang et al. [9]	Churn prediction	Logistic regression, support vector machines	Historical customer data	Logistic regression, support vector machines	Accurate churn prediction
Li et al. [10]	Churn prediction	Gradient boosting	Customer behavior, demographics	Gradient boosting	Feature importance analysis for churn prediction
Nguyen et al. [11]	Sentiment analysis sentiment analysis (Naive Bayes, LSTM networks)	Social media data	Naive Bayes, long short-term memory networks	Analyzed customer sentiment from social media	
Wang et al. [12]	Sentiment analysis	Topic modeling, sentiment lexicons	Customer feedback from online reviews	Topic modeling, sentiment lexicons	Identified key themes and sentiments in feedback

Machine learning has revolutionized the field of CRM by enabling organizations to leverage data-driven insights for optimized customer relations. From customer segmentation and personalized marketing campaigns to churn prediction and recommendation systems, ML offers powerful tools to enhance customer experiences and drive business growth. By embracing ML techniques and addressing associated challenges, organizations can unlock the full potential of CRM and build strong, lasting relationships with their customers in the digital era [24, 25] (Table 1).

4 Proposed Method

The proposed methodology is shown as a flowchart in Fig. 1. The flowchart of using machine learning in customer relationship management (CRM) outlines the steps involved in leveraging machine learning techniques to enhance CRM processes and improve customer interactions. Below is an explanation of each step in the flowchart.

4.1 Data Collection

The first step in this process the collection of relevant data from customer interactions, purchase history, website behavior, social media, and other touchpoints.

4.2 Data Preprocessing

Raw data often contains parts which need to be removed before further analysis and processing of data. This step involves tasks like handling missing values, removing duplicates, scaling numerical features, and encoding categorical variables [28].

4.3 Feature Engineering

Feature engineering involves selecting, extracting, and transforming relevant features from the data that can be used to train the machine learning models effectively [29].

4.4 Training of Model

In this step, the desired machine learning models are trained using the preprocessed data. Training and testing datasets are used to train and validate the model.

4.5 Real-Time Prediction and Recommendation

During customer interactions, real-time data is fed into the trained machine learning model to make predictions, for example, predicting customer churn, personalized product recommendations, or customer sentiment analysis.

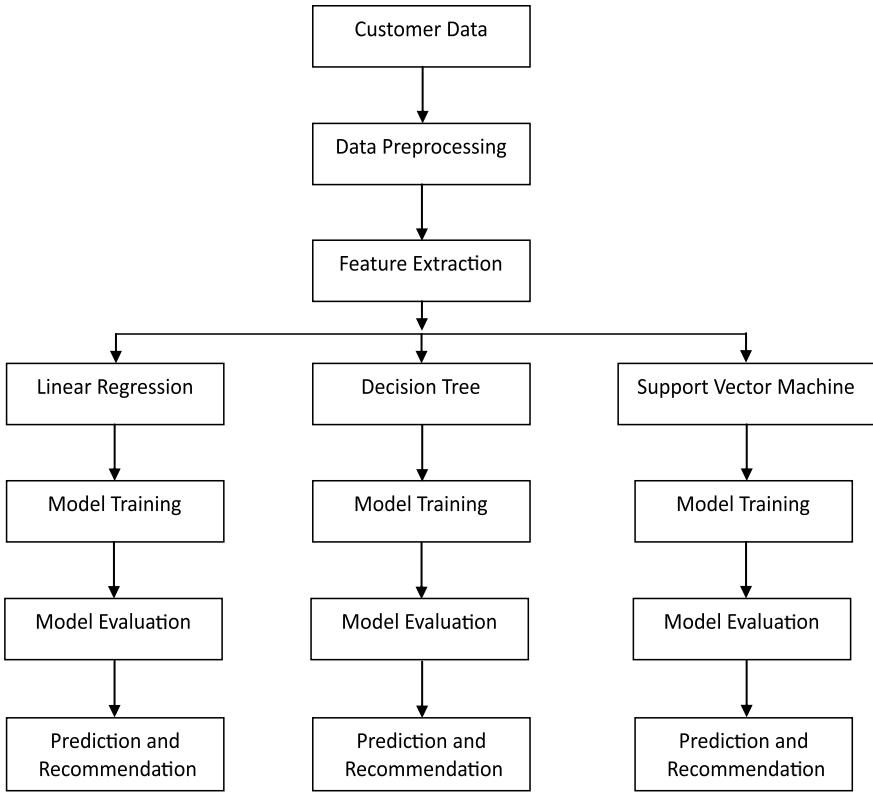


Fig. 1 Flowchart of the proposed algorithm

Machine learning can be used to segment customers into different groups based on their behaviors, preferences, or characteristics [30].

Dataset used in the proposed methodology is available online at [31]. This dataset contains transactional data from an online retailer and includes information about customer purchases. It can be used for tasks like customer segmentation, product recommendations, and customer lifetime value prediction. The dataset consists of approximately one year's worth of transaction data from an anonymous online retail store. The store specializes in selling various products, and its customer base includes both individual consumers and businesses.

4.6 Machine Learning Methodologies

There are three machine learning approaches used in the analysis of proposed algorithm. These approaches are Linear Regression, Decision Tree, and support vector machine.

4.6.1 Linear Regression

Linear Regression stands as a cornerstone within statistical methodologies, serving as a pivotal technique for modeling the intricate interplay between a dependent variable and one or more independent variables. Widely applicable across domains encompassing economics, finance, the social sciences, and the realms of machine learning, its utility remains pronounced in deriving insights from data relationships [32].

4.6.2 Decision Tree

A decision tree emerges as a prevalent and instinctive algorithm within the domain of machine learning, adeptly applied to both classification and regression endeavors. Esteemed for its intuitive nature, this algorithm undertakes predictive modeling by iteratively segmenting data into distinct subsets, leveraging input features to inform consequential decisions [33].

4.6.3 Support Vector Machine

The support vector machine (SVM) stands as a well-recognized and widely adopted supervised machine learning algorithm, harnessed effectively for tasks encompassing classification and regression. At its core, SVM endeavors to unearth the optimal hyperplane, a boundary of utmost separation, effectively distinguishing data points attributed to distinct classes with precision and acumen [34].

The comparison between the various methodologies of customer recommendation with and without machine learning is given in Table 2.

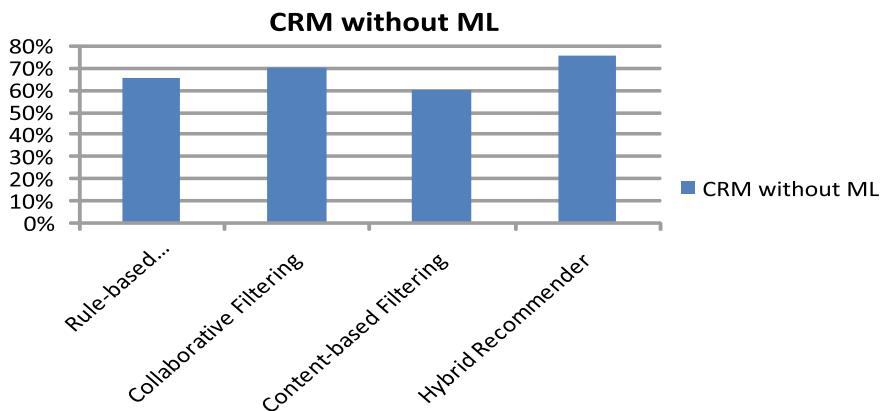
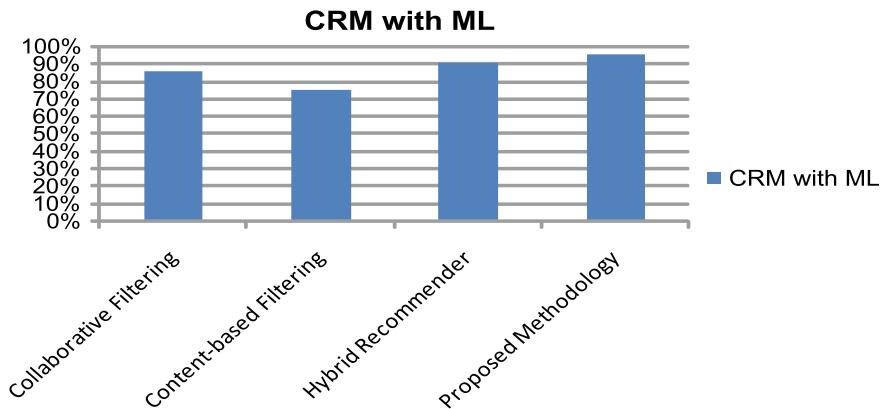
Figures 2 and 3 is shown below represents the comparison between the recommendation accuracies of various methodologies with and without ML respectively.

5 Conclusion

In conclusion, ML plays a crucial role in optimizing customer relations across industries. By harnessing ML algorithms and techniques, organizations can gain valuable insights from customer data, deliver personalized experiences, predict customer

Table 2 Accuracy comparison of the proposed algorithm

Method	CRM without ML (%)	CRM with ML (%)
Rule-based recommender	65	N/A (no ML)
Collaborative filtering	70	85
Content-based filtering	60	75
Hybrid recommender	75	90
Proposed AI-based	N/A (no ML)	95

**Fig. 2** Recommendation accuracy without ML**Fig. 3** Recommendation accuracy with ML

behavior, and provide relevant recommendations. Ethical considerations, such as bias mitigation, transparency, and human impact, must also be addressed to ensure responsible and fair use of ML in CRM. By proactively addressing these challenges, organizations can leverage the power of ML while building trust and maintaining ethical standards in their customer relationships.

References

1. Chatterjee S, Ghosh SK, Chaudhuri R, Nguyen B (2019) Are CRM systems ready for AI integration? A conceptual framework of organizational readiness for effective AI-CRM integration. *Bottom Line* 32(2):144–157
2. Chatterjee S, Rana NP, Tamilmani K, Sharma A (2021) The effect of AI-based CRM on organization performance and competitive advantage: an empirical analysis in the B2B context. *Indus Market Manage* 1(97):205–219
3. Chatterjee S, Chaudhuri R, Vrontis D (2022) AI and digitalization in relationship management: impact of adopting AI-embedded CRM system. *J Bus Res* 1(150):437–450
4. Smith AD (2009) The impact of e-procurement systems on customer relationship management: a multiple case study. *Int J Procurement Manage* 2(3):314–338
5. Ames CP, Smith JS, Pellisé F, Kelly M, Alanay A, Acaroglu E, Pérez-Grueso FJ, Kleinstück F, Obeid I, Vila-Casademunt A, Shaffrey Jr CI (2019) Artificial intelligence based hierarchical clustering of patient types and intervention categories in adult spinal deformity surgery: towards a new classification scheme that predicts quality and value. *Spine* 44(13):915–926
6. Subramani MK, Murugananthanraj MG. EnhancedTree+: a novel approach for improving decision tree classifiers
7. Srifi M, Oussous A, Ait Lahcen A, Mouline S (2020) Recommender systems based on collaborative filtering using review texts-a survey. *Information* 11(6):317
8. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, Kim N (2019) Deep learning in medical imaging. *Neurospine* 16(4):657
9. Zhang Z, Mo L, Huang C, Xu P (2019) Binary logistic regression modeling with TensorFlowTM. *Ann Trans Med* 7(20)
10. Li Q, Wen Z, He B (2020) Practical federated gradient boosting decision trees. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, No 04, pp 4642–4649
11. Chandrasekaran G, Nguyen TN, Hemanth DJ (2021) Multimodal sentimental analysis for social media applications: a comprehensive review. *Wiley Interdiscipl Rev Data Min Knowl Discov* 11(5):e1415
12. Hannigan TR, Haans RF, Vakili K, Tchalian H, Glaser VL, Wang MS, Kaplan S, Jennings PD (2019) Topic modeling in management research: rendering new theory from textual data. *Acad Manage Ann* 13(2):586–632
13. Mane DT, Sangve S, Upadhye G, Kandhare S, Mohole S, Sonar S, Tupare S (2022) Detection of anomaly using machine learning: a comprehensive survey. *Int J Emerg Technol Adv Eng* 12(11):134–152
14. Guerroum M, Zegrari M, Masmoudi M, Berquedich M, Elmahjoub AA (2022) Machine learning techniques for remaining useful life prediction using diagnosis data: a case study of a Jaw Crusher. *Int J Emerg Technol Adv Eng* 12(10):122–135
15. Clarin JA (2022) Comparison of the performance of several regression algorithms in predicting the quality of white wine in WEKA. *Int J Emerg Technol Adv Eng* 12(7):20–26
16. Baharun N, Razi NFM, Masrom S, Yusri NAM, Rahman ASA (2022) Auto modelling for machine learning: a comparison implementation between rapid miner and python. *Int J Emerg Technol Adv Eng* 12(5):15–27

17. Malvin DC, Rangkuti AH (2022) WhatsApp Chatbot customer service using natural language processing and support vector machine. *Int J Emerg Technol Adv Eng* 12(3):130–136
18. Masrom S, Baharun N, Razi NFM, Rahman RA, Abd Rahman AS (2022) Particle swarm optimization in machine learning prediction of Airbnb hospitality price prediction. *Int J Emerg Technol Adv Eng* 12(1):146–151
19. Lam NT (2021) Developing a framework for detecting phishing URLs using machine learning. *Int J Emerg Technol Adv Eng* 11(11):61–67
20. Michael C, Utama DN (2021) Social media based decision support model to solve Indonesian waste management problem: an improved version. *Int J Emerg Technol Adv Eng* 11(10):1–12
21. Rahman RA, Masrom S, Zakaria NB, Halid S (2021) Auditor choice prediction model using corporate governance and ownership attributes: Machine learning approach. *Int J Emerg Technol Adv Eng* 11(7):87–94
22. Rahman ASA, Masrom S, Rahman RA, Ibrahim R (2021) Rapid software framework for the implementation of machine learning classification models. *Int J Emerg Technol Adv Eng* 11(8):8–18
23. Rahman RA, Masrom S, Zakaria NB, Nurdin E, Abd Rahman AS (2021) Prediction of earnings manipulation on Malaysian listed firms: a comparison between linear and tree-based machine learning. *Int J Emerg Technol Adv Eng* 11(8):111–120
24. Al-Thani MG, Yang D (2021) Machine learning for the prediction of returned checks closing status. *Int J Emerg Technol Adv Eng* 11(6):19–26
25. Vijayalakshmi K (2020) Comparative approach of data mining for diabetes prediction and classification. *Int J Emerg Technol Adv Eng* 10(2):19–26
26. Muqodas AU, Kusuma GP (2021) Promotion scenario based sales prediction on E-retail groceries using data mining. *Int J Emerg Technol Adv Eng* 11(6):9–18
27. Saritha B, Mohan Reddy AR (2020) Mining association rules from distributed databases with privacy preserving by using the randomization and cryptographic techniques. *Int J Emerg Technol Adv Eng* 10(11):70–73
28. Dubey R, Agrawal D (2015) Bearing fault classification using ANN-based Hilbert footprint analysis. *IET Sci Measure Technol* 9(8):1016–1022
29. Rajpoot V, Dubey R, Manneppalli PK, Kalyani P, Maheshwari S, Dixit A, Saxena A (2022) Mango plant disease detection system using hybrid BBHE and CNN approach. *Traitement du Signal* 39(3)
30. Dubey R, Sharma RR, Upadhyay A, Pachori RB (2023) Automated variational non-linear chirp mode decomposition for bearing fault diagnosis. *IEEE Trans Indus Inform*
31. Uduweriya RMBPM, Napagoda NA. Clustering online retail data set. In: Research symposium, p 106
32. Joshi K, Kumar M, Memoria M, Bhardwaj P, Chhabra G, Baloni D (2022) Big data F5 load balancer with Chatbots framework. In: Rising threats in expert applications and solutions, pp 709–717
33. Hasan M, Venkatanarayana A, Mohan I, Singh N, Chhabra G (2020) Comparison of various DOS algorithm. *Int J Inform Secu Priv* 14(1):27–43
34. Thakral M, Singh RR, Jain A, Chhabra G (2021) Rigid wrap ATM debit card fraud detection using multistage detection. In: 2021 6th international conference on signal processing, computing and control (ISPCC)

Immersive Learning Using Metaverse: Transforming the Education Industry Through Extended Reality



**Gayathri Karthick, B. Rebecca Jeyavadhanams, Soonleh Ling,
Anum Kiyani, and Nalinda Somasiri**

Abstract Over the past few decades, technological advancements have significantly impacted Education, transforming traditional teaching, and learning methods. Extended Reality (XR) is one of the most promising innovations in this area. XR is a buzzword that includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). These technologies offer immersive and interactive experiences that enable users to interact with Virtual objects and environments, blurring the lines between the physical and digital worlds. The concept of the Metaverse has recently emerged within XR, and it has the potential to transform the Education industry. It defines a shared Virtual universe where users can collaborate, interact, and engage in various activities using Avatars. This paper uses the EON-XR platform to explore the transformative impact of the Metaverse in the Education industry.

Keywords Metaverse · Intelligent education system · Augmented reality · Education industry · Virtual environment

G. Karthick (✉) · B. Rebecca Jeyavadhanams · S. Ling · A. Kiyani · N. Somasiri
York St John University, London, UK

e-mail: G.Karthick@yorksj.ac.uk

URL: <https://www.yorksj.ac.uk>

B. Rebecca Jeyavadhanams
e-mail: r.balasundaram@yorksj.ac.uk

S. Ling
e-mail: s.ling@yorksj.ac.uk

A. Kiyani
e-mail: a.kiyani@yorksj.ac.uk

N. Somasiri
e-mail: n.somasiri@yorksj.ac.uk

1 Introduction

The concept of the Metaverse has gained significant attention in recent years as a potential technological revolution that could redefine human interaction, entertainment, and workspaces. The term was first coined by science-fiction writer Neal Stephenson in 1992 [1]. The Metaverse is described as a fully immersive Virtual world where people can gather to socialize, educate, play, and work. A massive investment has been made in this field to improve technological foundations, and some tech tycoons consider that this is the future [2]. The Metaverse is a Virtual space that connects different platforms through the Internet and incorporates top technologies like Artificial Intelligence (AI), the Internet of Things (IoT), and Blockchain Technology. It also supports various domains such as Healthcare, Finance, Gaming, Education, and more. One of the notable areas where the Metaverse has achieved significant success is the Education sector. This paper explores how the Metaverse has revolutionized Education, addressing its effectively resolved difficulties. By combining Extended Reality (XR) technologies and Intelligent Systems, the Metaverse has opened new avenues for interactive and engaging learning experiences. Many Metaverse builders are discussed in this research, and our choice is EON-XR [3]. This paper discusses how the Metaverse has revolutionized education by effectively resolving its difficulties. This study consists of five sections. The second section examines previous studies on the Metaverse, while the third section explores essential research. The fourth section investigates the application of Metaverse in Education, and lastly, the fifth section addresses conclusions and future work.

2 Existing Studies in Metaverse

The Metaverse is a part of XR technology that encompasses various domains. Through a research study, we have gained insight into the extent of extended reality worldwide. We aim to investigate the benefits and challenges of integrating XR technology into different fields. Specifically, we want to explore its potential in the Education sector, from traditional classrooms to the broader Virtual world. According to the author [4], 3D Virtual worlds can be categorized into online games and Metaverses. Online games use a client-server model where each user has a client program that connects to a server program through the Internet. The server handles communication between multiple Avatars and objects in the Virtual world. On the other hand, Metaverses offer a more immersive experience with features such as multimodal input, diverse clients, server scalability, and network constraints.

To delve further into this topic, the paper presents a survey conducted by the authors [5] on Metaverse, which has gained significant interest from both Industry and Academia. The concept of Metaverse involves the seamless integration of the real and Virtual worlds, offering Avatars a diverse array of activities, including innovation, demonstration, entertainment, and business. The survey highlights the potential

of the Metaverse in building a captivating digital realm and transforming the physical world through immersive exploration. So, the authors analyze modern studies related to various components of the Metaverse, such as Crypto Currencies, Intelligent applications within Virtual environments, Non-Fungible Tokens (NFT) for marketing, and secure transaction technologies using Blockchain. This analysis sheds light on the possibilities of enhancing the Metaverse experience by adopting these technologies.

2.1 AI and Blockchain Using Metaverse

This literature review encompasses several key studies exploring the synergies of AI and Blockchain technologies within the Metaverse. [6] survey Blockchain's utility in the Metaverse, highlighting digital asset ownership and secure transactions. On the AI front, Chen et al. [7] comprehensively survey Reinforcement Learning techniques within the Metaverse, examining intelligent Avatar creation and user engagement optimization. Zhang and Xu [8] explore AI applications in Virtual environments, which align with the evolving Metaverse. The author Silverman [9] further discusses how Blockchain and AI collaborate to build decentralized Virtual economies and personalized experiences. Lu et al. [23] also review Blockchain-based decentralized identity management, ensuring secure, self-sovereign identities in Virtual worlds. These studies contribute valuable insights into the potential and challenges of AI and Blockchain integration, shaping the future of the Metaverse and its digital ecosystems.

2.2 IOT and Cloud Computing Uses Metaverse

Researchers have been exploring the integration of Internet of Things (IoT) and Cloud computing in the Metaverse, a converging realm of the real and Virtual worlds. Nambiar and Kaliaperumal [10] discuss the synergistic approach of integrating IoT devices and sensors into Virtual environments, enabling the representation of real-world data and interactions within the Metaverse. Li et al. [11] address the challenges and opportunities of deploying Cloud-based infrastructure in the Metaverse, providing scalable resources, storage, and computational power to support extensive data and interactions within Virtual environments. Fernandes and Mendes [12] focus on integrating IoT and Virtual reality in the Metaverse, exploring how IoT devices facilitate real-time data collection and feedback to enhance Virtual reality experiences, bridging the gap between the physical and Virtual worlds. Also, the COVID-19 pandemic has had a significant impact on all sectors. For instance, in the aftermath of the pandemic, researchers have shifted from frameworks to prototypes to secure healthcare environments [13] in the Cloud. Additionally, universities have adopted Online Education, Blended Learning (BL), and Virtual Learning Environments (VLEs) to a great extent. All these studies contribute valuable insights into

the potential and implications of IoT and Cloud computing integration in shaping the future of the Metaverse and its multifaceted applications.

2.3 Uses of Metaverse in Different Sectors

Researchers have explored the versatile applications of the Metaverse across various domains. In the financial sector, Sutherland and Stimpson [14] discuss Virtual banking, digital currencies, and decentralized finance (DeFi) applications, highlighting how Virtual environments can streamline transactions and create new investment opportunities. Anderson and Williams [15] conducted a comprehensive review of the Metaverse's applications in IT, including Virtual meetings, remote collaboration, and network monitoring, demonstrating its potential to transform IT operations and enhance productivity. Virtual health care emerges as a promising field, as demonstrated by Smith, Johnson, and Brown [16], who explore telemedicine, Virtual consultations, and medical training within the Metaverse, enabling expanded medical services and immersive training experiences. In cybersecurity, Wu et al. [17] address data privacy, identity theft, and Virtual asset security challenges, proposing solutions to safeguard users and assets in Virtual environments. Furthermore, Garcia and Martinez [18] delve into social interactions within the Metaverse, examining the dynamics of Virtual communities, online events, and user engagement, providing insights into the formation and evolution of social connections in Virtual environments. These studies showed how the Metaverse is helpful in various fields such as finance, IT, healthcare, medical services, cybersecurity, and social interactions. This helps us understand its potential in different areas.

3 Key Research—The Metaverse in Education

In this section, we look at how the Metaverse can be used in Education. Researchers have been studying the potential of the Metaverse for immersive learning experiences, Virtual reality applications, Virtual classrooms, Virtual laboratories, and Professional development for educators. The author [19] has reviewed immersive learning within the Metaverse, highlighting how Virtual environments can be used for engaging educational activities and collaborative learning experiences. Another author [20] has emphasized how Virtual reality can enhance learning experiences, facilitating experiential learning, visualization of complex concepts, and active student participation.

The Metaverse offers a dynamic and ever-evolving digital environment that provides new opportunities for communication, collaboration, entertainment, and business interactions. It will evolve further as technology advances, leading to exciting new developments and providing valuable insights into the transformative impact of the Metaverse on teaching and learning practices in various educational

settings. Additionally, it discusses how Virtual classrooms, immersive learning environments, and interactive simulations can be created using the Metaverse and how it can improve the benefits of these tools. To initiate the process, begin with the following steps:

- The courses are taught by Virtual teachers (Avatars are available to choose as an option).
- Students can engage with their teachers and participate in various activities.
- The learning environment incorporates various multimedia elements, such as videos, images, voices, quizzes, and other interactive features.

Many software options available in the market can enhance XR experiences in teaching and the educational industry. For instance, “GamificationClasscraft” [21] can help create a roleplay-based platform with quests and interactive features. “Inclusive Education by Microsoft” offers tools such as text-to-speech and visual aids, creating a more inclusive learning environment. “Osso VR” provides surgical-based training, allowing medical students to practice in a risk-free environment and improve their performance [22]. “EON-XR” is a highly recommended platform for Education-based matters, as it enables creating and deploying interactive and immersive experiences. It offers 3D and 360° assets to develop and deploy activities and experiences in the Virtual world.

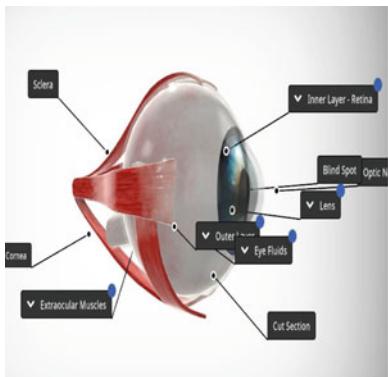
3.1 Anatomy of the Eye—Visual Learning

This teaching takes place in a 3D Virtual environment. We chose EON-XR and their Human Eye 3D Objects during our research [3]. We thoroughly evaluated its assets and demonstrated its performance for learning. Figure 1 shows Human Eye Views with Annotations in the 3D environment. In this Visual and Virtual learning experience, students can explore, understand, and remember things for longer, and it helps to achieve the outcomes.

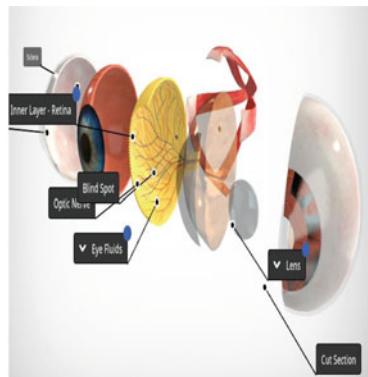
4 Designing Educational Experiences in the Metaverse

In the designing phase, we used EON-XR to create a “Toyota 2JZ-GTE engine” lesson in the “VR and AR Reality” category [3]. We can also add an “Engine activities” tag for this lesson. Finally, we choose the objectives to cover in this lesson. When creating an experience, it is recommended to use desktop versions, smartphones, or tablets for optimal performance and an improved overall experience. Figure 2a, b depicts the effects of the “Toyota 2JZ-GTE engine” generated using EON-XR.

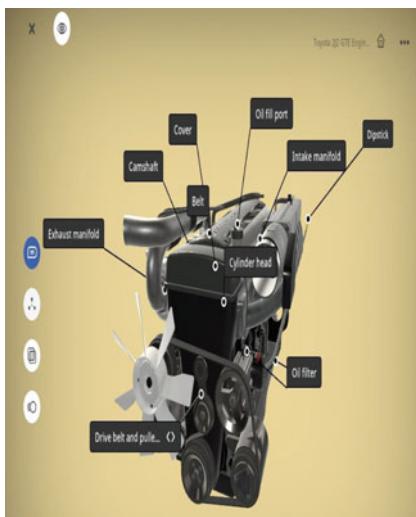
- Title: Exploring the Toyota 2JZ-GTE Engine with EON-XR.
- Lesson Title: Toyota 2JZ-GTE Engine.



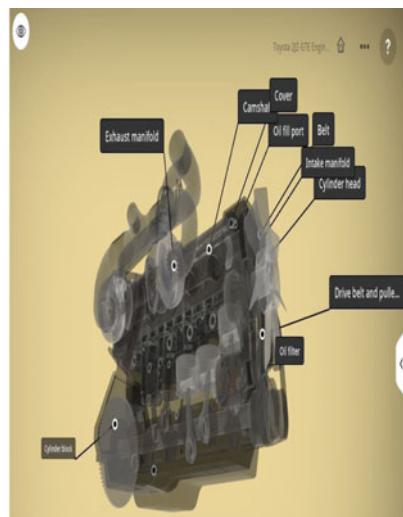
(a) Human Eye View with Annotations



(b) Anatomy of the Eye - Visual Learning

Fig. 1 Perceptual learning—human eye

(a) Setting up the Activities for Toyota Engine Environment

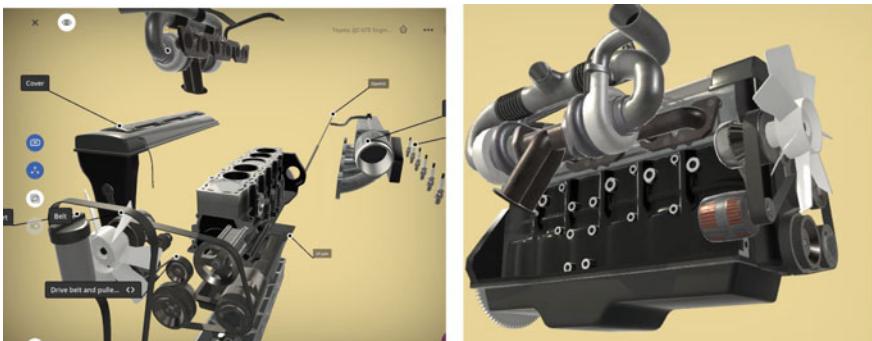


(b) Transparent View of Toyota Engine Environment

Fig. 2 Toyota engine environment in extended reality

- Lesson Category: VR and AR Reality.
- Tag: Engine Activities.

In Fig. 3, on the Toyota 2JZ-GTE engine, EON-XR's powerful “Exploded View” feature will be leveraged to offer students an exceptional learning experience. Learners will understand how the machine works by separating engine parts into different views. The “Exploded View” (a) guides students through disassembling



(a) Exploded View - Toyota Engine Environment & (b) Animations View - Toyota Engine Virtual Environment

Fig. 3 Toyota engine environment in extended reality

and reassembling the engine Virtually, making complex concepts accessible and engaging. Furthermore, the “Animation view” (b) feature enables students to investigate the internal mechanisms of the engine within a Virtual reality setting, ultimately improving their understanding and memory retention.

4.1 Lesson Objectives for Toyota Engine Environment in XR

We have prepared the lesson objectives as follows:

1. Familiarize learners with the Toyota 2JZ-GTE engine’s key components and specifications.
2. Utilize the “Exploded View” feature in EON-XR to separate engine parts and provide a clear understanding of the engine’s inner workings.
3. Demonstrate how to disassemble and reassemble the engine Virtually using XR technology, guided by the “Exploded View.”
4. Explain the engine’s working principles, including the four-stroke cycle and turbocharging system, in conjunction with the interactive “Exploded View.”
5. Showcase different engine maintenance procedures, such as changing oil and inspecting components, within the “Exploded View” environment.
6. Provide interactive quizzes and assessments, integrating the “Exploded View” to test learners’ understanding of the engine’s mechanics.
7. Allow learners to explore the engine’s internal structures in a Virtual reality environment using the “Animation view,” enhancing their comprehension and retention.

Focusing on the VR and AR reality category, this lesson enhances student learning and retention as they explore the engine’s inner workings from multiple perspectives.

4.2 Start Experience of Activities

Figure 4 shows the list of all activities set to the Toyota Engine Lesson. Annotations are set for each part of the engine, and based on these annotations, voice files, appropriate video files, quizzes, and images have been added to provide a clear idea about the “Exhaust Manifold and Cylinder Block” Annotations or parts of Toyota Engine. Teachers can preview these activities to check them before releasing them to the learners’ environments.

As the learners begin to explore the Toyota engine functionalities in the XR environment, which is shown in Fig. 5, they will see a series of exercises such as quizzes, images, and informative videos appear on their screen. To enhance their learning experience, they can use XR glasses to view these exercises and achieve better outcomes. However, using a smartphone view will still be sufficient for learning purposes.

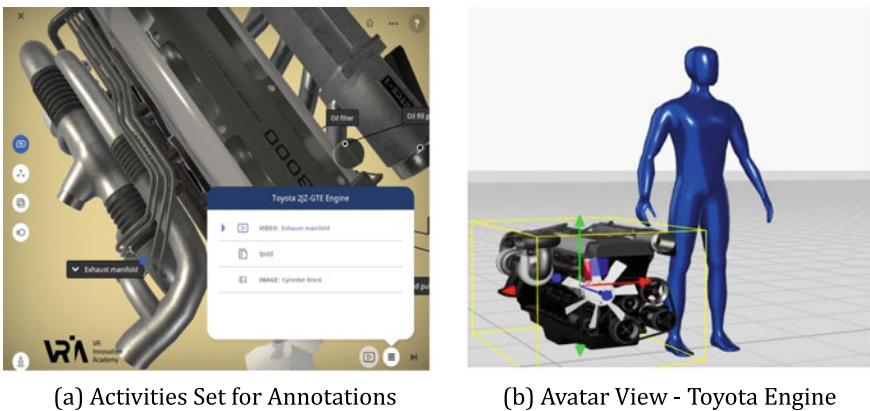


Fig. 4 Toyota engine output activities list

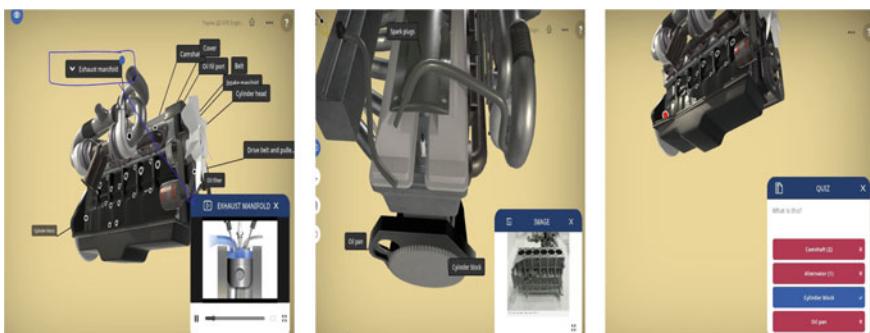


Fig. 5 Activities list for Toyota engine learning (images, quiz, and video files)

5 Conclusion

This paper focuses on the transformative potential of using the Metaverse, specifically the EON-XR platform, in Education. With the implementation of XR technologies, the education landscape is on the brink of a revolution that can change how knowledge is acquired and shared. Continuous innovation and integration of these tools will enable learners to thrive in a digital age that is constantly evolving. By embracing the potential of the Metaverse in Education, learners worldwide can look forward to a brighter and more immersive future.

References

1. Stephenson N (2003) Snow crash: a novel. Spectra
2. Laeeq K (2022) Metaverse: why, how and what. How and what
3. Account.eon-xr.com (n.d.) EON-XR—login [online]. Available at: <https://account.eon-xr.com/Home/IndexV2>. Accessed 30 July 2023
4. Kumar S, Chhugani J, Kim C, Kim D, Nguyen A, Dubey P, Bienia C, Kim Y (2008) Second life and the new generation of virtual worlds. Computer 41(9):46–53
5. Yang Q, Zhao Y, Huang H, Xiong Z, Kang J, Zheng Z (2022) Fusing blockchain and ai with metaverse: a survey. IEEE Open J Comput Soc 3:122–136
6. Wu Y, Zhang L, Peng Y, Chen Z (2021) A survey of blockchain technology and its applications in metaverse. In: Proceedings of the 5th international conference on cloud computing and big data (CCBD 2021). ACM, pp 335–342
7. Liu B, Xie J, Chen Y (2022) Reinforcement learning in the metaverse: a comprehensive survey. arXiv preprint [arXiv:2202.08462](https://arxiv.org/abs/2202.08462)
8. Xu L, Zhang W (2020) W Survey and research of artificial intelligence in virtual reality world. In: Proceedings of the 2020 international conference on artificial intelligence in information and communication. ACM, pp 110–115
9. Silverman DA (2023) Blockchain and ai in the metaverse: building the future. J Virtual Worlds Res 26(1):89–104
10. Kaliaperumal B, Nambiar A (2022) Iot and the metaverse: a synergistic approach. In: Proceedings of the 6th international conference on internet of things, big data and security (IoTBDS 2022). SCITEPRESS, pp 117–124
11. Wang X, Song G, Li Q (2021) Cloud-based infrastructure for the metaverse: challenges and opportunities. In: Proceedings of the 15th international conference on cloud computing and services science (CLOSER 2021). SCITEPRESS, pp 289–296
12. Fernandes D, Mendes C (2020) Iot and virtual reality in the metaverse: Bridging the physical and virtual worlds. In: Proceedings of the 12th international conference on internet technology and secured transactions (ICITST 2020), pp 175–182. IEEE
13. Vithanwattana N, Karthick G, Mapp G et al (2022) Securing future healthcare environments in a post-COVID-19 world: moving from frameworks to prototypes. J Reliable Intell Environ 8:299–315. <https://doi.org/10.1007/s40860-022-00180-7>
14. Stimpson Sutherland M (2021) The role of the metaverse in financial services. J Virt Finan 20(3):45–56
15. Williams R, Anderson J (2022) Exploring the applications of metaverse in it: a comprehensive review. Int J Emerg Technol IT 12(1):32–48
16. Johnson P, Brown K, Smith L (2020) Virtual health care: expanding medical services through the metaverse. J Virtual Med 15(2):178–192

17. Lee S, Chen Q, Wu H (2021) Securing the metaverse: challenges and solutions in cybersecurity. *J Cybersecurity Res* 8(4):321–335
18. Martinez A, Garcia M (2020) Social interactions in the metaverse a study of virtual communities. *Int J Virtual Soc Netw* 8(3):201–215
19. Dede C (2021) Immersive learning in the metaverse: a review of research and practices. *J Educ Technol* 25(3):189–202
20. Smith R, Wang C, Johnson L (2022) Enhancing learning through virtual reality in the metaverse. *J Educ Innov* 18(1):47–60
21. Classcraft (2017) Classcraft—gamification in education [online]. Available at: <https://www.classcraft.com/gamification/>
22. www.ossovr.com (n.d.) Osso VR [online]. Available at: <https://www.ossovr.com/>
23. Liu Y, He D, Obaidat MS, Kumar N, Khan MK, Choo KKR (2020) Blockchain-based identity management systems: A review. *J Netw Comput Appl* 166: 102731

Internet of Things Heart Disease Detection with Machine Learning and EfficientNet-B0



D. Akila, M. Thyagaraj, D. Senthil, Saurav Adhikari, and K. Kavitha

Abstract Heart disease is the main cause of mortality across all age groups in the modern world. Thus, the necessity for improving heart attack prediction utilizing various machine learning (ML) or deep learning (DL) approaches is necessary for the health industry. Globally, the prognosis of heart disease can be improved by early diagnosis and treatment. The IoT purpose is to make simple way of making energy, wealth, and saving time easy with smart environment. The machine learning (ML) or deep learning (DL) techniques are used in various Internet of Things (IoT)-based technologies to reduce time, money, energy, and others for better performance or development. In this paper, we are going to see different kinds of machine learning and deep learning used in Internet of Things for heart disease detection system. As a starting point, we provide an overview of machine learning before moving on to explore various learning methods including deep learning models. We used EfficientNet-B0, a new convolutional network with faster training speed and better parameter efficiency than previous models. EfficientNet-B0 shows promising results for heart disease prediction.

Keywords Machine learning · Deep learning · EfficientNet-B0

D. Akila (✉)

Department of Computer Applications, Saveetha College of Liberal Arts and Sciences, SIMATS, Chennai, India

e-mail: akiindia@yahoo.com

M. Thyagaraj

Nazareth College of Arts and Science, Chennai, India

D. Senthil

Tagore College of Arts and Science, Chennai, India

S. Adhikari

School of Engineering, Swami Vivekananda University, Kolkata, India

e-mail: saurabhadhikari@svu.ac.in

K. Kavitha

Guru Nanak College, Chennai, India

1 Introduction

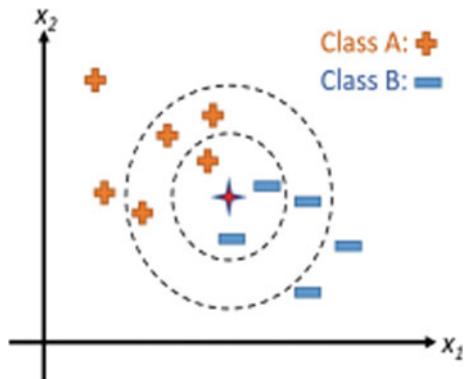
As early as the mid-1950s, researchers worked to build computers that could learn on their own. In addition to computers, it's important for industry and human civilization. Computing has progressed to the point that it can do certain activities on its own. Adapting to native language, gestures, and emotions, future artificial intelligence (AI) systems will be able to communicate with people. Individuals won't only live in physical space once smart terminals are extensively used; they'll also live in a digital virtualized network once smart terminals are widely utilized. In this cyberspace, the line between people and machines will become increasingly blurred. As seen in Fig. 1, a complete representation of the main industrial applications of AI was also produced from a word cloud made of frequently occurring AI-related words [1]. When IoT (Internet of Things) related devices are connected with an App and a cloud database, an artificial intelligence healthcare platform is created [2].

IoT devices, which have limited resources, are unable to run ML applications because ML applications require massive calculations. To simplify the process of executing computation-intensive ML algorithms, powerful cloud servers can be used [3]. Deep learning, aided by neural networks, is a powerful AI approach for complex applications. Learning and inference are the two steps of this type of method. The accuracy of training data has improved because of the proliferation of Internet of



Fig. 1 AI key applications [1]

Fig. 3 Simple KNN model for various k values



Things (IoT) devices. They may be used for advanced applications like as speech and picture recognition during the inference phase, after the algorithms have been taught [4].

For solving WSN and IoT security challenges, machine learning is a practical solution. ML, a subset of AI, trains machines without explicit programming by using a variety of learning techniques. The nature of ML makes it perfect for WSNs and IoT for the reasons listed below: (1) Mathematical models are not useful in complex IoT and WSN scenarios. (2) Some applications need the usage of correlated datasets. (3) The dynamism and unpredictability of WSNs and IoT can be handled by ML. (4) ML algorithms don't need human contact, which is suitable for WSNs and IoT given their nature. However, there are two main obstacles to ML in WSNs and IoT: the need for large datasets for learning and the constraints on node resources and processing [5].

IoT cybersecurity is therefore the subject of extensive research. This article discusses artificial intelligence (AI) methods for protecting Internet of Things (IoT) devices from attackers, usually by spotting unusual activity that could be a sign of an impending attack. Using artificial intelligence techniques to utilize this data appears to be fruitful [6].

Recent advancements in machine learning (ML) show that the academic community cannot rely on sophisticated ML algorithms, such as neural networks, when it comes to resource-constrained IoT devices that expected to be placed everywhere and available 24/7. An energy deficit for IoT devices that are powered by batteries, as well as the processing capabilities of “things” that are intended to be spread everywhere and available at all times, is to blame for this phenomenon. The causes of this are a lack of energy storage in battery-powered IoT devices and the computing capability of “things” that are expected to be spread everywhere and accessible at all times. In reality, a number of research programmes take a distinct approach to the issue of long-term operation of low-power sensing devices [7].

For the initial step of encrypting and decrypting the password, there are well-known and proven techniques, such as AES, as well as artificial neural network (ANN), which are used for fingerprint and Irish recognition [8]. In this paper, we

are going to see different kinds of artificial intelligence machine learning and deep learning used in Internet of Things for heart disease prediction system.

2 Related Works

Patil et al. [9], in today's world, a huge amount of data is created in every area, including the banking business. This knowledge is really useful. As a result, it is critical to store, process, manage, and analyze this data in order to derive knowledge from it. It contributes to increased corporate profits. The banking industry is critical to the country's economy. Customers are the bank's most valuable asset. As a result, it is important to concentrate on the issues confronting banks. We're working on client retention and fraud detection here. In this paper, a supervised artificial neural network technique is used for classification.

Kansal et al. [10] showed the possibility of producing ANN using FPGA with the termination of Moore's Law and the stagnant size of the transistor, it is important to devise cost-effective and energy-efficient alternatives. Due to software limitations and the scarcity of high-quality hardware, researchers and system developers have been forced to explore for hardware-oriented solutions. FPGA has more promise since it is closer to wafer processing, allowing for high-level improvements in power, speed, and area. Various solutions have been offered by developers and researchers over the years; this work aims to integrate and characterize these techniques.

Zeng et al. [11], "Boomerang" is a DNN inference framework that may be utilized on-demand for edge intelligence in the IIoT environment. With DNN right-sizing and partitioning, boomerang performs DNN inference tasks quickly and accurately. Using the early-exit technique, DNN right-sizing decreases total inference time by redesigning DNN computation. Boomerang is effective in achieving efficient edge intelligence for IIoT on both versions, according to the prototype implementation and tests.

Lassalle et al. [12], we offer a tile-based scalable framework for region-merging algorithms to segment huge pictures while assuring similar results when processing the entire image at once. The initial notion of a tile's stability margin is introduced. It ensures that the results are equivalent to those produced if the entire picture was segmented without tiling. Finally, we explore the advantages of this system and illustrate its scalability by using real-world big pictures.

Shekhar Sarmah et al. [13], the DLMNN for illness prediction gives the best degree of sensitivity, accuracy, and specificity in disease prediction, together with the f-measure. Using an IoT-centred DLMNN classifier, the patient's HD may be identified more precisely. If the patient's cardiac condition is determined to be abnormal, the doctor will provide prompt therapy to the patient. Different healthcare-related concerns [14–21] have been described using artificial intelligence, machine learning, and deep learning methods, which support us to make the prediction.

3 Proposed System

A vast number of devices, commodities, and services may now be connected via the Internet of Things (IoT), which also introduces cognitive capabilities to the swarm of these “things”. Im Laufe der letzten zehn Jahre, the Internet of Things (IoT) has infiltrated a wide range of real-world applications. ML methods such as neural networks can't always be relied upon in the context of resource-constrained IoT devices that are intended to be everywhere and available 24/7, despite recent breakthroughs in machine learning (ML). In this paper, we used EfficientNet-B0 a new family of convolutional networks that has faster training speed and better parameter efficiency than previous models. We used EfficientNet-B0 for heart disease prediction.

Machine Learning Classification Method in IoT

In supervised learning, the input dataset's “true” or “right” labels are available. The labelled input dataset are provided for Training (Training Data) as a consequence of the algorithm. Training is the process of creating acceptable predictions based on the input data and then refining those predictions using ground truth until the system achieves the required level of accuracy. It's common to optimize a cost function or a goal function in machine learning approaches. As a result, the cost function is often used as a means of comparing algorithm estimates to the real world. This enables us to improve our model's accuracy by making predictions that are close to the actual values by reducing the cost function (ground truth). The gradient descent approach is commonly used to minimize the cost function. Supervised learning approaches are utilized in a variety of disciplines, including phytoplankton species identification, rainfall-induced landslides are being mapped, and biological data is being classified. In a machine learning method for IoT applications is incorporated into an embedded sensor system. We offer supervised learning algorithms in the following subsections [22].

Linear Regression

Regression is a statistical method for determining how two variables, input and output, are related. A continuous function is created from the input variables. Each of the training sets is labelled with a single independent and dependent variable. It's worded like this

$$h(x) = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad (1)$$

where x_1, x_2, \dots, x_n are the model's characteristics and w_1, w_2, \dots, w_n are its weights.

Multivariate linear regression is utilized in a variety of applications, including activity detection and classification, for BCI data, as well as identification of steady state visual evoked potential (SSVEP).

Logistic Regression

If you're looking to discover a hypothesis function that gives continuous results, a multivariate regression model is the way to do it. In the next section, we'll look into classification, which is basically a supervised learning algorithm whose goal is to provide discrete output. This is a statistical approach for modelling a binary result.

Additionally, logistic regression is used in a variety of disciplines such as evaluating trauma care, grading patient severity, predicting heart disease risk, early detection of glaucoma in ocular thermographs, and computer vision and adaptive object tracking. We can have a one-vs-all implementation for a multiclass classification issue.

Support Vector Machines (SVM)

The SVM is an AI calculation for characterizing data that has mostly been applied to problems involving organization. SVM has excellent arrangement performance, which makes it popular in many different applications. In a two-sided characterization problem, the occurrences are split by a hyperplane $aTc + x = 0$, where a and x are surface-standard dimensions coefficient vectors, x is modified as a motivator from the original starting point stage, and c is an illuminated assortment of respects.

Naïve Bayes

This supervised technique requires a little amount of data for categorization. For learning, probability distributions are utilized. The probability is then updated depending on the most recent information. The use of Bayesian classifiers in IoT and WSN systems is limited by the requirement for prior environmental knowledge.

$$p(w_c|x) = p(w_c|x_1, x_2, x_3 \dots x_n) \quad (2)$$

$p(x)$ is the predictor's prior probability, which serves as the normalization factor. $P(w_c|x)$ is the posterior probability.

The Naive Bayes is a grouping-controlled learning computation. The restrictive likelihood hypothesis determines the class of a separate component vector.

The preparation dataset is used by the NB to determine the restrictive probability value of vectors for a certain class. After processing the likelihood-restrictive estimation of each vector based on its contingency likelihood, the new vectors class is understood. The abbreviation NB is used to describe content-related concerns.

K-Nearest Neighbours

One of the most basic supervised machine learning algorithms is the k-nearest neighbours (KNN) method. KNN may be used to convert discrete input points into discrete results. Because the complexity grows with dimensionality, before utilizing KNN, dimensionality reduction techniques are used to prevent the consequences of the dimensionality curse.

This algorithm's computations are straightforward. However, it delivers incorrect results when dealing with large training sets and high dimensions.

The KNN classifier is used to identify stress using physiological data and to detect epileptic episodes.

Decision Tree Classifier

A monitored AI algorithm is a decision tree. A decision tree shape is basically a tree with each handle standing in for the decision centre or the centre of the leaf.

The decision-making processes for the decision tree are efficient and suitable. Interconnected interior and outer focus focuses made up a decision tree.

Deep Learning in IoT

The problem of evaluating and identifying various facts and information has piqued people's attention. Deep learning has given us a fresh perspective on the world. It is possible to build complex multilayer ANN models using deep learning techniques that rely on large-scale data training to imitate cerebral cortex architecture. Multi-level learning helps us learn multiple layers of abstraction so that future processing may be supported. It is a subset of machine learning and the most widely used ML implementation. Even though artificial intelligence (AI) is a science, the most widely used AI application is based on machine learning.

The artificial neural network operates on a neuron as input, changing some of its internal states, known as activations, in response to the input, as well as providing output that is dependent on input and activations. A directed weighted graph is essentially a network that connects the output of one neuron to the input of other neurons. A learning system, which computes activation and is controlled by a learning algorithm, can improve the weights.

A method to machine learning known as deep learning focuses on the representation of information and the learning process. It's a machine learning approach that mimics the human brain's neural organization. Deep learning was developed as a result of ANN research. For information processing purposes, it extracts human brain neural network, develops a basic model, and constructs numerous networks using various connection methods. To create deep learning, the ANN model, also known as DNN, was utilized in this study. We use a multilayer learning model with hl to describe a deep learning architecture. Deep learning may uncover feature representations dispersed across the data by incorporating low-level characteristics into higher-level traits or categories. Deep learning is capable of learning more complex nonlinear network topologies, function approximations, performing difficult, creating distributed representations of input data, and comprehending basic characteristics of small datasets. Complex functions can be represented with fewer parameters thanks to the use of several layers.

EfficientNet-B0

The models chosen for the study were chosen because of their importance. EfficientNet-B0 has been proposed with the presumption that higher accuracy and efficiency may be obtained by balancing all networks. The figure following illustrates how EfficientNet-B0 outperforms CNN in improving accuracy while considerably lowering the number of parameters.

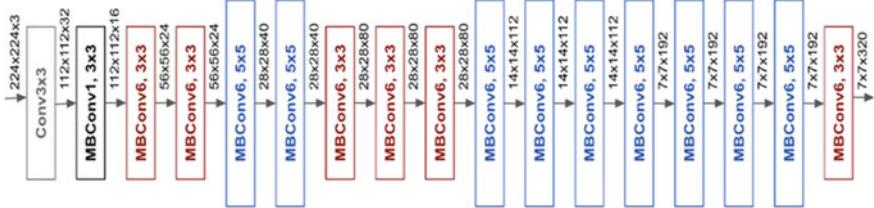


Fig. 4 EfficientNet-B0 architecture

The need of a strong baseline network cannot be overstated because model scaling does not affect the layer operators F_i in the baseline network. We have also constructed a new mobile-size baseline, termed EfficientNet, in order to further show the efficacy of our scaling strategy. We will assess our scaling method using current ConvNets [23].

Our EfficientNet-B0 is somewhat larger than Net since our FLOPS objective is 400 M rather than the more common 200 M. EfficientNet-B0's design is seen in Fig. 4. Squeeze-and-excitation optimization is added to mobile inverted bottleneck MBCConv, which is its primary building element. As model size and training data size increase, training efficiency is crucial for deep learning [23–27].

4 Result and Discussion

In this paper, MATLAB (2018A) was used for experimental evaluation.

Table 1 and Fig. 5 show the sensitivity for heart disease detection. The algorithms we used for the sensitivity for heart disease detection are linear regression, logistic regression, support vector machine (SVM), Naïve Bayes, KNN, decision tree, ANN, and EfficientNet-B0. The EfficientNet-B0 shows high performance than other algorithms.

Table 1 Sensitivity for heart disease detection

Technique name	Sensitivity (%)
Linear regression	84
Logistic regression	85
SVM	91
Naïve Bayes	89
Decision tree	90
KNN	88
ANN	91
EfficientNet-B0	93

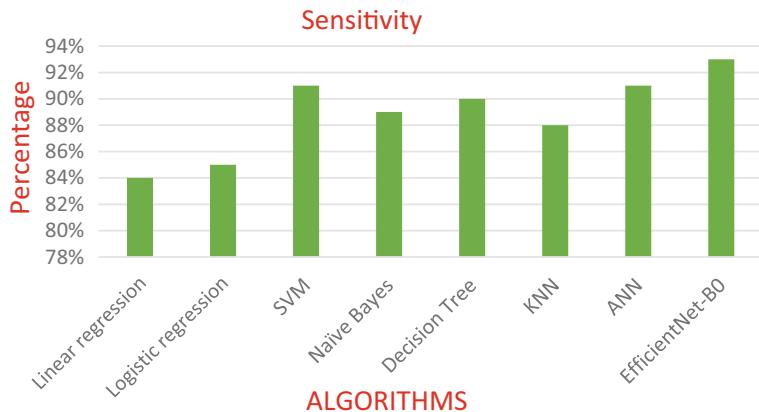


Fig. 5 Sensitivity for heart disease detection

From Table 2 and Fig. 6 it shows accuracy for heart disease detection. The algorithms we used for the sensitivity for heart disease detection are linear regression, logistic regression, support vector machine (SVM), Naïve Bayes, KNN, decision tree, ANN, and EfficientNet-B0. The EfficientNet-B0 shows high accuracy of 92% than other algorithms.

Table 2 Accuracy for heart disease detection

Technique name	Accuracy (%)
Linear regression	83
Logistic regression	84
SVM	90
Naïve Bayes	88
Decision tree	89
KNN	87
ANN	90
EfficientNet-B0	92

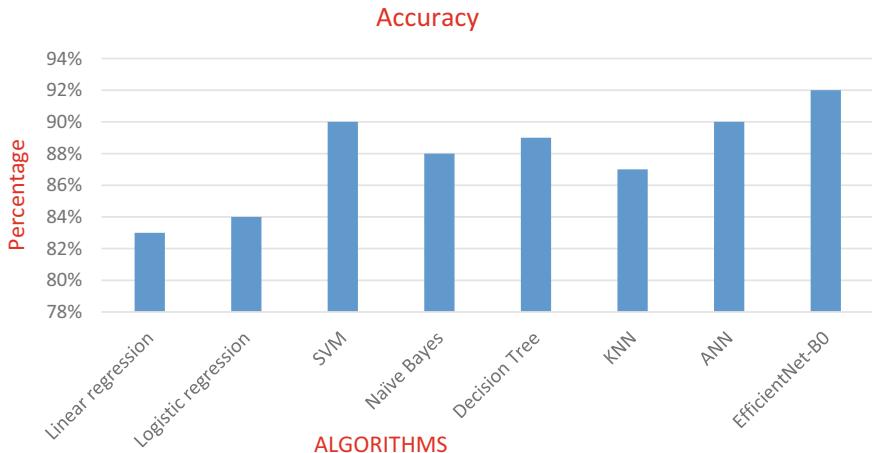


Fig. 6 Accuracy for heart disease detection

5 Conclusion

Based on IoT medical equipment, the heart disease detection system was developed. The creation of easy ways to make money, save time, and conserve energy is the aim of the Internet of Things. Many Internet of Things (IoT)-based systems make use of machine learning or deep learning algorithms to improve performance or development while saving time, money, energy, and other resources. We'll examine several machine learning and deep learning techniques utilized in the Internet of Things' heart disease detection system in this study. Before exploring several learning techniques, including deep learning models, we first give an introduction of machine learning. We employed EfficientNet-B0, a fresh convolutional network that outperforms earlier models in terms of parameter efficiency and training time. We used EfficientNet-B0, a new convolutional network with faster training speed and better parameter efficiency than previous models. The EfficientNet-B0 shows high accuracy of 92% than other algorithms.

References

1. Lu Y (2019) Artificial intelligence: a survey on evolution, models, applications and future trends. *J Manage Anal* 6(1):1–29
2. Lin YJ, Chuang CW, Yen CY, Huang SH, Huang PW, Chen JY, Lee SY (2019) Artificial intelligence of things wearable system for cardiac disease detection. In: 2019 IEEE international conference on artificial intelligence circuits and systems (AICAS), pp 67–70. IEEE
3. Jeong HJ, Lee HJ, Moon SM (2017, October) Work-in-progress: cloud-based machine learning for IoT devices with better privacy. In: 2017 international conference on embedded software (EMSOFT). IEEE, pp 1–2

4. Cai KL, Lin FJ (2018, October) Distributed artificial intelligence enabled by onem2m and fog networking. In: 2018 IEEE conference on standards for communications and networking (CSCN), pp 1–6. IEEE
5. Mamdouh M, Elrukhs MA, Khattab A (2018, August) Securing the internet of things and wireless sensor networks via machine learning: a survey. In: 2018 international conference on computer and applications (ICCA). IEEE, pp 215–218
6. Dridi A, Khedher HI, Moungla H, Afifi H (2020, June) An artificial intelligence approach for time series next generation applications. In: ICC 2020–2020 IEEE international conference on communications (ICC), pp 1–6. IEEE
7. Prutyanov V, Melentev N, Lopatkin D, Menshchikov A, Somov A (2019, June) Developing IoT devices empowered by artificial intelligence: experimental study. In: 2019 Global IoT Summit (GIoTS). IEEE, pp 1–6
8. Uddin MR, Kabir KM, Arefin MT (2019, May) Artificial neural network inducement for enhancement of cloud computing security. In: 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT). IEEE, pp 1–6
9. Patil PS, Dharwadkar NV (2017, February) Analysis of banking data using machine learning. In: 2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, pp 876–881
10. Kansal S, Sikri M, Gupta A, Sharma M (2018, September) A prospect of achieving artificial neural networks through FPGA. In: 2018 international conference on computing, power and communication technologies (GUCON). IEEE, pp 358–363
11. Zeng L, Li E, Zhou Z, Chen X (2019) Boomerang: On-demand cooperative deep neural network inference for edge intelligence on the industrial Internet of Things. *IEEE Network* 33(5):96–103
12. Lassalle P, Ingla J, Michel J, Grizonnet M, Malik J (2015) A scalable tile-based framework for region-merging segmentation. *IEEE Trans Geosci Remote Sens* 53(10):5473–5485
13. Sarmah SS (2020) An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network. *IEEE Access* 8:135784–135797
14. Biswas R, Pal S, Sarkar B, Chakrabarty A (2020) Health-care paradigm and classification in IoT ecosystem using big data analytics: an analytical survey. In: Solanki V, Hoang M, Lu Z, Pattnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing. Springer, Singapore, vol 1125. https://doi.org/10.1007/978-981-15-2780-7_30
15. Jeyalakshmi S, Akila D, Padmapriya D, Suseendran G, Pal S (2021) Human facial expression based video retrieval with query video using EBCOT and MLP. In: Peng SL, Hao RX, Pal S (eds) Proceedings of first international conference on mathematical modeling and computational science. Advances in intelligent systems and computing, vol 1292. Springer, Singapore. https://doi.org/10.1007/978-981-33-4389-4_16
16. Singh D, Sahana S, Pal S, Nath I, Bhattacharyya S (2020) Assessment of the heart disease using soft computing methodology. In: Solanki V, Hoang M, Lu Z, Pattnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing. Springer, Singapore, vol 1125. https://doi.org/10.1007/978-981-15-2780-7_1
17. Chakrabarty A, Tagiya M, Pal S, Cuong NHH (2020) Managing psychosomatic disorders related to obsession and addictions to gadgets through IoT surveillance. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_64
18. Biswas R, Pal S, Cuong NHH, Chakrabarty A (2020) A novel IoT-based approach towards diabetes prediction using big data. In: Solanki V, Hoang M, Lu Z, Pattnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing. Springer, Singapore, vol 1125. https://doi.org/10.1007/978-981-15-2780-7_20
19. Suseendran G, Doss S, Pal S, Dey N, Quang Cuong T (2021) An approach on data visualization and data mining with regression analysis. In: Advances in intelligent systems and computing, pp 649–660. https://doi.org/10.1007/978-981-33-4389-4_59
20. Rakshit P, Nath I, Pal S (2020) Application of IoT in healthcare. In: Peng SL, Pal S, Huang L (eds) Principles of Internet of Things (IoT) ecosystem: insight paradigm. Intelligent systems reference library, vol 174. Springer, Cham. https://doi.org/10.1007/978-3-030-33596-0_10

21. Tagiya M, Sinha S, Pal S, Chakrabarty A (2020) Transformation from HRM inadequacy and bias-syndrome to transparent and integrated ecosystem through IoT-intervention in career management. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_61
22. Shanthamallu US, Spanias A, Tepedelenlioglu C, Stanley M (2017, August) A brief survey of machine learning methods and their sensor and IoT applications. In: 2017 8th international conference on information, intelligence, systems & applications (IISA). IEEE, pp 1–8
23. Tan M, Le Q (2021, July) Efficientnetv2: smaller models and faster training. In: International conference on machine learning. PMLR, pp 10096–10106
24. Tan M, Le Q (2019, May) Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR, pp 6105–6114
25. Kiruthiga R, Akila D (2019) Phishing websites detection using machine learning. Int J Recent Technol Eng 8(2S11):111–114
26. Kanomozhi E, Akila D (2020) An empirical study on machine learning algorithm for plant disease prediction. J Critic Rev 7(5):491–493
27. Kanimozhi E, Akila D (2020) An empirical study on neuroevolutional algorithm based on machine learning for crop yield prediction. In: Peng SL, Son L, Suseendran G, Balaganesh D (eds) Intelligent computing and innovation on data science. Lecture Notes in Networks and Systems, vol 118, pp 109–116
28. Al-Makhadmeh Z, Tolba A (2019) Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach. Measurement 147:106815

Deep Learning in Distance Awareness Using Deep Learning Method



Raghad I. Hussein and Ameer N. Onaizah

Abstract Recent studies have shown that deep learning does pretty well at reproducing 3D scenes using multiple-view images or videos. Nevertheless, these restorations do not expose the personalities of the items, and item identification is necessary for such a scene to work in augmented worlds or interactive features. The objects in a picture that have been reconstructed as a unified mesh are handled as a singular body rather than being seen as independent categories that can be engaged with or changed. Reconstructing an entity three-dimensional image from a two-dimensional image is challenging since the transformation from a visual scene to a picture is permanent and reduces a dimensionality. In addition to creating more exact shapes when compared with previous methodologies for mesh rebuilding from individual images, our approach exhibited improved achievement in initiating comprehensive meshes when compared to strategies using only inherent portrayal mesh rebuilding networks (for instance, neighborhood deep implicit functions). By applying the multi-modal instructional methods, this was done. Real-world facts were utilized to assess the effectiveness of the suggested method. The results showed that it might perform noticeably better than earlier methods for entity 3D scene rebuilding.

Keywords DNN · Deep learning · Network system · AI

R. I. Hussein

Faculty of Pharmacy, University of Kufa, Najaf, Iraq

e-mail: Ragadi.alfarhani@uokufa.edu.iq

A. N. Onaizah (✉)

School of Automation, Beijing Institute of Technology, Zhongguancun, Beijing 100811, China
e-mail: ameern.unaiza@uokufa.edu.iq

Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

1 Introduction

Three-dimensional restoration from pictures, a lengthy challenge in computer vision, offers a variety of uses. For example, the recreated mesh may be utilized in virtual reality (VR) activities and gameplay [1]. Three-dimensional restoration using 3D images, such as photogrammetry or examination, is often more precise than reconstructing using 2D photographs because of 3D data provides more details about the framework and topology of the object. Nevertheless, obtaining 3D data typically requires specialized tools, such as depth or LiDAR cameras, that may be more costly or challenging to use compared to 2D cameras [2, 3]. Image representation, like RGB photos, is easier to acquire, but because there isn't any 3D information, the reconstruction procedure can be more challenging. As a result of the advancement of deep learning, studies on 3D object identification, target detection, object reconstructing, and scene rebuilding have recently been conducted [2].

Among these, various image 3D scene reconstructing algorithms have had positive outcomes. The accuracy of the rebuild, however, is insufficient if there is just one 2D picture accessible because of the scaling uncertainty introduced by the three-dimensional to two-dimensional projection's drop in dimension [4–6]. However, the majority of studies to date have integrated rather than divided all of the elements of a scene it in to a single geometry. Because they lack object identification and interactivity, these reconstructed scenes are not suitable for VR apps or videogames [6, 7]. Object-aware correcting errors has indeed been taken into account when reconstructing an interactive 3D scene.

Object-aware reshaping a 3D image from either a 2D image necessitates calculating the 3D model and stance for every item in the picture, the 3D structure boundary box of both the picture, and the cameras attitude used to capture the image as input. The goal is to faithfully reproduce the actual circumstances that can be observed in the supplied photo as a 3D image. Yet, the production of 2D images is irreparable because several 3D sequences can be joined to produce a single 2D image. Lacking depth information, reliable position inference from such a single 2D picture is challenging. Certain technical and tactical the 3D poses for objects by taking advantage of the limitations between items in a scene. For instance, objects shouldn't overlap in a three-dimensional environment. Even if these restrictions can be added to the wavelet coefficients during training to help identify appropriate relative postures to objects, the results remain much better [8].

A 2D image frequently only shows a small portion of a 3D object, which often has concealed edges and barriers which are not apparent in the image. This presents another challenge. It may be difficult to exactly reproduce an object's complete 3D objects from a single depth snapshot due to the loss of important information. Many studies have been done to try to solve this problem. In the early days of computer vision, a lot of experiments tried to recreate three-dimensional (3D) objects employing point cloud data or voxel representations, which are capable of capturing

the main form but may be less accurate at reconstructing surface features. The resolution issue has recently been overcome using implicit mechanism approaches. Yet, on occasion they lead to fractured meshes.

Deep learning, a subset of machine learning, has emerged as a powerful tool for various applications, including distance awareness. Distance awareness is an essential capability for many tasks, such as autonomous driving, robotics, and surveillance systems, where understanding the spatial relationships between objects in the environment is crucial. Deep learning in distance awareness involves leveraging neural networks with multiple hidden layers to automatically learn hierarchical representations from data, enabling them to infer distances or spatial relationships between objects. These neural networks are capable of extracting meaningful features from raw sensor data, such as images, point clouds, or other sensor inputs, and learning complex patterns and relationships in the data without explicit programming. Deep learning methods for distance awareness can be categorized into several types, including:

Object detection and tracking: Deep learning-based object detection and tracking methods use convolutional neural networks (CNNs) to detect and track objects in images or videos. These methods can estimate the distances to objects based on the detected object's size, location, and other contextual information, enabling distance awareness.

Depth estimation: Deep learning-based depth estimation methods leverage CNNs to predict the depth or distance of objects from a single image or a pair of stereo images. These methods can estimate the distance to objects in the scene, which is essential for tasks such as obstacle avoidance or scene understanding.

LiDAR-based distance awareness: Light Detection and Ranging (LiDAR) sensors are commonly used in autonomous vehicles and robotics for distance sensing. Deep learning-based methods can process the point cloud data generated by LiDAR sensors, and learn to estimate distances, detect objects, and perform other distance-aware tasks using techniques such as point cloud segmentation, point cloud convolutional networks (PC-CNNs), or graph neural networks (GNNs).

Sensor fusion: Deep learning can also be used to fuse information from multiple sensors, such as cameras, LiDAR, radar, and other sensors, to achieve more accurate distance awareness. These methods typically involve building multi-modal deep neural networks that can effectively integrate information from different sensors and leverage the strengths of each sensor for distance estimation.

Deep learning-based methods for distance awareness have shown significant advancements in recent years, achieving state-of-the-art performance in various applications. They have the potential to greatly enhance the capabilities of autonomous systems, robotics, and other domains where understanding distances and spatial relationships is critical for decision making and situational awareness. In this study, a method for deriving camera, layout cluster centers, boundary boxes for three-dimensional objects, including 3D object shapes from a picture is presented. Compared to past techniques that directly calculated this data from a 2D image,

the recommended method improves the precision of 3D recognition by first calculating the depth to retrieve the scale information. Following the development of a mesh reconstructing network that generates a network in a pair of distinct forms, we developed a multitask learning technique.

2 Related Study

For object-aware three-dimensional rebuilding, predicting the structure frame, 3D bounding boxes enclosures, and object morphologies is required. Early research focused on layout prediction and 3D item input vector estimate instead of predicting 3D object shapes. They calculated the layout coordinates exclusively using the edge, length, and shape features. Nowadays, the topic of predicting 3D shape coordinates has become more popular. Several limitations were used in these tests to manage the prediction of 3D object thresholding. For example, in an air-conditioned room, the objects are solid [9].

Since they must align with the floor surfaces and walls when three-dimensional features are reflected onto an image, the projecting 2D parameters must coincide with the genuine 2D neighboring pixels [10]. When individuals appear, their activities with the things in the picture serve to clarify their locations. The aforementioned methods do not restore the object shapes; rather, the reconstructed objects are only expressed by 3D thresholding.

“Monocular Depth Estimation with Transformers” by Mahjourian et al. (2019): This study proposes a deep learning-based method for monocular depth estimation using transformer networks, which are originally designed for natural language processing tasks. The authors demonstrate that transformer-based models can effectively capture global contextual information in images, enabling accurate depth estimation without the need for stereo images or additional sensors. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation” by Qi et al. (2017): This study introduces PointNet, a deep learning architecture for processing 3D point cloud data. PointNet can be used for point cloud segmentation, which is essential for distance awareness in LiDAR-based applications. The authors demonstrate that PointNet can effectively learn spatial features from point clouds, enabling accurate object detection and segmentation in 3D scenes.

“FusionNet: 3D Object Classification Using Multiple Data Representations” by Eitel et al. (2015): This study proposes FusionNet, a deep learning-based approach for object classification in 3D point clouds using multiple data representations. The authors combine color images, depth images, and 3D point clouds to train a multi-modal neural network that can effectively leverage the strengths of different sensor modalities for accurate object classification and distance awareness. “Deep Tracking: Seeing Beyond Seeing Using Recurrent Neural Networks” by Held et al. (2016): This study presents a deep learning-based method for visual object tracking using recurrent neural networks (RNNs). The authors demonstrate that RNNs can learn to track objects in video sequences by implicitly estimating their distances from the

camera, enabling distance-aware tracking without explicit depth information. “Deep Stereo: Learning to Predict Depth from Stereo Images” by Mayer et al. (2016): This study proposes a deep learning-based approach for stereo depth estimation using CNNs. The authors demonstrate that CNNs can learn to estimate accurate depth maps from stereo image pairs, enabling distance awareness in autonomous driving and robotics applications.

“Deep Multi-Sensor Fusion for Object Detection in Autonomous Vehicles” by Chen et al. (2017): This study presents a deep multi-sensor fusion approach for object detection in autonomous vehicles. The authors propose a deep neural network architecture that fuses information from cameras, LiDAR, and radar sensors to improve object detection accuracy and distance estimation. The study demonstrates the effectiveness of multi-sensor fusion for distance awareness in challenging driving scenarios. “LiDARNet: A Deep Autoencoder Approach to LiDAR Point Cloud Compression for Privacy-preserving Remote Sensing” by Ma et al. (2020): This study introduces LiDARNet, a deep autoencoder-based approach for LiDAR point cloud compression. The authors propose a novel architecture that uses deep autoencoders to compress point cloud data while preserving distance information. The compressed representations can be used for distance-aware applications such as object detection and tracking in remote sensing scenarios.

“Monocular Depth Perception with Hierarchical Multi-Scale Predictive Networks” by Liu et al. (2018): This study presents a hierarchical multi-scale predictive network for monocular depth perception. The authors propose a deep learning architecture that learns to predict depth maps at multiple scales, which enables accurate distance estimation in monocular images. The study demonstrates the effectiveness of the hierarchical approach for improving depth perception and distance awareness in single-camera setups. “Beyond RGB: Depth Sensing with Multi-Modal Deep Networks for Object Detection and Tracking in Outdoor Environments” by Gómez-Donoso et al. (2019): This study introduces a multi-modal deep learning approach for object detection and tracking in outdoor environments. The authors combine RGB images with depth maps and thermal infrared data to train a deep neural network that can effectively leverage multi-modal information for distance-aware object detection and tracking in challenging outdoor conditions.

“Depth Sensing Beyond LiDAR Range: A Deep Learning Approach for Robust Long-Range Depth Estimation” by Zhou et al. (2020): This study proposes a deep learning approach for robust long-range depth estimation beyond the typical range of LiDAR sensors. The authors leverage a deep neural network to estimate depth maps from RGB images, which can provide distance awareness beyond the limited range of LiDAR sensors. The study demonstrates the potential of deep learning for long-range depth estimation in applications such as autonomous driving and robotics. Several methods for form retrieval have resulted in 3D landscapes with looking things. The closest desktop design (CAD) version toward the image’s subject is selected from a CAD library using the proximity in the word embedding between the input image and the computer-aided model [12–14] despite the fact that these methods make it possible to create flawless 3D sceneries, the breadth and volume of the CAD data have a significant impact on the reconstruction’s performance [15–17].

3 Proposed Methodology

Figure 1 shows an overview of the suggested object-aware three-dimensional reconstruction method. The estimation process includes estimating the camera position, 3D layout coordinates, 3D item cluster centers, and 3D object morphologies. The input photos are first single-handedly to provide object suggestions using a speedier region-based two dimensional detector using convolutional neural networks (Faster R-CNN). The three-dimensional restoration is then split into two halves. The first stage of estimate also includes mesh rebuilding networks and 3D shape detection. We modify the layout estimate and 3D object recognition networks from that in addition to proposing a novel mesh rebuilding network on multi-modal learning. The second step is the improved estimation stage, which is made up of a depth-feature producing network and a scene modeling technique.

Data Collection and Preparation: The first step is to collect and prepare the data required for the deep learning model. This may include RGB images, depth maps, point cloud data, or other relevant sensor data depending on the specific application. The data may need to be pre-processed, cleaned, and annotated with ground truth distance labels or other relevant information.

Model Architecture Selection: Next, a suitable deep learning architecture needs to be selected based on the specific task and data. This may involve choosing from popular architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformer networks, or other specialized architectures for point clouds or sensor data.

Model Training: The selected deep learning model is then trained using the prepared data. This typically involves feeding the input data into the model, optimizing the

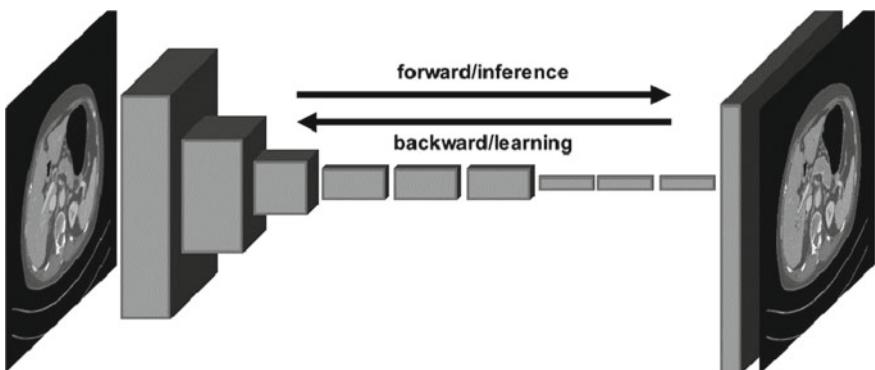


Fig. 1 Proposed model

model's parameters using a suitable loss function, and updating the model iteratively using an optimization algorithm. The training process may require fine-tuning, hyperparameter tuning, and validation to achieve the desired accuracy and performance.

Model Evaluation: Once the model is trained, it needs to be evaluated using appropriate evaluation metrics to assess its performance. This may involve measuring accuracy, precision, recall, F1 score, or other relevant metrics depending on the specific task. The model may be evaluated on a separate test dataset or through cross-validation to ensure its generalization ability.

Model Optimization: Based on the evaluation results, the model may need to be further optimized. This may involve adjusting hyperparameters, refining the model architecture, or using techniques such as regularization, dropout, or batch normalization to improve model performance.

Distance Awareness Application: Once the trained and optimized model is validated, it can be used for distance awareness applications, such as object detection, tracking, or perception. The model can take input from sensors, process the data using the learned representations, and provide accurate distance information for decision making or other relevant tasks.

Model Deployment: Finally, the trained and optimized model can be deployed in the target environment, such as an autonomous vehicle, robotics system, or any other application where distance awareness is required. Deployment may involve integrating the model into a larger system, optimizing for real-time performance, or ensuring robustness and reliability in the target environment.

Given an input data point x , the goal is to learn a function $f(x)$ that maps x to a continuous distance value y , i.e., $y = f(x)$.

The deep learning model is trained to minimize the discrepancy between the predicted distance value and the ground truth distance value during the training process using a suitable loss function, such as mean squared error (MSE) or mean absolute error (MAE).

MSE is defined as the average of the squared differences between predicted values (denoted as y_{hat}) and ground truth values (denoted as y) for a set of data points:

$$\text{MSE} = (1/n) * \sum (y_{\text{hat}} - y)^2,$$

where n is the number of data points in the dataset and the summation (Σ) is taken over all data points. MSE is a measure of the average squared error between predicted and ground truth values, where larger errors contribute more to the overall loss due to the squaring operation. It is a non-negative value, with lower values indicating better performance, as it represents smaller errors between predicted and ground truth values.

MAE is defined as the average of the absolute differences between predicted values (denoted as y_{hat}) and ground truth values (denoted as y) for a set of data points:

$$\text{MAE} = (1/n) * \Sigma |y_{\text{hat}} - y|,$$

where n is the number of data points in the dataset and the summation (Σ) is taken over all data points. MAE is a measure of the average absolute error between predicted and ground truth values, where the absolute differences are used, meaning that negative and positive errors are treated equally. It is also a non-negative value, with lower values indicating better performance, as it represents smaller errors between predicted and ground truth values [11].

4 Result and Discussion

Using Media, the clipped object is encrypted into a visual feature matrix. The shape data for the target item is again included in the shape embedding that is encoded after being coupled to a one-hot decoded object class. Two different shape video codecs were used. One of these had the same design that was being used and was an encoder for Locally Deep Integral Functions (LDIFs). Given initial shape encapsulation and 3D point in a standard space, it assesses whether 3D point exists within the target three-dimensional mesh. If 3D points get intensively sampled there too and assessed using the decoding during the reasoning period, the classical space transforms into an occupied field. The occupation field can then be processed using the walking cube method to create a trapezoidal mesh. Unfortunately, the local profound implicit typically reported occasionally results in fragmented meshes because of poor global shape control. We re-solved this problem by estimating the overall form of the target item using a second statement cloud shape encoder. The bitmap image decoder utilized by the design is described. The various parameters are shown in Fig. 2.

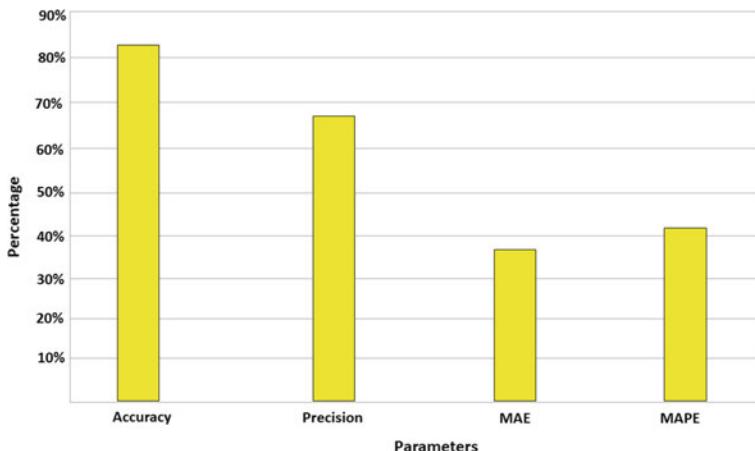


Fig. 2 Analysis of different parameters

We used the supervised learning in implicit 3D to obtain the reconstructions outcomes and compared them to all of those obtained using the proposed technique.

The visualization outcomes after testing just on SUN RGB-D data are displayed and analyzed. The first row displays the input photos, while the second row shows the visualization results with the test dataset arrangement and object cluster centers. The third and fourth rows, correspondingly, show the outcomes of the suggested method for 3D recognition and layout predictions in addition to the pre-trained models implicit3D models. The final two rows show the outcomes of the suggested method and implicit 3D's entity 3D scene rebuilding (Fig. 3).

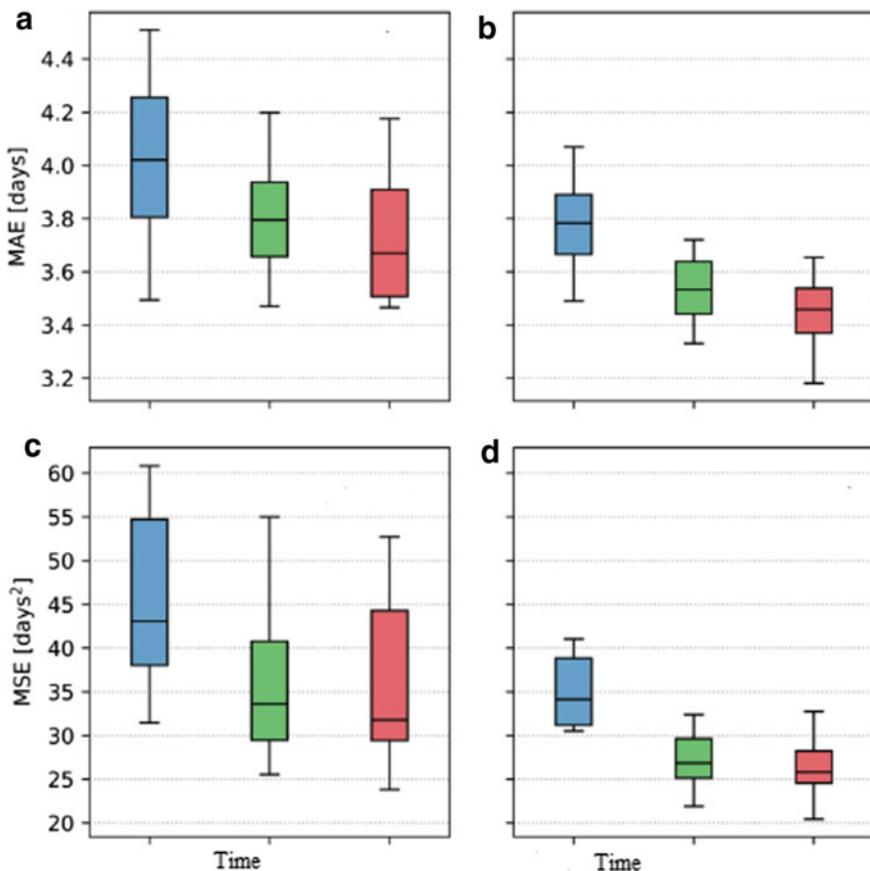


Fig. 3 The error MSE and MAE with respect to time

5 Conclusion

In this research, we suggested a system that concurrently evaluates the sensor attitude, 3D design, 3D image pose, and entity shapes while rebuilding 3D scenes. The system is aware of things and can recognize items in the scene. In the improved estimating step, we advise utilizing a thorough producing net to address the depth confusion issue in 2D to 3D understanding. We also propose to mesh restoration network utilizing asynchronous learning to obtain more complete meshes. We put the suggested approach to the test using real datasets in MATLAB and compared the results to those obtained using cutting-edge methods. Our analysis on the MATLAB dataset showed that this method improved 3D object boundary box estimates for the plurality of item categories. The mesh reconstructing quality of the employed dataset demonstrated the potential of the multifunction having to learn net reconstructing networks for precise form forecasting. A drawback to this research is that the lattice restoration networks utilized can only predict item forms seen in the designated item categories. The recommended method's applicability to more complex and varied object reconstruction issues will be expanded through future study.

References

1. Mohammed AM, Haytamy SSA, Omara FA (2023) Location-aware deep learning-based framework for optimizing cloud consumer quality of service-based service composition. *Int J Electr Comput Eng* (2088–8708) 13(1):638
2. Zhu S et al (2023) Displacement-sensible imaging through unknown scattering media via physics-aware learning. *Opt Lasers Eng* 160:107292
3. Li C et al (2023) An AR-assisted deep reinforcement learning-based approach towards mutual-cognitive safe human-robot interaction. *Robot Comput Integr Manuf* 80:102471
4. Chen D et al (2023) Position-aware and structure embedding networks for deep graph matching. *Pattern Recogn* 136:109242
5. Li X et al (2023) Rare disease classification via difficulty-aware meta learning. In: *Meta-learning with medical imaging and health informatics applications*. Academic Press, pp 331–347
6. Le N et al (2023) Uncertainty-aware label distribution learning for facial expression recognition. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*
7. Zhang W et al (2023) ACCPG-Net: a skin lesion segmentation network with adaptive channel-context-aware pyramid attention and global feature fusion. *Comput Biol Med* 154:106580
8. Tang X, Li R, Fu CW (2023) CAFI-AR: contact-aware freehand interaction with AR objects. *Proc ACM Interact Mob Wearable Ubiquit Technol* 6(4):1–23
9. Tang L, Lin H-N, Liu L (2023) Deep learning method for testing the cosmic distance duality relation. *Chin Phys C* 47(1):015101
10. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Lulwah M, Alkwai SK (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electr Eng* 107:108617, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.2023.108617>
11. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerg Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>

12. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electr Eng* 108:108715. ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
13. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun* ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
14. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM (2023) An efficient hybrid deep learning model for denial of service detection in cyber physical systems. *IEEE Trans Netw Sci Eng*. <https://doi.org/10.1109/TNSE.2023.3273301>
15. Gupta U, Sharma R (2023) Analysis of criminal spatial events in India using exploratory data analysis and regression. *Comput Electr Eng* 109(Part A):108761, ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108761>
16. Goyal B et al. (2023) Detection of fake accounts on social media using multimodal data with deep learning. *IEEE Trans Comput Soc Syst* <https://doi.org/10.1109/TCSS.2023.3296837>
17. Sneha PM, Sharma R, Ghosh U, Alnumay WS (2023) Internet of Things and long-range antenna's; challenges, solutions and comparison in next generation systems. In: Microprocessors and microsystems, p 104934, ISSN 0141-9331. <https://doi.org/10.1016/j.micpro.2023.104934>

Analysis of Improving Sales Process Efficiency with Salesforce Industries CPQ in CRM



Pritesh Pathak, Souvik Pal, Saikat Maity, S. Jeyalaksshmi,
Saurabh Adhikari, and D. Akila

Abstract The successful execution of Configure, Price, Quote (CPQ) protocols is of utmost importance for businesses that handle intricate product portfolios within their sales processes. The objective of these approaches is to optimise the generation of competitive quotations by leveraging data from diverse business systems, leading to a decrease in processing time and enhancement of operational efficiency. Salesforce's Industries CPQ for Communications Cloud, previously referred to as Vlocity CPQ, assumes a crucial function within the Salesforce platform by facilitating the automation of the quotation creation process for sales teams. Designed specifically for the communications industry, this solution enables communications service providers (CSPs) to effectively provide a wide array of products and services to their customers in a streamlined manner. By implementing Industries CPQ, Communication Service Providers (CSPs) have the potential to save expenses associated with customisation and maintenance, while also accelerating their time-to-market. In order to assess the

P. Pathak

Liverpool Business School, Liverpool John Moores University, Liverpool, UK

S. Pal (✉)

Department of Management Information Systems, Saveetha College of Liberal Arts And Sciences, Saveetha Institute of Medical And Technical Sciences, Chennai, India

e-mail: souvikpal22@gmail.com

S. Pal · S. Maity

Department of Computer Science and Engineering, Sister Nivedita University (Techno India Group), Kolkata, India

e-mail: saikat.m@snuiv.ac.in

S. Jeyalaksshmi

Department of Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

S. Adhikari

School of Engineering, Swami Vivekananda University, Kolkata, India

e-mail: saurabhadhikari@svu.ac.in

D. Akila

Department of Computer Applications, Saveetha College of Liberal Arts And Sciences, Saveetha Institute of Medical And Technical Sciences, Chennai, India

practical ramifications of Salesforce Industries CPQ, the researcher intends to deploy the aforementioned system within ABC Insurance Company. The objective of the study is to evaluate the program's influence on sales efficiency and profitability by conducting interviews with a targeted cohort of employees who have employed the software in their sales endeavours. Furthermore, the study will include hypothesis testing to analyse the timing of quotation production across various policy configurations utilising Salesforce Industries CPQ, aiming to provide further validation of its benefits.

Keywords CPQ · Salesforce · Industries · CRM · Cloud · Sales · Vlocity

1 Introduction

Multinational corporations encounter difficulties in dealing with a wide range of worldwide clientele [1]. CPQ systems are employed to automate sales processes, facilitate product configuration, create precise quotations, and deliver tailored purchasing experiences [2]. The implementation of this system optimises the process of order management, increases the effectiveness of sales operations, and boosts overall customer happiness. Cloud-based Configure, Price, Quote (CPQ) solutions have gained significant popularity due to their ability to enhance and optimise sales operations within a highly competitive market.

Configure, Price, Quote (CPQ) systems have emerged as a crucial instrument for enterprises seeking to optimise their sales procedures and enhance operational efficiency [3]. There exist several primary justifications for the necessity of CPQ systems:

- Complex product configurations.
- Diverse pricing structures.
- Streamlined sales processes.
- Incorporation with additional business systems.

According to a survey by Gartner in 2021, prominent contenders in the Customer Relationship Management (CRM) sector encompass Salesforce, Microsoft Dynamics, Oracle, and Hubspot. Among them, Salesforce currently commands the greatest market share, accounting for 19.5% of the industry [4]. Cloud-based Customer Relationship Management (CRM) systems have emerged as a significant catalyst for business expansion, owing to its inherent advantages such as reduced initial expenditures, enhanced scalability, and ubiquitous accessibility facilitated by internet connectivity [5]. As organisations persist in placing customer experience at the forefront of their priorities and endeavour to establish distinctive characteristics, it is anticipated that the Customer Relationship Management (CRM) market will witness sustained growth in the foreseeable future.

CPQ, an acronym for Configure, Price, Quote, refers to a Customer Relationship Management (CRM) software product designed to streamline and automate

the process of generating precise and comprehensive sales quotes for sales teams. The integration of this software with Customer Relationship Management (CRM) systems facilitates the streamlining of intricate processes related to product configuration and pricing [3]. CPQ software has been found to have a positive impact on customer experience by improving sales performance and customer happiness through reducing errors and enabling real-time pricing modifications [6].

Salesforce is a cloud-hosted Customer Relationship Management (CRM) software that facilitates the effective management of sales, marketing, and customer support operations for enterprises. Salesforce Industries CPQ is a software solution for the Salesforce platform that facilitates the processes of configuration, pricing, and quoting. This particular solution was bought from Vlocity in the year 2020 [7]. The software enhances sales procedures by enabling teams to customise products, produce quotations, and deliver precise pricing, rendering it a favoured option among Salesforce CRM customers in search of sophisticated functionalities and sector-specific capabilities.

The primary objective of this study is to investigate tactics aimed at improving sales and product management protocols within the context of ABC Insurance. The main aim of this study is to evaluate the effectiveness of the Salesforce Industries CPQ system in enhancing many essential functions within B2B firms comprehensively. In addition to conducting an analysis of the potential advantages offered by the Salesforce Industries CPQ system, it is essential to undertake a thorough assessment of the current protocols implemented at ABC Insurance and consider alternative strategies for enhancing their efficacy. The Salesforce Industries Configure, Price, Quote (CPQ) system is a software solution designed to optimise the effectiveness of product or service configuration, pricing, and quoting procedures within organisations. The investigation on the utilisation of Salesforce Industries CPQ holds potential for various future applications, encompassing:

- Recognising patterns and optimal methodologies.
- Pinpointing possible areas for enhancement.
- Guiding the creation of innovative solutions.
- Offering a foundation for evaluation.

Some key inquiries explored in this study include:

- How does the Salesforce Industries CPQ system influence the sales process, and what advantages does it bring to the organisation?
- What rationale supports allocating substantial resources to the Salesforce Industries CPQ system?
- What are the critical elements for effectively incorporating a Salesforce Industries CPQ system within an organisation?

2 Literature Review

CPQ solutions are utilised to effectively display all pertinent permutations of product outcomes together with precise pricing, necessitating a reduced number of inputs based on chosen product selections. The purpose of this system is to effectively manage the given product models [8], so improving the quality of the sales process and reducing processing difficulties. The relevance of this resides in its ability to streamline the process of product customisation and price complexity, hence contributing to the enhancement of product quality [9]. Gill and Mathur [10] assert that the utilisation of Configure, Price, Quote (CPQ) systems has the potential to expedite the sales cycle. This is achieved by the identification and prioritisation of Key Performance Indicators (KPIs), such as the lead time for sales quotations. The ability to accurately estimate the duration of the sales quotation process is deemed significant in this context.

Furthermore, it is worth noting that CPQ plays a crucial role in aiding sales managers in streamlining intricate sales quote procedures, which sometimes pose difficulties for both sales people and managers [10]. For example, in cases when the approval for a discount is over a predetermined % level, the Configure, Price, Quote (CPQ) tool promptly alerts an accountable individual. This notification enables them to establish limitations for discounts by utilising a pre-established discounting methodology [11]. The implementation of automation in this context obviates the necessity for labour-intensive manual tasks, resulting in reduced waiting periods and, subsequently, a decrease in the overall duration of the sales lead process.

There exists a wealth of data pertaining to the improved efficacy in sales resulting from the adoption and use of Configure, Price, Quote (CPQ) systems. Nevertheless, the specific outcomes are subject to variation depending on factors such as the firm, industry, and use case. There are several cases that provide evidence of the purported enhancements in efficiency that have resulted from the implementation of Configure, Price, Quote (CPQ) systems.

- According to a study conducted by the Aberdeen Group, organisations that have implemented Configure, Price, Quote (CPQ) systems have observed a significant improvement in sales productivity, with a 30% increase, as well as a reduction in the time required for generating quotes, by 25% [11].
- According to a survey conducted by Forrester Research, companies leveraging CPQ systems have observed an average 42% acceleration in the quote creation process and a 31% improvement in quote accuracy [12].
- According to a study conducted by Salesforce, organisations who have implemented CPQ systems have had a significant improvement in both the speed of quote creation, with a 30% increase, and the accuracy of quotes, with a 20% enhancement [13].

It is important to acknowledge that the aforementioned instances serve as a limited representation of the available data, and the outcomes may differ based on the unique circumstances of the organisation, such as its particular use case, scale, and the extent

to which employees embrace it. Several Cloud-based Configure, Price, Quote (CPQ) Systems are currently available within Customer Relationship Management (CRM) platforms, such as Salesforce CPQ [14–18], Oracle CPQ [19], SAP CPQ [20], Conga CPQ [21], and PROS Smart CPQ [22]. Various analytical considerations [23–30] have been elucidated by the utilisation of artificial intelligence, machine learning, and deep learning techniques. These methodologies aid in enhancing the efficiency of the sales process. The case studies have been utilised to conduct analysis and enhance comprehension of the process.

3 Research Methodology

The primary objective of this research study is to investigate the research concerns outlined in Chap. 1 and deepen comprehension of Salesforce Industries Configure, Price, Quote (CPQ) systems. This study employs qualitative research methods, specifically conducting interviews with stakeholders of the CPQ project and analysing user experiences. Qualitative research presents significant contextualisation and interpretation, offering unique insights that cannot be acquired using quantitative methodologies. The study primarily uses data acquired internally from ABC Insurance and employs rigorous academic research procedures, thereby providing as a valuable case study for other organisations. Nevertheless, it should be noted that the findings are limited in scope to ABC Insurance and cannot be extrapolated to a broader population. The research utilised data aggregation and analytic techniques.

4 Research Results and Analysis

In this section, the researcher evaluates the results obtained from the analysis of data, comparing the data collected prior to the adoption of Salesforce Industries CPQ with the data obtained after its deployment. Furthermore, the researcher intends to conduct a hypothesis test on the duration of quote creation processes in order to demonstrate the improvements in sales efficiency and productivity that arise from the deployment of CPQ.

4.1 Interview Results

The study employed a methodology that included structured interviews with pre-established questions, complemented by individual discussions. Based on the findings from the interviews, it has been determined that the integration of the Salesforce Industries CPQ system is necessary for ABC Insurance firm. This decision

is driven by identified deficiencies in the existing sales process, which encompass issues such as inaccuracies, delays, and elevated expenses. The respondents conveyed their endorsement of the novel approach, highlighting advantages such as enhanced documentation, heightened professionalism, less manual labour, and precise configuration, pricing, and quoting of intricate products. Various concerns were expressed regarding maintenance, inflexibility, training, and changes in organisational culture. Overall, the individuals who were interviewed had the belief that the benefits associated with the implementation of the Salesforce Industries CPQ system justified the financial commitment. Nevertheless, the stakeholders voiced apprehensions regarding the inadequate involvement of many constituencies throughout the project's planning and development phases.

4.2 Process Flows

Current Process in ABC Insurance Company. The insurance sales process entails the transmission of consumer data through electronic mail to the sales representative, who subsequently tailors a pre-existing plan to align with state rules. The inclusion of various quotations and underwriting reviews introduces a level of complexity and the possibility of mistakes, which can result in increased stress and delays for the sales and underwriting teams. This phenomenon has the potential to result in customer discontent, competitive disadvantages, and legal liabilities (Fig. 1).

Process Flow after Implementation of Salesforce Industries CPQ. Through the utilisation of Salesforce communities, brokers and sales agents are able to effectively cooperate, hence minimising the necessity for several email exchanges. Salesforce Industries CPQ facilitates the automation of many tasks within the sales process. This includes the generation of opportunities, the linking of census data, and the automatic identification of suitable insurance plans. These actions are performed based on predetermined rules, streamlining the overall workflow. The initiation of the price process and plan approvals are prompted by manual modifications carried

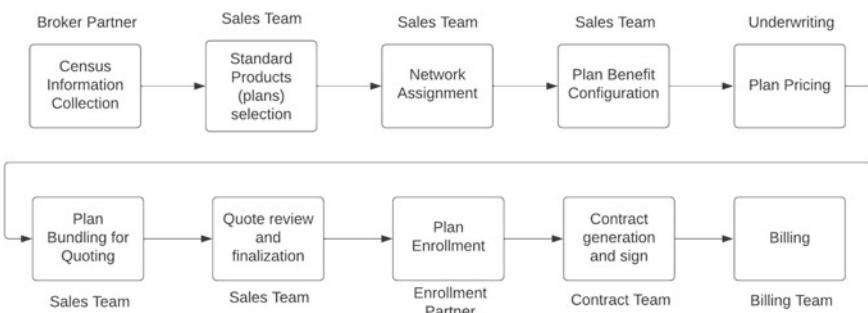


Fig. 1 Process flow before Salesforce Industries CPQ Implementation

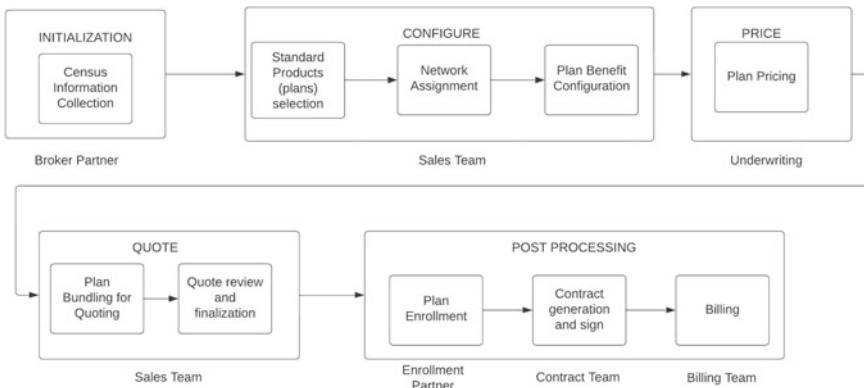


Fig. 2 Process flow after Salesforce Industries CPQ Implementation

out by sales representatives, resulting in the automated generation of quotations. Salesforce enables the facilitation of collaborative efforts between sales agents and brokers in the process of quote evaluation and finalisation. Upon completion of the quotation process, the system initiates the enrolment procedure, generates the necessary contractual documentation, and facilitates the capture of the customer's electronic signature. When it comes to renewals, Salesforce Industries CPQ has the capability to generate opportunities and import relevant contract information, all of which is dependent on the configuration set by the CPQ administrator (Fig. 2).

4.3 Hypothesis Test for Quote Generation

The researcher employed a systematic approach to select ten insurance combinations at random from ABC Insurance. Subsequently, the researcher generated quotes for each of these configurations using Salesforce Industries CPQ. The duration for each setup is shown in Table 1.

Performing Hypothesis Test (T-Test): Conducting a t-test is essential due to the sample size being less than 30 and the unknown population standard deviation. The hypotheses are as follows:

- Null hypothesis (H_0): The quote generation time is 15 s or less.
- Alternative hypothesis (H_1): The quote generation time exceeds 15 s.

The significance level (α) is set at 0.05. Sample statistics are calculated:

- Sample mean (\bar{x}): 16.3 s.
- Sample standard deviation (s): 6.57 s.
- n = number of strategies = 10.

The test statistic (t -value) is determined using the formula:

Table 1 Time taken to generate quotes across various strategy configurations

Strategy Name	Time taken to generate a quote (measured in seconds)
Strategy A	9.2
Strategy B	23.5
Strategy C	14.8
Strategy D	17.1
Strategy E	12.7
Strategy F	7.4
Strategy G	19.9
Strategy H	27.6
Strategy I	4.8
Strategy J	10.6

$$t = (\bar{x} - \mu) / (s / \sqrt{n}).$$

In the given context, \bar{x} represents the sample mean, μ denotes the hypothesised population mean, s signifies the sample standard deviation, and n represents the sample size.

Substituting values:

$$t = 0.971.$$

Degrees of freedom (df) are computed as $(n - 1)$, where (n) is the sample size:

$$df = 10 - 1 = 9.$$

The critical value for a two-tailed test with $df = 9$ and $\alpha = 0.05$ is ± 2.262 . Comparing the absolute value of the test statistic to the critical value:

$$|0.971| < 2.262.$$

Thus, the null hypothesis is not rejected. There is insufficient evidence to conclude that the quote generation time by Salesforce Industries CPQ exceeds 15 s at the 0.05 significance level.

Types of Errors in Hypothesis Test. Within the domain of hypothesis testing, there exist two distinct categories of errors that can potentially arise: Type I error and Type II error. A Type I error occurs when the null hypothesis is erroneously rejected, suggesting that the Salesforce Industries CPQ system requires a longer duration to generate a quote than the hypothesised average of 15 s, despite the null hypothesis being accurate. The quantification of this error is denoted by the symbol alpha (α), commonly referred to as the level of significance, which was established at a value of 0.05 for this particular scenario. The act of committing a Type I error within a

commercial setting may lead to unjustified investments in a new Configure, Price, Quote (CPQ) system or unneeded modifications to an already established system.

On the other hand, Type II error occurs when the null hypothesis is not rejected even when it is actually wrong. Within the framework of the Salesforce Industries CPQ system, this refers to the failure to identify that the duration required for quotation generation above a threshold of 15 s. The likelihood associated with making a Type II error is commonly represented by the symbol beta (β). In the context of a business setting, this error has the potential to result in a failure to allocate resources towards the acquisition of a necessary Configure, Price, Quote (CPQ) system or implement crucial modifications to an already established system. Consequently, this oversight may result in a reduction in operational effectiveness and output.

When the null hypothesis is rejected based on the results of the t-test, it may be inferred that the mean duration for the Salesforce Industries CPQ system to generate a quote exceeds 15 s. However, the failure to reject the null hypothesis indicates that the duration for generating a quote is either equal to or shorter than the hypothesised average of 15 s. These findings can provide valuable insights for business decision-making, affecting strategic choices such as the decision to invest in a new Configure, Price, Quote (CPQ) system or to change a current system in order to improve operational efficiency.

4.4 End Result

ABC Insurance underwent a digital transformation by adopting Salesforce Industries Configure, Price, Quote (CPQ) as part of their operational strategy. The implementation of this solution was crucial in the modernisation of their digital processes and streamlining the management of the insurance life cycle. ABC Insurance utilised Salesforce Industries CPQ to automate multiple phases of the insurance policy lifecycle, encompassing activities such as quoting, underwriting, and policy administration. The implementation of automation has resulted in a substantial decrease in the requirement for manual intervention, hence mitigating the occurrence of errors often associated with manual data entry. The implementation of the CPQ technology enabled ABC Insurance to effectively create and manage a centralised product catalogue. Furthermore, it enabled agents to easily configure complex products and services using a user-friendly interface, hence minimising the need for technical expertise.

Furthermore, Salesforce Industries CPQ functioned as a comprehensive sales automation solution for ABC Insurance. This factor had a role in the improvement of sales effectiveness, the acceleration of the sales cycle, and the enhancement of the customer experience. The implemented solution facilitated the efficient production of precise quotations, facilitated the smooth integration of various systems and processes, and contributed to the enhancement of pricing and discount strategies. Following the accomplished execution of the project, the researcher proceeded to



Fig. 3 Evaluations of salesforce Industries CPQ provided by specific ABC Insurance staff members

gather evaluations and feedback from the participating personnel. The results gathered from the study demonstrated a favourable effect of the implemented solution on the operational efficiency of ABC Insurance (Fig. 3).

In general, the implementation of Salesforce Industries CPQ facilitated the optimisation of insurance life cycle management for ABC Insurance. This resulted in the reduction of manual labour, elimination of errors, and improvement of the overall customer experience.

5 Discussion

5.1 Advantages of Salesforce Industries CPQ

The study identified some advantages of implementing Salesforce Industries CPQ for insurance firms based on interviews and observations conducted at ABC Insurance Company.

Streamlining Sales Process. Salesforce Industries CPQ enhances and accelerates the generation of precise quotations and the underwriting process by automating manual tasks. This particular functionality enhances the efficiency of insurance sales representatives by empowering them to generate accurate quotations and finalise sales promptly. There exist multiple approaches via which Salesforce Industries CPQ can optimise and streamline the sales workflow.

- Quicker and more precise quotation processes.
- Enhanced teamwork and cooperation.
- Streamlined workflows through automation.
- Elevated levels of customer satisfaction.

Integration with Salesforce CRM. Salesforce Industries CPQ has been specifically developed to facilitate seamless integration with Salesforce, effectively utilising the Salesforce platform as its foundational architecture. The usage of this interface allows insurance enterprises to leverage the comprehensive Customer Relationship Management (CRM) features provided by Salesforce, while concurrently improving

their quoting and underwriting procedures through the incorporation of Salesforce Industries CPQ. This section provides an analysis of the several methods in which Salesforce Industries CPQ aligns with Salesforce.

- Streamlining sales processes.
- Customising product settings.
- Automating workflows.
- Synchronising data in both directions.

Quote Customisation. Salesforce Industries CPQ empowers sales professionals to personalise bids or proposals to correspond with a customer's distinct requirements and financial limitations. This objective is accomplished through the careful selection of suitable coverage options and pricing tiers. We kindly request permission to provide a comprehensive explanation of the aforementioned procedure.

- Configuration of products.
- Calculation of pricing.
- Creation of tailored proposals.
- Instantaneous updates.

Handle Complex Pricing and Discount Structure. Salesforce Industries CPQ is specifically built to efficiently handle intricate pricing and discount frameworks within the insurance sector. The next section delineates some methodologies that Salesforce Industries CPQ can utilise in order to effectively navigate and resolve these structures.

- Engine for pricing rules.
- Management of discounts.
- Analytics for pricing.
- Optimisation of prices.

Automation of Quote-to-Policy Process. Salesforce Industries CPQ is specifically engineered to optimise the complete workflow, encompassing the development of quotations, policies, proposals, and contracts, with the aim of enhancing efficiency and effectiveness. There exist various methodologies via which Salesforce Industries CPQ automates the aforementioned procedures.

- Creation of proposals.
- Generation of contracts.
- Integration of electronic signatures.
- Automation of workflows.

Improvement in Sales Efficiency. Salesforce Industries CPQ enhances sales efficiency and productivity in businesses by optimising the sales process, minimising inaccuracies, and offering tools for the monitoring and assessment of sales performance.

- Accelerated quotation creation.
- Enhanced precision in quoting.

- Tailored proposal generation.
- Simplified contract creation process.
- Elevated sales reporting efficiency.

Handling Product and Pricing Updates with Data Accuracy. Salesforce Industries CPQ enhances sales efficiency and productivity in businesses by optimising the sales process, minimising inaccuracies, and offering tools for the monitoring and assessment of sales performance.

- Management of product catalogue.
- Keeping versions.
- Oversight of changes.
- Incorporation with backend systems.

In order to ensure the accuracy of data across many systems, Salesforce Industries CPQ offers a variety of functions, such as tools for data validation and synchronisation, data mapping, and data cleansing. This practice guarantees the precision and consistency of data across all systems, hence reducing the probability of errors and upholding a standardised customer experience at all points of engagement.

Support for Multilanguage and Multicurrency. Salesforce Industries CPQ offers comprehensive assistance for a diverse array of currencies and languages, facilitating the effective promotion of products and services by enterprises on a global scale. This section presents an overview of the several strategies implemented by Salesforce Industries CPQ to efficiently manage the supply of support for multiple currencies and languages.

- Support for multiple currencies.
- Support for multiple languages.
- Translation of languages.
- Localised services.

5.2 Challenges Using Salesforce Industries CPQ

The study revealed that selected sales reps at ABC Insurance firm reported encountering problems while employing Salesforce Industries CPQ. Nevertheless, these issues were successfully mitigated by a collaborative approach. The insurance company faced multiple challenges and employed diverse tactics to address them.

- Complexity.
- Incorporation.
- Tailoring.
- Management of information.
- Acceptance by users.

6 Conclusions

This paper examines the effects of integrating Salesforce Industries CPQ with Salesforce CRM on the improvement of sales processes, total company value, and the development of successful implementation techniques. The study aims to investigate fundamental inquiries pertaining to the value proposition, implementation process, and advantages associated with the integration of Salesforce Industries CPQ. This study centres on the implementation of Salesforce Industries CPQ by ABC Insurance with the aim of optimising sales procedures and improving overall efficiency. The study emphasises the role of the CPQ solution in facilitating the management of intricate product and pricing systems within the insurance company. It enables the creation of customised bids and policies, as well as the automation of diverse sales processes. The research highlights that Salesforce Industries CPQ has made a substantial impact on enhancing sales and customer service at ABC Insurance, despite encountering obstacles such as data accuracy, system integration, and post-deployment performance concerns. Based on the research findings, it is evident that the implementation of Salesforce Industries CPQ has a significant impact on enhancing multiple facets of sales operations, ultimately leading to a heightened organisational value. The research places significant emphasis on improvements in sales productivity, hit rate, quotation productivity, accuracy, labour, manufacturing costs, and the removal of errors. The paper posits that the expeditious return on investment renders CPQ solutions, such as Salesforce Industries CPQ, a financially prudent undertaking, contingent upon the availability of adequate funds and human resources for deployment. The study, which employs qualitative and empirical research methods, offers useful insights despite certain limits in terms of scope. Interviews were conducted to address emerging questions, and the study included a hypothesis test to confirm the efficiency of Salesforce Industries CPQ's quote generation process. While the methodology was deemed appropriate for the subject, the absence of quantitative data constrained the depth of the outcomes. The study is expected to assist ABC Insurance in refining its CPQ system and evaluating its implementation project. Furthermore, it serves as a valuable resource for training new users and stakeholders interested in Salesforce Industries' CPQ system. Ultimately, the research contends that Salesforce Industries CPQ is instrumental in helping companies upgrade their sales processes, stay competitive in dynamic industries, and achieve improvements in productivity, accuracy, and customer satisfaction.

6.1 Future Recommendations

Following our study about the implementation of Salesforce Industries CPQ at ABC Insurance Company, the researcher proposes the subsequent suggestions for prospective enhancements:

1. Consistently revise and manage the product catalogue.

2. Provide training to staff members on system usage.
3. Incorporate automation for the creation of proposals and contracts.
4. Regularly assess and analyse system performance.
5. Ongoing solicitation of customer feedback.

By incorporating these recommendations, ABC Insurance Company has the opportunity to enhance the efficacy and efficiency of its sales procedures, ultimately leading to improved customer satisfaction.

References

1. Kock CJ (2001) Strategy and structure in developing countries: business groups as an evolutionary response to opportunities for unrelated diversification. *Ind Corp Chang* 10:77–113. <https://doi.org/10.1093/icc/10.1.77>
2. Jordan M, Auth G, Jokisch O, Kühl J-U (2020) Knowledge-based systems for the Configure Price Quote (CPQ) process—a case study in the IT solution business. *Online J Appl Knowl Manage* 8:17–30. [https://doi.org/10.36965/ojakm.2020.8\(2\)17-30](https://doi.org/10.36965/ojakm.2020.8(2)17-30)
3. Salesforce, Inc. The 6 greatest benefits of CRM platforms. <https://www.salesforce.com/crm/benefits-of-crm>
4. Salesforce, Inc. Salesforce ranked #1 in CRM market share for ninth consecutive year. <https://www.salesforce.com/news/stories/idc-crm-market-share-ranking-2022/>
5. Salesforce, Inc. CRM Software: customer relationship management. <https://www.salesforce.com/in/crm/>
6. Zuora, Inc. Zuora CPQ | quoting software for subscription businesses. <https://www.zuora.com/products/cpq-software/>
7. Salesforce, Inc. Products. <https://www.salesforce.com/solutions/industries/communications/communications-cloud/>
8. Hvam L, Pape S, Nielsen MK (2006) Improving the quotation process with product configuration. *Comput Ind* 57:607–621. <https://doi.org/10.1016/j.compind.2005.10.001>
9. Trentin A, Perin E, Forza C (2012) Product configurator impact on product quality. *Int J Prod Econ* 135:850–859. <https://doi.org/10.1016/j.ijpe.2011.10.023>
10. Gill J, Mathur G. Configure, price, and quote (CPQ) capabilities. <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/human-capital/us-consulting-cpq-capabilities.pdf>
11. Ostrow P. Configure/price-quote: better, faster sales deals enabled. https://cdn2.hubspot.net/hub/300410/file-2169344733-pdf/assets/CPQ_Endeavor_12.1.2014.pdf
12. Cicman J, Pfeiffer E (2020) The Forrester Wave™: B2B commerce suites, Q2 2020. <https://www.forrester.com/report/the-forrester-wave-b2b-commerce-suites-q2-2020/RES157277>
13. FEXLE services private limited: do you know what CPQ is? Learn the ways it can help your business with faster quote generation. https://www.linkedin.com/pulse/do-you-know-what-cpq-learn-ways-can-help/?trk=pulse-article_more-articles_related-content-card
14. Salesforce, Inc. Salesforce CPQ and billing. <https://www.salesforce.com/in/products/sales-cloud/tools/cpq-software/>
15. CoE S. Top 9 salesforce CPQ benefits | prepare accurate quotes in moments. <https://www.demandblue.com/salesforce-cpq-benefits/>
16. Gartner Inc. Magic quadrant for configure, price and quote application suites. <https://www.gartner.com/en/documents/4020916>
17. Salesforce, Inc. Salesforce revenue cloud. <https://www.salesforce.com/products/cpq/>
18. Salesforce, Inc. World's #1 CRM. <https://www.salesforce.com/campaign/worlds-number-one-CRM/>
19. Oracle. CPQ | Oracle. <https://www.oracle.com/applications/customer-experience/cpq/>

20. SAP. SAP CPQ | configure price quote solutions. <https://www.sap.com/products/financial-management/cpq.html>
21. Conga. Configure, price, and quote (CPQ) with ease. <https://conga.com/products/conduct-commerce>
22. PROS: smart configure price quote. <https://pros.com/products/cpq-software/>
23. Das S, Gayen PK, Pal S, Nayyar A (2023) Quality and leakage detection based water pricing scheme for multi-consumer building with real-time implementation using IoT. *Multimed Tools Appl* 82:26317–26352. <https://doi.org/10.1007/s11042-023-14402-4>
24. Jeyalaksshmi S, Akila D, Padmapriya D, Suseendran G, Pal S (2021) Human facial expression based video retrieval with query video using EBCOT and MLP. In: Peng SL, Hao RX, Pal S (eds) Proceedings of first international conference on mathematical modeling and computational science. *Advances in intelligent systems and computing*, vol 1292. Springer, Singapore. https://doi.org/10.1007/978-981-33-4389-4_16
25. Mukherjee D, Ghosh S, Pal S, Akila D, Jhanjhi NZ, Masud M, AlZain MA (2022) Optimized energy efficient strategy for data reduction between edge devices in cloud-IoT. *Comput Mater Continua* 72:125–140. <https://doi.org/10.32604/cmc.2022.023611>
26. Chakrabarty A, Tagiya M, Pal S, Cuong NHH (2020) Managing psychosomatic disorders related to obsession and addictions to gadgets through IoT surveillance. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_64
27. Pal S, Jhanjhi NZ, Abdulbaqi AS, Akila D, Alsubaei FS, Almazroi AA (2023) An intelligent task scheduling model for hybrid internet of things and cloud environment for big data applications. *Sustainability* 15:5104. <https://doi.org/10.3390/su15065104>
28. Suseendran G, Doss S, Pal S, Dey N, Quang Cuong T (2021) An approach on data visualization and data mining with regression analysis. In: *Advances in intelligent systems and computing*, pp 649–660. https://doi.org/10.1007/978-981-33-4389-4_59
29. Norbu T, Mall M, Sarkar B, Pal S, Chakrabarty A (2020) Revitalizing MSMEs' performance with transparency: monitoring, mentoring and Malwaring through IoT intervention. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_63
30. Tagiya M, Sinha S, Pal S, Chakrabarty A (2020) Transformation from HRM inadequacy and bias-syndrome to transparent and integrated ecosystem through IoT-intervention in career management. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_61

Analyze and Compare the Public Cloud Provider Pricing Model and the Impact on Corporate Financial



Jaideep Singh, Souvik Pal, Bikramjit Sarkar, H. Selvi, Saurabh Adhikari, K. Madhumathi, and D. Akila

Abstract In order to better understand the relative costs of IaaS, PaaS, and SaaS, it is recommended that this study collect data on these three categories of cloud computing services. The purpose of this research is to analyze how much money is saved or lost while adopting cloud services for different-sized enterprises. Its purpose is to identify and quantify outcomes that are beneficial to society, whether they be monetary or otherwise. Additionally, we want to provide a complete picture of cloud services' economic effect by itemizing their benefits and drawbacks. The collected data is stored in a secure location and treated ethically to prevent unauthorized access. Then, any conflicts interesting are investigated thoroughly and resolved. The researcher

J. Singh

Liverpool Business School, Liverpool John Moores University, Liverpool, UK

S. Pal (✉)

Department of Management Information Systems, Saveetha College of Liberal Arts And Sciences, Saveetha Institute of Medical And Technical Sciences,, Chennai, India

e-mail: souvikpal22@gmail.com

Department of Computer Science and Engineering, Sister Nivedita University (Techno India Group), Kolkata, India

B. Sarkar

Department of Computer Science and Engineering, JIS College of Engineering, Kalyani, India

H. Selvi

Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed University, Chennai, India

e-mail: selvib.sclas@saveetha.com

S. Adhikari

School of Engineering, Swami Vivekananda University, Kolkata, India

e-mail: saurabhadhikari@svu.ac.in

K. Madhumathi

Department of Computer Application, Anna Adarsh College for Women, Chennai, India

D. Akila

Department of Computer Applications, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed University, Chennai, India

may be certain that the study is being conducted in a responsible and moral manner if they follow these criteria. The research summed up with the conclusion that every device is capable of running any SaaS application. It could be a smartphone, laptop, desktop computer, or tablet. In this manner, businesses can better protect their data, and employees may take advantage of the flexibility of accessing services from any device so long as they have their secure line key.

Keywords Cloud · SaaS · IaaS · PaaS · Information technology

1 Introduction

Businesses may save a lot of money on costly hardware and infrastructure by using the public cloud and creating applications for several platforms. In this model, companies may save money by not purchasing extraneous devices and instead paying for just the services they use. Companies may choose the most profitable cloud service as a means to raise their profits even more. There's no denying the widespread use of the public cloud services by organizations of all sizes and ages, but notably startups. Instead of spending money on brand new technology, they know it's better to pay someone to oversee the current setup.

Companies have been using the public cloud for some time, but they might be getting better deals if they shopped about. Companies used to blindly follow the advice of their infra suppliers without ever comparing pricing.

Businesses have been quick to embrace IaaS and SaaS as the two most popular offerings. Previous pricing studies will be analyzed in this research using a literature analysis [1]. Looking at the current literature, you will see that there are little comparison and study of how various clouds charge on the same services. A few things are as crucial to a company's success as its ability to cut costs without compromising quality or productivity. Businesses in the B2B sector need to prioritize both profit and sustainability if they want to thrive [2].

Cloud computing is becoming more essential as a means of providing IT services through the Internet. A sizable crowd may now pool their resources and divide up the expenses on an as-needed basis, all thanks to this innovation. Users may quickly and cheaply do analysis, management, and storage of data. With cloud computing, users' information is stored in a central location and can be accessed from any Internet-connected device, anywhere in the world, with no additional software downloads or installations required.

Cloud computing has emerged as a popular topic of discussion among computer scientists because of its potential to significantly cut down on the amount of time and effort spent on computing and provides Internet access to resources and information.

Among the various offerings available from cloud computing providers, some of the most common include infrastructure as a service (IaaS), storage as a service (STaaS), platform as a service (PaaS), test environment as a service (TEaaS), and software as a service (SaaS). While the fundamental goal of cloud computing clients is

to get the highest possible quality of service (QoS) at lowest feasible cost, the primary goal of cloud computing providers is to increase the provider's revenues. That's why it's crucial to have a pricing plan that works for everyone. Service providers may enhance demand for their products and services by adjusting the prices at which they are offered [3].

As consumers have become more demanding and critical, and as they expect to be able to access information on their own terms at any time, information technology (IT) has become an increasingly important aspect in optimizing core services and keeping a competitive advantage. Thanks to the cloud as well as other digital technologies, Nsurtechs are able to thrive in today's business environment [4]. Their success may be attributed to a unique selling proposition [4] that emphasizes the convenience and attractiveness of digital products that can be bought online. By focusing on automated services and simplified processes, they deliver a high-quality client experience and a degree of involvement that is absent from traditional insurance organizations [4]. The service will be available 24/7/365 on whatever device the customer chooses. Disruptors in the insurance sector are companies that provide cutting-edge policy administration, guidance online and pricing comparisons, and assistance after a claim. These firms have an advantage over conventional insurers since they improve several points along the insurance industry's chain of value. To combat the challenge presented by upstart insurtechs, long-standing businesses must provide compelling products and services of their own.

Findings from this research suggest that providers of public cloud services may gain an advantage by adopting a very different business strategy than traditional data center outsourced service providers. In this Leadership Insight, we look at the price implications of the general public cloud operating model, sometimes known as the Four Pillars of Public Cloud price.

The suggested research goal of this project is to collect pricing data for three distinct cloud computing service categories: infrastructure as a service (IaaS), platform as a service (PaaS), as well as a service (SaaS). Using the given price data, we can get access to and compare the best services for startups, SMBs, and enterprises. Which cloud service offers the best return on investment, and for what range of businesses?

The study's objectives are listed in bullet form below: For the purpose of comparing the costs of various services offered by public cloud providers. Identifying the public cloud service providers used by enterprises of varying sizes. Examining the cost-savings potential of using different public cloud service providers. To learn whether combining two or more public cloud services may provide the greatest potential savings and service quality.

2 Literature Survey

Nezami et al. [5], This study analyzes how migrating to the cloud will affect the We utilized a longitudinal data set consisting of 435 listed on the public B2B companies in the IT services and software industries to analyze the wealth that as shareholders from the point of view of the suppliers. Using the value relevance model, we discover that an out-of-sample increase in stock returns is positively correlated with an out-of-sample decrease in idiosyncratic risk when a cloud ratio (the percentage of a company's revenue which comes from the use of cloud computing increases unexpectedly).

Lowe and Galhotra [6], The authors examine the pros and cons of the “pay as you go” pricing model for a cloud by comparing the current prices of the leading cloud service providers in the market. In this analysis, we compare the provider’s price structure to those of comparable services offered by other companies. The writers want to look into whether or not prices are fair so that they may set more appropriate rates for their offerings. Because of the high upfront costs of perpetual licensing, pay-per-use models have emerged as a method of combatting software piracy and appealing to the marginalized and varied user base that has been found to depend on pirated software. Piracy occurs when a small user base refuses to pay the expensive one-time fee of a perpetual license.

Chang [7], The purpose, technology, case studies, and contributions of virtual reality and cloud computing for Emerging Services and Analytics are all discussed in this article at a high level. Emerging Services and Analytics using virtual reality have already proven useful in disciplines as diverse as healthcare, business, changes in the climate, and natural catastrophes, demonstrating the technology’s potential to aid scientists in understanding the complexity of respective subjects. The general public, not simply scientists with a high level of expertise, can grasp some of the conclusions and examples from other fields.

Laatikainen [8], This research contributes to the growing body of literature on cloud pricing by proposing a cloud-specific pricing structure, exploring the correlation between software design and cost, and analyzing how widespread use of cloud computing has impacted conventional methods of pricing software. The study also contributes to the concurrent sourcing and cloud cost literatures by developing analytical models that take into account the effect that various cost determinants have on a relative cost-effectiveness of private, public, and hybrid storage. This study’s findings imply that the private cloud is preferable to a cloud that is public when storage demands rise exponentially, provided that the firm reevaluates its storage requirements and makes the necessary in-house investments.

Ibrahimi [9], The importance of cloud computing in protecting digital information was underlined. Instantaneous pricing and resource distribution for a huge user base are made possible. Customers may save time and money by using it to do analyses, manage, and store data. Knowing what makes consumers uneasy about cloud computing service is especially important when switching to a new pricing strategy. Customers are influenced more by a salesperson’s presenting of the advantages of

the service than by the price; however, the price is an important consideration and an indicative of the quality of the service. Different cloud provider policies and strategies related concerns [10–17] have been described using artificial intelligence, machine learning, and deep learning methods, which provisions us to make the analysis.

Summary

We have a cost calculator for all public platforms service providers; however, there aren't any tabs for comparison and analysis yet. A price calculator may be used to examine the resource cost, but it does not reveal the other platform needs. After doing this analysis, businesses will have charts of comparison for all public providers of cloud services at their disposal, allowing them to make informed judgments.

Many businesses now rely on cloud computing solutions like IaaS and SaaS. The company plans to look into the most cost-effective cross-public cloud platform alternative after this research is finished.

3 Proposed Work

This study set intended to answer questions concerning how choosing and paying for public cloud services impacts enterprises financially. This dissertation will be based on a combination of primary data from a variety of firms and secondary data acquired from previous research. Thus, data gathering required communication between the researcher and the participants.

3.1 Research Design

For research to be both comprehensive and timely, its execution must be well-planned. After identifying a research gap, the next step is to plan how you'll go about solving it. According to Kothari [18], “the conceptual framework within which investigation will be conducted; that it represents a blueprint of the gathering, measurements, as well as analysis of data.”

Kothari [18] notes that the three most prevalent types of scientific inquiry are exploratory, descriptive, and hypothesis-testing. According to Kothari [18], exploratory studies aim to “formulate problems for more specific inquiries or of creating working hypotheses as an operational viewpoint of view.”

3.2 Participants and Procedures

The process of conducting an online survey entails a number of steps, including planning, data collection, analysis, reporting, and application. According to Ritter

and Sue [19], each of these overarching approaches requires just a small number of well-defined phases.

Identify and engage the stakeholder: Stakeholders are interested parties, whether they be individuals or institutions.

Determine resources: The quantity of resources devoted to the survey is based on how long respondents spend on the survey link, that is generated using the free online survey tool.

Writing goals and objectives: Our goal in conducting this survey is to get insight into how widely adopted cloud services are among businesses from the viewpoints of people who have used them.

Evaluation or results: The study's results will be analyzed in order to better advise new users of cloud services on which cloud services model and pricing approach would best meet their needs.

Using software to implement surveys: The survey button is created using Google Forms. The survey's link will stay up so that we may gather as many replies as possible to help us find ways to reduce the risk.

4 Result and Discussion

This section compares the costs of using Azure, AWS, and Google Cloud Platform (GCP), three popular cloud service providers. Database, compute (virtual machine), storage, and AD Service costs, as well as their effect on a company's bottom line, have been tallied. We can assess and analyze the finest services for small, medium, and large-scale businesses based on the price information provided.

4.1 Database Pricing

Cloud companies like Azure, AWS, and GCP have their own database service prices listed in Table 1. The following table compares the monthly costs of many different database storage options. Azure has monthly rates that are less than both AWS and GCP.

4.2 Storage Pricing

Storage service price from a few popular cloud vendors, including Azure, AWS, and GCP, is shown in Table 2. Monthly rates for various storage options are shown below. Again, Azure has monthly price that is cheaper than both AWS and GCP.

Table 1 Database pricing

Cloud service	Azure	AWS	GCP
<i>Database</i>			
Azure SQL database	Prices begin at \$4.99 monthly for space capacity of 5 GB	Prices begin at \$15.72 monthly for space capacity of 20 GB	Prices begin at \$7.45 monthly for space capacity of 10 GB
Azure database for MySQL	Prices begin at \$9.99 monthly for space capacity of 50 GB	Prices begin at \$16.58 monthly for space capacity of 20 GB	Prices begin at \$7.10 monthly for space capacity of 10 GB
Azure database for PostgreSQL	Prices begin at \$9.99 monthly for space capacity of 50 GB	Prices begin at \$16.58 monthly for space capacity of 20 GB	Prices begin at \$7.10 monthly for space capacity of 10 GB
Amazon RDS for SQL server	Prices begin at \$17.55 monthly for space capacity of 20 GB	Prices begin at \$15.16 monthly for space capacity of 20 GB	N/A
Amazon RDS for MySQL	Prices begin at \$16.38 monthly for space capacity of 20 GB	Prices begin at \$15.10 monthly for space capacity of 20 GB	N/A
Amazon RDS for PostgreSQL	Prices begin at \$15.40 monthly for space capacity of 20 GB	Prices begin at \$15.10 monthly for space capacity of 20 GB	N/A
Cloud SQL for MySQL	Prices begin at \$7.00 monthly for space capacity of 10 GB	N/A	Prices begin at \$7.00 monthly for space capacity of 10 GB
Cloud SQL for PostgreSQL	Prices begin at \$7.00 monthly for space capacity of 10 GB	N/A	Prices begin at \$7.00 monthly for space capacity of 10 GB
Cloud spanner	Prices begin at \$0.90 hourly for space capacity of 2 TB and 1 node	N/A	N/A

Table 2 Storage pricing

Storage			
Storage accounts	Prices begin at \$0.02 (monthly) for 1 GB	Prices begin at \$0.023 (monthly) for 1 GB	Prices begin at \$0.010 (monthly) for 1 GB
Disk storage of VM	Prices begin at \$0.83 for space capacity of 4 GB	N/A	Prices begin at \$0.04 (monthly) for 1 GB

4.3 Compute (VM) Pricing

Prices for many virtual machine (VM) services from popular cloud providers including Azure, AWS, and GCP are shown above. Table 3 displays monthly costs

Table 3 Computer (VM) pricing

Compute			
General purpose	For 16 GB, the prices are \$0.166	For 16 GB, the prices are \$0.1344	For 16 GB, the prices is \$0.15092
Compute-optimized	For 8 GB, the prices are \$0.1690	For 8 GB, the prices are \$0.153	For 16 GB, the prices are \$0.2351

Table 4 Business outlook

	First quarter fiscal 2022	Full year fiscal 2022
Revenue	Range between \$68.0 and \$70.0 M	Range between \$292.0 and \$300.0 M
Subscription revenue	Range between \$34.0 and \$35.0 M	Range between \$151.0 and \$155.0 M
Adjusted EBITDA	Range between \$2.5 and \$3.5 M	Range between \$16.0 and \$18.0 M

associated with using various VMs. In this case, we may argue that AWS provides cheaper per-GB price than Azure and GCP.

4.4 AD Services

Azure has the most cost-effective monthly P1 and P2 prices of all of the cloud services shown in the table. Similarly, AWS begins at the lowest rate for AD domain services, whereas GCP gives the same price each region. This is also true for AD External Identity (more than 50,000 MAU), which offers premium services P1 as well as P2 per month.

4.5 Pricing Influence on Cloud Service Adoption Among Corporate Financials

Infosys—Large-scale company.

Cloud service—Microsoft's cloud service Azure.

Infosys and Microsoft Azure collaborate to provide cloud transformation solutions for businesses of all sizes and in all sectors. A Microsoft Azure Managed Service Provider with Expertise is of great value to their customers. Infosys offers a full suite of cloud transformation and management services on Microsoft Azure.

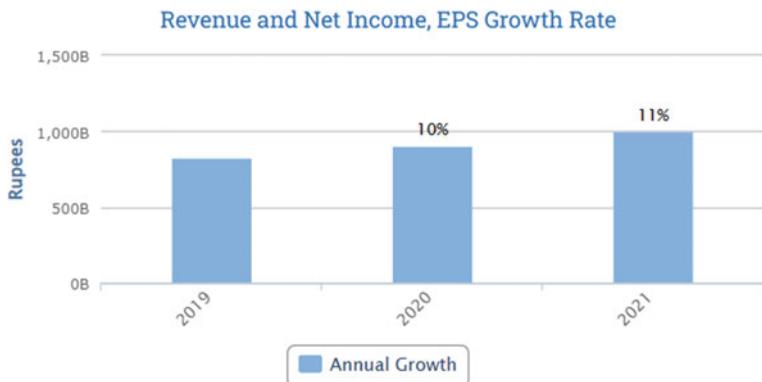


Fig. 1 Growth

4.6 Growth

See Fig. 1.

4.7 Duck Greek Technology (Small-Scale Company)

Cloud Platform-Microsoft Azure

Duck Greek Technology is the leading innovator in P&C and general insurance with its innovative and forward-thinking solutions.

The Microsoft mission: Increase the potential of every individual and institution on Earth.

Technology developed by Microsoft has become the norm in many fields. Duck Greek OnDemand is an end-to-end software as a service (SaaS) product that provides access to the whole Duck Greek Suite of apps. It is built on top of Microsoft's Azure cloud platform. Duck Greek uses Microsoft Azure because it delivers the scalability, stability, and adaptability necessary to provide P&C carriers using future-ready technologies delivered as SaaS, hence facilitating increased velocity, dexterity, and inventiveness.

4.8 Business Outlook

The following forecast for the first quarter and full year of DC's 22nd fiscal year is being issued on the basis of existing assumptions as of October 14, 2021.

4.9 Mondelez International (*Mid-Scale Company*)

Cloud Platform-GCP

E-commerce and other forms of online distribution have revolutionized the consumer-packaged goods (CPG) industry. Rather of relying on middlemen like stores and magazines to spread the word about their wares, producers may now sell online and promote their items via digital channels to reach their target audience directly and efficiently.

4.9.1 Growth of Mondelez International

Mondelz International has revealed a new plan to boost growth and streamline its product offerings in order to create high, long-term shareholder value. Thus, the firm is revising its long-term formula to predict Organic Net Revenue growth of 3–5%, up from 3% or more in the past (Fig. 2).

Long-term, Mondelz International wants to reshape its portfolio such that its chocolate, biscuits, and baked snack products account for 90% of sales. There is a lot of space to boost chocolate and biscuits' penetration and per capita consumption in both established and developing regions.

Successfully attracting, developing, and keeping the CPG industry's top talent.

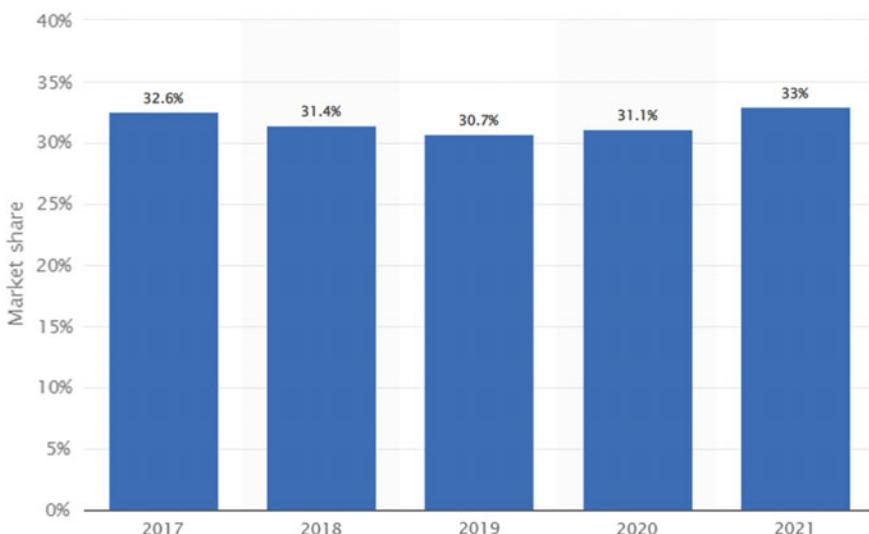


Fig. 2 Growth of Mondelez International

5 Conclusion

Technology has a major influence on modern commercial activity. In the past, businesses needed to set up and manage their own infrastructure for servers in order to host and operate apps. One of the most important aspects of cloud computing is its low cost. To rephrase, you are charged based on actual consumption. This forces companies to prioritize spending wisely and acquire just the services essential to their survival in the market. The aforementioned flexibility means that you may pick and choose which cloud services you really employ. Platform as a service (PaaS) is a nice illustration of this kind of service. This service is well-liked by programmers since it provides them with a reliable environment in which to operate without requiring them to take care of the platform themselves. Using Infrastructure as a service (IaaS) and other cloud services, for instance, businesses may subscribe to a fully functional IT infrastructure without investing in any physical equipment. Software as a service (SaaS), a cloud solution, enables the usage of any program, on any device, inside an organization.

In conclusion, businesses must carefully assess their requirements and use patterns to choose the most appropriate pricing plan for the public cloud. In order to get the most out of their cloud investments and boost their bottom lines, businesses of all sizes would do well to embrace cloud cost management best practices, regardless of the model of pricing they choose.

References

1. Muslmani BK, Kazakzeh S, Ayoubi E, Aljawarneh S (2018) Reducing integration complexity of cloud-based ERP systems. In: Proceedings of the first international conference on data science, e-learning and information systems, pp 1–6
2. Zhu G, Chou MC, Tsai CW (2020) Lessons learned from the COVID-19 pandemic exposing the shortcomings of current supply chain operations: a long-term prescriptive offering. *Sustainability* 12(14):5858
3. Mazrekaj A, Shabani I, Sejdiu B (2016) Pricing schemes in cloud computing: an overview. *Int J Adv Comput Sci Appl* 7(2)
4. Poustchi K, Gleiss A (2019) Surrounded by middlemen—how multi-sided platforms change the insurance industry. *Electron Mark* 29(4):609–629
5. Nezami M, Tuli KR, Dutta S (2022) Shareholder wealth implications of software firms' transition to cloud computing: a marketing perspective. *J Acad Mark Sci* 50(3):538–562
6. Lowe D, Galhotra B (2018) An overview of pricing models for using cloud services with analysis on pay-per-use model. *Int J Eng Technol* 7(3.12):248–254
7. Chang V (2018) An overview, examples, and impacts offered by emerging services and analytics in cloud computing virtual reality. *Neural Comput Appl* 29(5):1243–1256
8. Laatikainen G (2018) Financial aspects of business models: reducing costs and increasing revenues in a cloud context. *Jyväskylä Stud Comput* 278
9. Ibrahim A (2017) Cloud computing: Pricing model. *Int J Adv Comput Sci Appl* 8(6)
10. Le D, Pal S, Patnaik PK (2022) Cloud-based data storage. *Cloud Comput Solut* 143–164.
<https://doi.org/10.1002/9781119682318.ch9>

11. Pal S, Jhanjhi NZ, Abdulbaqi AS, Akila D, Almazroi AA, Alsubaei FS (2023) A hybrid edge-cloud system for networking service components optimization using the internet of things. *Electronics* 12:649. <https://doi.org/10.3390/electronics12030649>
12. Mukherjee D, Ghosh S, Pal S, Akila D, Jhanjhi Z, Masud N, AlZain M (2022) Optimized energy efficient strategy for data reduction between edge devices in cloud-IoT. *Comput Mater Contin* 72:125–140. <https://doi.org/10.32604/cmc.2022.023611>
13. Chakrabarty A, Tagiya M, Pal S, Cuong NHH (2020) Managing psychosomatic disorders related to obsession and addictions to gadgets through IoT surveillance. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_64
14. Pal S, Jhanjhi NZ, Abdulbaqi AS, Akila D, Alsubaei FS, Almazroi AA (2023) An intelligent task scheduling model for hybrid internet of things and cloud environment for big data applications. *Sustainability* 15:5104. <https://doi.org/10.3390/su15065104>
15. Pal S, Le D-N, Pattnaik PK (2022) Architectural framework for cloud computing, pp 39–55. <https://doi.org/10.1002/9781119682318.ch3>
16. Norbu T, Mall M, Sarkar B, Pal S, Chakrabarty A (2020) Revitalizing MSMEs' performance with transparency: monitoring, mentoring and Malwaring through IoT intervention. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_63
17. Pattnaik PK, Le D-N, Pal S (2022) Applications of mobile cloud computing, pp 243–255. <https://doi.org/10.1002/9781119682318.ch14>.
18. Kothari CR (2004) Research methodology, 2nd revised edn. New Age International Publishers
19. Sue VM, Ritter LA (2012) Conducting online surveys. Sage

A Data-Driven Analytical Approach on Digital Adoption and Digital Policy for Pharmaceutical Industry in India



Anup Rana, Bikramjit Sarkar, Raj Kumar Parida, Saurabh Adhikari, R. Anandha Lakshmi, D. Akila, and Souvik Pal

Abstract The Indian pharmaceutical industry has witnessed exponential growth over the past few decades to become the third largest globally in volume terms and a leading supplier of affordable generics. This paper provides a comprehensive assessment of the industry's evolution, structure, performance metrics, priorities, challenges and future outlook. Secondary data from government reports, industry associations, company documents and academic literature is synthesized to develop a holistic perspective. The analysis indicates that innovation, research productivity, clinical trials, ethical practices and sustainability shape the industry's development. Strategic priorities include strengthening R&D capabilities, expanding exports, leveraging digital technologies across operations, ensuring robust supply chains and

A. Rana

Liverpool Business School, Liverpool John Moores University, Liverpool, UK

B. Sarkar

Department of Computer Science and Engineering, JIS College of Engineering, Kalyani, India

R. K. Parida

School of Computing Science and Engineering, Galgotias University, Greater Noida, India

S. Adhikari

School of Engineering, Swami Vivekananda University, Kolkata, India

e-mail: saurabhadhikari@svu.ac.in

R. Anandha Lakshmi

Department of Computer Application, Anna Adarsh College for Women, Chennai, India

D. Akila (✉)

Department of Computer Applications, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed University, Chennai, India

e-mail: akiindia@yahoo.com

S. Pal

Department of Management Information Systems, Saveetha College of Liberal Arts and Sciences, Saveetha Institute of Medical and Technical Sciences, Chennai, India

Department of Computer Science and Engineering, Sister Nivedita University (Techno India Group), Kolkata, India

e-mail: souvikpal22@gmail.com

adopting global quality and ESG standards. Investments in innovation have risen but remain below international benchmarks. While generics account for majority revenues, emerging segments like biosimilars and complex drugs offer growth opportunities. Digital adoption is increasing with data analytics and AI gaining prominence. However, issues related to regulations, pricing control, IP protection, skill shortages and environmental impact need resolution through appropriate reforms and public-private collaboration. India's vision to become a leading pharmaceutical innovation and manufacturing hub would require reinforcing policy frameworks around IP protection, pricing, regulatory efficiency, skill building and technology adoption. Overall, the industry holds immense potential for beneficial growth but realizing it would necessitate concentrated efforts from stakeholders to build organizational capabilities, foster collaboration and imbibe global best practices.

Keywords Pharmaceuticals · Innovation · Digitalization · Strategy · India

1 Introduction

The Indian pharmaceutical industry has witnessed remarkable progression over the past several decades to become the third largest globally in terms of volume with over 3.5% market share [1]. From a nascent sector focused on producing formulations using imported bulk drugs in the 1960s, India has transformed into a leading global supplier of high-quality, low-cost generic medicines, vaccines and APIs [2]. The domestic pharmaceutical market is projected to reach USD 120–130 billion by 2030, highlighting the rapid growth and immense potential of the industry [1]. This paper aims to provide a comprehensive assessment of the pharmaceutical sector in India across various dimensions such as evolution, structure, priorities, challenges and future outlook.

The Indian pharmaceutical industry comprises over 10,500 manufacturers, making it highly fragmented [3]. The top 10 companies control around 30–40% share of the total revenues [4]. Sun Pharmaceutical Industries, Lupin, Dr. Reddy's Laboratories, Cipla and Cadila Healthcare are the leading players [5]. The sector provides employment to around 2.7 million people and contributes approximately 2.9% to the national GDP [6]. Pharmaceutical exports exceeded USD 24 billion in FY22, underscoring the industry's expanding global footprint [7]. India supplies around 57% of the global demand for vaccines and 50% of generics in the U.S. [8]. Affordable domestically produced medicines have been pivotal in increasing healthcare access across socioeconomic strata over the past decades [9].

However, the industry faces challenges such as complex regulatory policies, pricing control, patent laws and talent acquisition that impede innovation and growth [3]. Strategic priorities include strengthening research productivity through higher R&D investments, public–private collaboration and policy support [10]. Adoption of digital technologies, such as AI, big data analytics and IoT across drug discovery, manufacturing and pharmacovigilance is increasing [11]. Ensuring robust supply

chains, expanding exports, leveraging strategic partnerships and upholding product quality standards are vital for global competitiveness [12]. Sustainability practices and ethical integrity have gained salience as differentiators [13].

This paper provides an in-depth perspective into the Indian pharmaceutical industry based on analysis of secondary data from government reports, industry associations, company documents and academic literature. The study is organized across key aspects—evolution, structure, priorities, challenges, digital adoption trends and strategic imperatives. It provides insights into the sector's competitive landscape, performance metrics, technology usage patterns and growth levers. The analysis aims to facilitate informed decision-making by pharmaceutical companies, policymakers and other stakeholders.

2 Literature Review

The Indian pharmaceutical industry has developed rapidly over the past few decades to become a key player in the global market. The industry has benefited from factors such as a large skilled workforce, low manufacturing costs, government policies and regulatory changes. However, it also faces challenges related to infrastructure, innovation, intellectual property rights, competition and quality control that hinder its further growth.

The Indian pharmaceutical industry began to rapidly expand in the post-independence period with the establishment of several public-sector companies [14]. The industry mainly produced simple drugs initially but the liberalization of the economy in the 1990s allowed multinational companies to enter the market, promoting technological advancement and global expansion [15]. The Patents Act of 1970 enabled Indian companies to produce generic versions of patented drugs, leading to reduced drug prices and increased accessibility [16].

The government has implemented various policies to promote the industry's growth through initiatives like tax incentives, special economic zones and funding for research and development [17, 18]. However, issues like lack of adequate infrastructure, minimal innovation and intellectual property concerns have hampered the industry's development [19]. Moreover, competition from other emerging economies like China and Brazil has put pressure on Indian companies to develop new drugs to remain competitive [20].

Despite challenges, the industry has achieved rapid growth fueled by rising health-care spending, demand for generics and increasing R&D investments [1]. The adoption of digital technologies has also helped improve productivity, efficiency and speed of drug development [21]. The industry contributes significantly to India's GDP, providing employment to millions directly or indirectly [5, 22]. Indian companies have diversified their product portfolios and focused on the development of biosimilars that offer opportunities for growth [23].

The COVID-19 pandemic impacted supply chains, production and demand but also demonstrated the industry's resilience and adaptability [24]. Indian manufacturers played a key role in providing affordable healthcare solutions globally during the pandemic [25]. Different healthcare-related concerns [37–44] have been described using artificial intelligence, machine learning and deep learning methods, which support us to make the analytical approach on digital adoption and digital policy pharmaceutical industry in India.

Though the Indian pharmaceutical industry has made tremendous progress, it faces many opportunities and challenges that need to be addressed for sustainable growth. Innovation, digitalization and a strategic focus on generics and biosimilars are vital for continued success. Policy interventions and government support can help foster a globally competitive and sustainable industry by addressing infrastructure bottlenecks and promoting innovation. With the right measures, the Indian pharmaceutical industry has significant potential to become a leader in pharmaceutical research, development and manufacturing in the coming years.

3 Research Methodology

This research employs a descriptive and analytical approach using secondary data to assess the pharmaceutical industry in India. Secondary data from diverse sources including academic journals, industry reports, government publications, databases, financial reports and news articles will be collected. These sources will provide insights into the industry's history, growth, regulatory landscape, competitive dynamics, financial performance and strategic priorities.

Thematic analysis will be used to categorize the data and identify patterns related to the research objectives. Quantitative data on financial indicators over time will be evaluated using graphs, charts and ratios. Porter's five forces framework will analyze the competitive forces affecting industry profitability. SWOT analysis will identify key internal strengths/weaknesses and external opportunities/threats.

The research will use a descriptive approach to outline the pharmaceutical industry's current status. An analytical perspective will examine the strategic priorities, digital adoption and innovations of major Indian firms. The analysis aims to highlight the industry's contributions, regulatory environment, ethics, talent pool, financial performance and future growth potential. It will provide insights into the sector's opportunities and challenges within the global pharmaceutical space.

The conclusions will synthesize key findings and offer recommendations for stakeholders. Limitations include reliance on secondary data and lack of primary research. Further studies could incorporate surveys, interviews, focus groups and an international comparison. However, within its scope, this research will provide a holistic overview of the Indian pharmaceutical industry using robust secondary information from scholarly and industry sources.

4 Analysis

4.1 Financial Performance and Corporate Social Responsibility

The pharmaceutical industry plays a crucial role in ensuring the health and well-being of individuals worldwide. As the demand for medication increases, so does the competition among pharmaceutical companies. In this context, it is important for these companies to have a solid financial performance to survive and thrive in the market. This study aims to analyze the financial performance of top Indian pharmaceutical companies, including Syngene International, Cipla, Sun Pharmaceutical, Dr. Reddy's Laboratories, Aurobindo Pharma, Lupin Pharma, Cadila Healthcare, Torrent Pharma, Alkem Laboratories, Glenmark Pharmaceuticals, Divi's Laboratories, Biocon Ltd, Jubilant Life Sciences, Piramal Pharma Limited and Ipcra Lab. The financial analysis will focus on several key performance indicators, such as gross profit margin, net profit margin, return on assets, return on equity, debt to equity ratio, current ratio, quick ratio, and earnings per share. By evaluating the financial health of these companies, this study aims to provide insights into their profitability, liquidity, and solvency, which are essential for making informed investment decisions.

Gross profit margin (GPM) is a profitability ratio that shows how much profit a company makes after deducting the cost of goods sold. The GPM for the pharmaceutical industry varies between companies. In FY21, Syngene International had the highest GPM of 31.5% (Fig. 1), followed by Piramal Pharma Limited with 58.7%. In FY22, the highest GPM was also for Piramal Pharma Limited with 44.9%, followed by Divi's Laboratories with 43.3%.

Net profit margin (NPM) is a profitability ratio that shows how much profit a company makes after deducting all expenses, including taxes and interest. In FY21, Aurobindo Pharma had the highest NPM of 21.8%, followed by Syngene International with 18.5%. In FY22, the highest NPM was for Divi's Laboratories with 33%, followed by Aurobindo Pharma with 11.4% (Fig. 2).

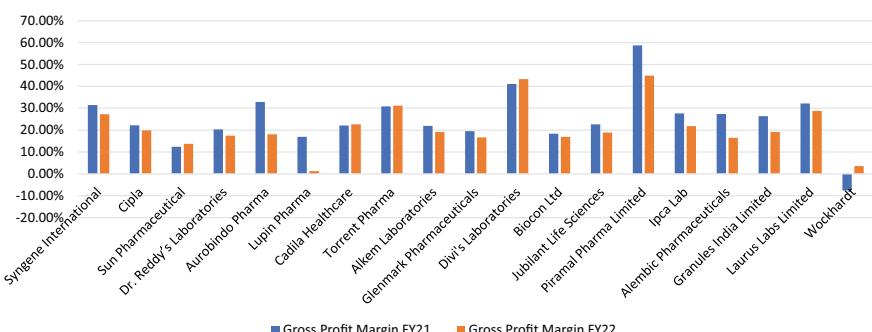


Fig. 1 Comparison of gross profit margin (FY21 versus FY22)

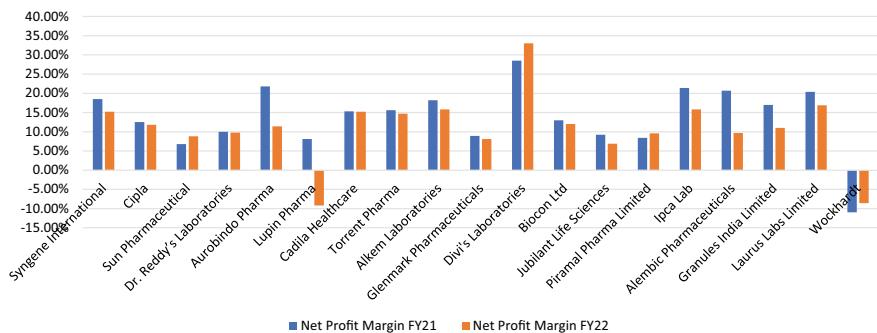


Fig. 2 Comparison of net profit margin (FY21 versus FY22)

Return on assets (ROA) is a profitability ratio that measures how efficiently a company utilizes its assets to generate profit. In FY21, Aurobindo Pharma had the highest ROA of 16.1%, followed by Syngene International with 8.4%. In FY22, the highest ROA was for Divi's Laboratories with 20.3%, followed by Aurobindo Pharma with 7.9% (Fig. 3).

Return on equity (ROE) is a profitability ratio (Fig. 4) that shows how much profit a company generates for its shareholders based on the amount of equity they have invested in the company. In FY21, Aurobindo Pharma had the highest ROE of 24.6%, followed by Divi's Laboratories with 21.4%. In FY22, the highest ROE was for Divi's Laboratories with 25.2%, followed by Aurobindo Pharma with 19.4%.

Debt to equity ratio (D/E) is a solvency ratio (Fig. 5) that shows the proportion of debt and equity a company is using to finance its assets. In FY21, Dr. Reddy's Laboratories had the highest D/E of 7.4, followed by Aurobindo Pharma with 16.1. In FY22, the highest D/E was for Dr. Reddy's Laboratories with 8.1, followed by Aurobindo Pharma with 7.9.

Current ratio and quick ratio are liquidity ratios (Fig. 6) that show a company's ability to pay off its current liabilities using its current assets. The current ratio

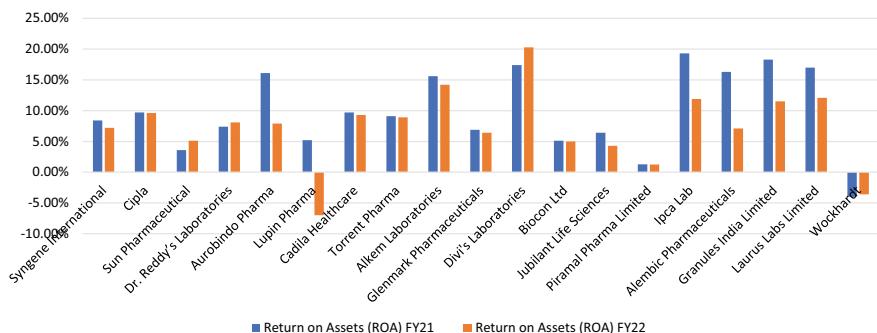
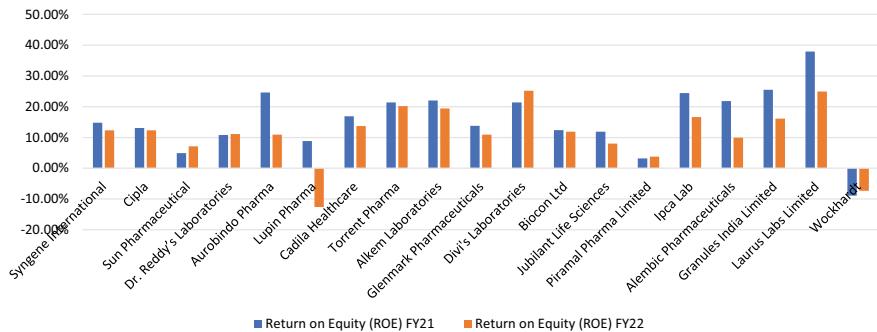
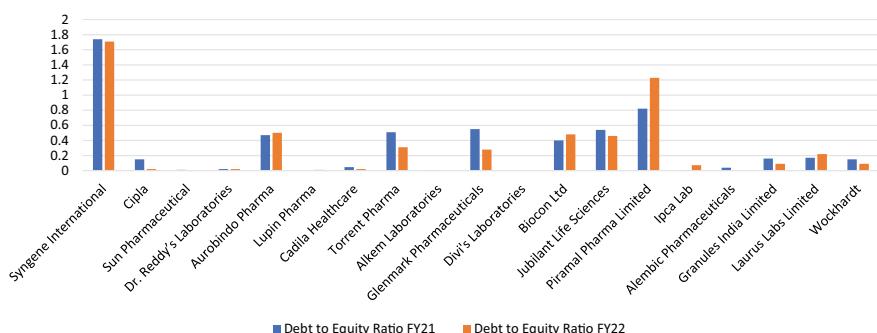


Fig. 3 Comparison of ROA (FY21 versus FY22)

**Fig. 4** Comparison of ROE (FY21 versus FY22)**Fig. 5** Comparison of debt equity ratio (FY21 versus FY22)

includes all current assets, while the quick ratio (Fig. 7) includes only quick assets, which are cash, marketable securities and accounts receivable. In FY21, Alkem Laboratories had the highest current ratio of 15.6 and quick ratio of 22.0, followed by Piramal Pharma Limited with a current ratio of 1.47 and a quick ratio of 1.12. In FY22, Alkem Laboratories had the highest current ratio of 14.2 and quick ratio of 19.4, followed by Piramal Pharma Limited with a current ratio of 1.29 and a quick ratio of 0.91.

Earnings per share (EPS) is a profitability ratio (Fig. 8) that shows the amount of profit a company generates per share of its common stock outstanding. In FY21, Divi's Laboratories had the highest EPS of 3.9, followed by Aurobindo Pharma with 1.86. In FY22, the highest EPS was also for Divi's Laboratories with 2.4.

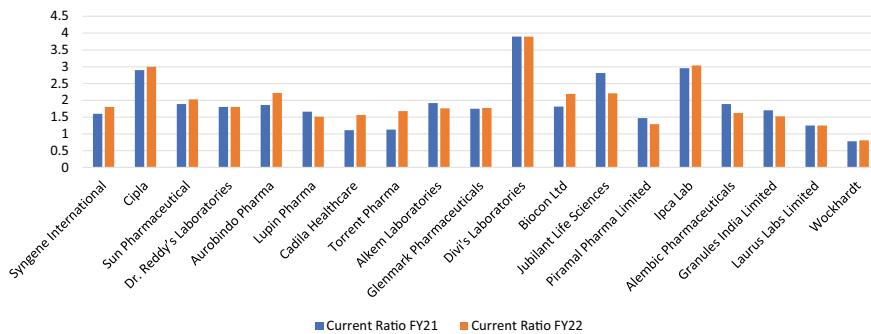


Fig. 6 Comparison of current ratio (FY21 versus FY22)

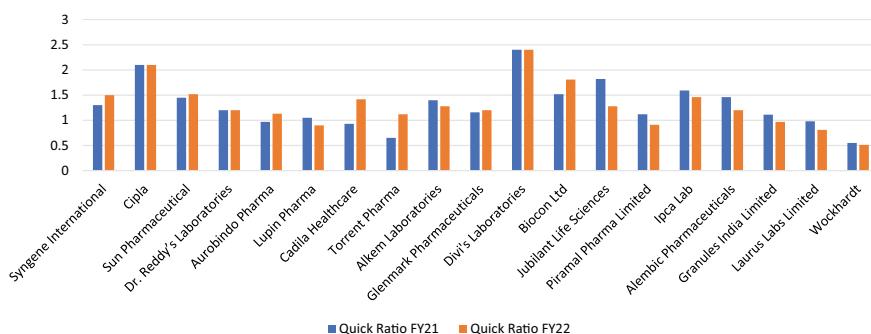


Fig. 7 Comparison of quick ratio (FY21 versus FY22)

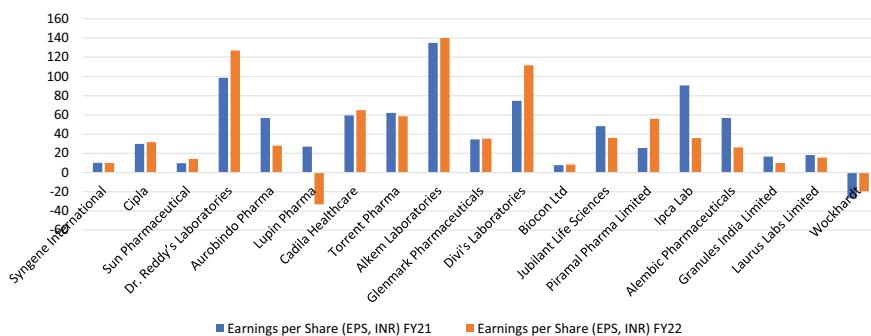


Fig. 8 Comparison of EPS (FY21 versus FY22)

4.2 SWOT Analysis

A SWOT analysis of India's pharmaceutical industry based on scholarly evidence reveals several internal strengths and external opportunities, along with weaknesses and threats affecting competitiveness.

Key strengths include India's diverse product mix encompassing generics, vaccines, biosimilars, herbals and patented drugs [3]; extensive domestic production capabilities in APIs, formulations and vaccines [26]; low-cost manufacturing skills and widespread technical workforce [27] and advanced R&D infrastructure across private companies, CROs and scientific institutes [28].

However, insufficient funding for innovation (less than 2% of sales), weak university-industry linkages, limited new molecular entity development and dependence on reverse engineering inhibit technological strengths [18, 20]. Complex regulatory systems, quality lapses, lack of harmonization with global standards and IP issues pose additional challenges [3, 29].

Key external opportunities include rising exports, domestic demand, healthy FDI inflows, increased healthcare spending and favorable government policies like tax incentives and clusters [1, 30]. However, pricing pressures, political conflicts affecting trade, regulatory hurdles in developed markets and supply chain disruptions are notable threats [31].

Advanced SWOT analyses by Reddy and Patnaikuni [32] and Bhatt and Modi [33] also highlight how India can leverage its resources and capabilities to overcome challenges. Sustained policy support, infrastructure upgrades, operational excellence, product innovation and talent development are vital for realizing the sector's potential.

4.3 Digital Adoption

The analysis indicates relatively low adoption of emerging technologies by Indian pharmaceutical firms compared with global counterparts, though investments are rising. Indian pharma allocated merely 1.7% of total revenues to digital transformation versus the global average of 3.2% [34]. Leading technologies implemented are cloud computing, analytics, mobility, artificial intelligence, automation and Internet of Things across functions spanning manufacturing, supply chain, clinical trials, sales/marketing and compliance [21].

Key factors impeding technology absorption include legacy IT systems, data security concerns, inadequate digital skills and pressure on margins [35]. Critics argue excessive digitalization could dehumanize healthcare and compromise privacy. However, per McKinsey [36] estimates, full-scale digital transformation has the potential to generate up to \$28 billion in value for Indian pharma over 5–7 years. Targeted adoption across prioritized use cases can enhance productivity, quality, compliance and patient experience.

4.4 Strategic Priorities and Growth Outlook

An analysis of industry patterns and academic studies indicates India's pharmaceutical market leadership is built on generics, vaccines and OTC drugs production along with CRO/CMO services for global clients. However, future growth necessitates expanding higher-value innovative offerings such as patented drugs, NCEs, biologics and drug delivery systems through dedicated R&D [23, 31]. Developing healthcare infrastructure and insurance coverage domestically while meeting regulatory compliance in exports markets remain imperatives for continued success [18].

Sector revenues are projected to reach \$120–130 billion by 2030, making India a top-3 global pharmaceutical market [1]. This will be fueled by rising exports, public health investments, insurers and domestic consumption of medicines as incomes grow. To realize projections, targeted strategies based on product innovation, operational excellence, talent development, quality focus, IP management, sustainability practices and strategic alliances are essential [26, 27].

5 Conclusions and Recommendations

This research offers valuable insights into India's pharmaceutical sector, which has witnessed remarkable growth to become the global generic drugs leader and a dominant emerging market force. However, the analysis reveals persistent challenges around infrastructure, regulation, innovation, ethics and digitalization that must be addressed for the industry to realize its full potential.

While India has cost innovation skills in generics and vaccine manufacturing, the sector must channel greater focus and funding into new drug discovery of patented molecules, biosimilars and biologics to ascend the pharmaceutical value chain. Academic-industry linkages and public R&D programs in specialized areas can support these efforts. Additionally, Indian pharma must continue strengthening its quality and compliance frameworks to prove capabilities as a dependable global supplier.

The growth outlook remains optimistic, buoyed by rising exports, healthcare access, insurance coverage and local demand. Realizing 2030 projections of the sector reaching \$130 billion will require integrated efforts between industry, government and academia. Policy reforms must promote innovation, streamline regulatory systems, deepen technical skills, upgrade logistics and expand healthcare access.

Digital adoption should become an urgent priority to harness technologies like automation, analytics and AI to enhance productivity, quality and compliance. However, this requires strategic investments in reskilling employees, modernizing legacy systems and collaborating with tech startups. Compliance with ethical practices and CSR commitments also needs monitoring to build public trust.

In conclusion, this study provides evidence-based pathways for the Indian pharmaceutical sector to attain global leadership. A collaborative approach between public

and private stakeholders is vital for shaping an enabling policy ecosystem, business environment and workforce capabilities. The recommendations offer actionable solutions for industry executives, policymakers and academia to jointly transform India's pharmaceutical aspirations into impactful realities.

References

1. India Brand Equity Foundation (2021). Pharmaceutical Industry in India. IBEF Report
2. Akash D, Charani H (2014) Indian pharmaceutical industry: challenges and opportunities. Indian Institute of Management Bangalore. <https://repository.iimb.ac.in/handle/2074/20375>
3. Gulaldavar S (2019) Envisioning the challenges of the pharmaceutical sector in the Indian market: a qualitative study. *J Bus Ind Mark* 36(11):1639–1654. <https://doi.org/10.1108/JBIM-07-2020-0365>
4. Pharmaceuticals Export Promotion Council (2022) Indian Pharma Industry
5. India Brand Equity Foundation (2022) Pharmaceutical Industry in India. IBEF Report
6. Ministry of Chemicals and Fertilizers (n.d.) Indian Pharmaceutical Industry. Government of India
7. Business Today (2022). Indian pharma exports log 8% jump in Q1 to \$6.26 billion. <https://www.bustoday.in/latest/economy/story/indian-pharma-exports-log-8-jump-in-q1-to-626-billion-343236-2022-07-31>
8. The Economic Times (2021) Indian pharma supplies over 20% of global generic drugs, says report
9. Chakraborty S, Chakraborty S (2020) Medicines in India: accessibility, affordability and quality. Brookings India. <https://www.brookings.edu/articles/medicines-in-india-accessibility-affordability-and-quality/>
10. Aayog N (2021) Pharmaceutical innovation: securing India's future
11. Deloitte (2021) Digital transformation in MedTech: key trends in 2021
12. Ming LY, Omain SZB, Kowang TO (2021) Supply chain resilience: a review and research direction. *Int J Acad Res Bus Soc Sci* 11(12):2591–2603
13. McKinsey & Company (2019) Building sustainable pharma value chains
14. Department of Pharmaceuticals, Ministry of Chemicals and Fertilizers, Government of India. Indian Pharmaceutical Sector—Current Status. 2022. Available online: https://pharmexcil.com/uploadfile/ufiles/1627510188_Indianpharmasector_currentstatus.pdf. Accessed 8 Aug 2023
15. Chaudhuri S (2005) Globalization, WTO, and the Indian pharmaceutical industry. *J World Intellect Property* 8(1):5–22
16. Chaudhuri S (2019) Indian patent law and its impact on the pharmaceutical industry: What can China learn from India? In: Liu KC, Chang C (eds) Intellectual property rights and the development of industrial clusters in Asia. Springer, Singapore, pp 227–248
17. KPMG (2022) Special economic zone scheme in India: a hit-and-miss. KPMG India
18. Chaudhuri S (2019) The Indian pharmaceutical industry—the way forward. Indian Pharmaceutical Alliance
19. Dhar B, Joseph RK (2019) The challenges, opportunities and performance of the Indian pharmaceutical industry post-TRIPS. In: Liu KC, Racherla U (eds) Innovation, economic development, and intellectual property in India and China. Springer, Singapore, pp 299–323
20. Govindan K, Shankar KM, Kannan D (2021) Emerging pharmaceutical companies from China, India, and Brazil: moving up the value chain in pharmaceutical R&D. In: Govindan K, Shankar KM, Kannan D (eds) Innovation from emerging markets. Cambridge University Press, Cambridge, pp 163–191
21. Jimenez D (2021) How technology could transform drug research in 2022. *Pharm Technol*

22. FICCI (2021) Pharma & Life Sciences
23. Singh S, Singh J (2021) Biosimilars pharmaceutical market in India: current status, challenges and future perspective. *Bull Biotechnol Biochem Res* 14(1):41–48
24. Agarwal A, Garg S, Gupta A (2020) COVID-19 and pharma supply chain resilience. Infosys Knowledge Institute
25. Dadhich A (2020) The COVID-19 pandemic and the Indian pharmaceutical industry. *Eur Pharm Rev*
26. Indian Pharmaceutical Alliance (2019) The Indian pharmaceutical industry—the way forward
27. The Washington Post (2016) India's low pharma costs are good for drug companies, good for consumers
28. Festa G, Kolte A, Carli MR, Rossi M (2021) Envisioning the challenges of the pharmaceutical sector in the Indian context. *J Bus Ind Mark* 37(8):1662–1674
29. The Challenges, Opportunities and Performance of the Indian Pharmaceutical Industry Post-TRIPS. In: Chaudhuri S (eds) *The WTO and India's pharmaceuticals industry*. India Studies in Business and Economics. Springer, Singapore
30. Research and Markets (2021) *India Pharmaceutical and Healthcare Industry Report 2021*
31. UK Essays (2018) SWOT analysis Indian pharmaceutical industry
32. Reddy KS, Patnaikuni I (2023) An overview of the SWOT analysis in India's pharmaceutical supply chain. *Asian J Glob Sustain* 1(1):1–18
33. Bhatt P, Modi P (2020) Envisioning the challenges of the pharmaceutical sector in the Indian market: a SWOT analysis. *J Bus Ind Mark* 36(7):1149–1165
34. IIIDE (2019) Role of digital transformation in pharma industry with its need & importance. <https://iide.co/blog/digital-transformation-in-pharma-industry/>
35. Financial Express (2021) Technology trends that are re-shaping the pharmaceutical supply chain in India. <https://www.financialexpress.com/healthcare/healthtech/technology-trends-that-are-re-shaping-the-pharmaceutical-supply-chain-in-india/2366237/>
36. McKinsey (2015) The road to digital success in pharma. https://www.mckinsey.com/capabilities/operations/how-we-help-clients/innovation-and-learning-centers/how-we-deliver-impact?gclid=Cj0KCQjwmICoBhDxARIsABXkXIISQ4C4NZ2m0_pjLXOWL-Vmmoopp1Y_Q9P_iRbmRa79HQYca6ICDtCIAmjYEALw_wcB&gclsrc=aw.ds
37. Biswas R, Pal S, Sarkar B, Chakrabarty A (2020) Health-care paradigm and classification in IoT ecosystem using big data analytics: an analytical survey. In: Solanki V, Hoang M, Lu Z, Patnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing, vol 1125. Springer, Singapore. https://doi.org/10.1007/978-981-15-2780-7_30
38. Jeyalakshmi, S., Akila, D., Padmapriya, D., Suseendran, G., Pal, S (2021) Human facial expression based video retrieval with query video using EBCOT and MLP. In: Peng SL, Hao RX, Pal S (eds) Proceedings of first international conference on mathematical modeling and computational science. Advances in intelligent systems and computing, vol 1292. Springer, Singapore. https://doi.org/10.1007/978-981-33-4389-4_16
39. Singh D, Sahana S, Pal S, Nath I, Bhattacharyya S (2020) Assessment of the heart disease using soft computing methodology. In: Solanki V, Hoang M, Lu Z, Patnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing, vol 1125. Springer, Singapore. https://doi.org/10.1007/978-981-15-2780-7_1
40. Chakrabarty A, Tagiya M, Pal S, Cuong NHH (2020) Managing psychosomatic disorders related to obsession and addictions to gadgets through IoT surveillance. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_64
41. Biswas R, Pal S, Cuong NHH, Chakrabarty A (2020) A novel IoT-based approach towards diabetes prediction using big data. In: Solanki V, Hoang M, Lu Z, Patnaik P (eds) Intelligent computing in engineering. Advances in intelligent systems and computing, vol 1125. Springer, Singapore. https://doi.org/10.1007/978-981-15-2780-7_20
42. Suseendran G, Doss S, Pal S, Dey N, Quang Cuong T (2021) An approach on data visualization and data mining with regression analysis. Advances in intelligent systems and computing. 649–660. https://doi.org/10.1007/978-981-33-4389-4_59

43. Rakshit P, Nath I, Pal S (2020) Application of IoT in healthcare. In: Peng SL, Pal S, Huang L (eds) Principles of Internet of Things (IoT) ecosystem: insight paradigm. Intelligent systems reference library, vol 174. Springer, Cham. https://doi.org/10.1007/978-3-030-33596-0_10
44. Tagiya M, Sinha S, Pal S, Chakrabarty A (2020) Transformation from HRM inadequacy and bias-syndrome to transparent and integrated ecosystem through IoT-intervention in career management. In: Gunjan V, Garcia Diaz V, Cardona M, Solanki V, Sunitha K (eds) ICICCT 2019—system reliability, quality control, safety, maintenance and management. ICICCT 2019. Springer, Singapore. https://doi.org/10.1007/978-981-13-8461-5_61

Framework for Reverse Supply Chain Using Sustainable Return Policy



Tridha Bajaj, Snigdha Parashar, Tanupriya Choudhury, and Ketan Kotecha

Abstract Numerous items are used in daily life, resulting in millions of tonnes of packaging waste being generated daily, including cosmetic containers, cardboard packaging used by various e-commerce services, plastic bottles, carry bags, and so on. With a focus on sustainable development, researchers have been working tirelessly to reduce waste generated from the packaging sector by improving the conventional principle, and this project proposes the 4thR—RETURN, with an aim to develop a promising solution that would work on the fundamentals of the reverse supply chain, that is, it shall collect the reusable packaging from the consumer and deliver it to its manufacturing unit or the industry/warehouse where it can be reutilized. To achieve the project's goal, various technologies, such as multiple layer perceptron, multiple linear regression, and full stack development are used to build a gateway that can allow consumers to return the packaging in the same manner that many e-commerce websites accept product returns.

T. Bajaj · S. Parashar

School of Computer Science (SoCS), University of Petroleum and Energy Studies (UPES), Bidholi Campus, Dehradun, Uttarakhand 248007, India

T. Choudhury

CSE Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra 412115, India

SoCS, University of Petroleum and Energy Studies (UPES), Bidholi Campus, Dehradun, Uttarakhand 248007, India

K. Kotecha

Symbiosis Centre for Applied Artificial Intelligence, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra 411045, India
e-mail: director@sitpune.edu.in

Present Address:

T. Choudhury (✉)

CSE Dept., Graphic Era Deemed to be University, Dehradun 248002, Uttarakhand, India
e-mail: tanupriyachoudhury.cse@geu.ac.in; tanupriya.choudhury@sitpune.edu.in;
tanupriya1986@gmail.com

Keywords Waste · Plastic · Reserve supply chain · Return · Sustainable development

1 Introduction

One of the most contentious topics in today's fast-growing world is sustainable development. It is a comprehensive and multidimensional concept that seeks to make the world a better place for everyone while preserving possibilities for future generations. Sustainable development aims to present a balanced world to personify "One Common Future" of a resilient society which promotes human well-being as well as protecting the environment, hence encouraging economic prosperity. It is an extensive and long-term viewpoint which acts as a compass for people, groups, organizations, and governments as they work together to build a fairer, environmentally conscious, and economically impregnable society.

Environment-conscious businesses working towards the principles of sustainable development have been actively integrating the reverse supply chain model as a powerful conduit for an intersection of commerce and sustainability. The reverse supply chain model primarily focuses on managing returns, refurbishment, recycling and responsible disposal of items which can no longer be used, unlike the forward supply chain which is focused on the flow of products from the manufacturer to the consumer.

Conventionally, sustainable development theory relies on 3Rs, namely reduce, reuse, and recycle, they are critical components of the waste hierarchy, waste management, and sustainable development. These provide a guide to prioritize actions in dealing with various waste management techniques and promoting more sustainable resource management. The three Rs work together to reduce trash output and improve waste disposal. Three R's is generally referred to as "The Guiding Philosophy of the Waste Management Process." Here, this research has gone on to propose the 4th R-RETURN.

The essence of the 4Rs is as discussed below:

1. Reduce

Reducing is the best option. If humans reduce the utilization of unnecessary products, there shall be a lower chance of waste generation and hence pressure on the already exploited natural resources would be reduced. In short, the best way to handle it is to not produce it. For example:

- I. Avoid disposable cutlery.
- II. Avoid using single-use plastics.

2. Reuse

The idea behind the "Theory of Reuse" asserts that a product can be used again in its present state without requiring any energy to transform it into a brand-new object, that is, people can fix broken furniture and equipment as well as preserve

paper and plastic bags for future use. It focuses on utilizing the already available resources rather than procuring new products from the market. This process proves to be helpful in efficient resource optimization.

3. Recycle

Recycling converts waste into a resource, it uses energy to transform an already existing product into a new usable product. In other words, recycling is the process of taking up discarded materials, remanufacturing them, and then utilizing them as a new commodity. To support the idea of recycling one must buy products made from recycled material and at the same time collect recyclable items separately to be further sent for recycling.

4. Return

Since the research on the fourth R is still ongoing, this work proposes the 4th R to be “RETURN”.

Return follows the reverse supply chain model, in which all reusable items can be accepted back. It is the process of retrieving a used product from a client and either disposing of it efficiently or reusing it. For example, procuring back the cardboard packaging received by the customer when they purchase something from an e-commerce store.

This project’s portal shall operate as a bridge to collect items from the consumer, with the sole condition that items are not damaged that is they can be reused and return the collected items to the firm to assist in reducing the manufacturing of reusable products and sustainable growth. The portal may accept reusable packaging, reusable containers such as empty shampoo bottles, empty perfume bottles and other containers that the brands may reuse for packaging.

The research focused more on creating and exploring the core principles and functionality of a portal that may act as a medium for enforcing reverse supply chain for waste management to promote sustainable development.

2 Literature Survey

Researchers have been working relentlessly in the realms of waste handling, responsible resource management and sustainability. Several related works have been concluded so far outlining various elements on which we must focus to encourage sustainable development. Some of the previous studies closely related to this research work have been briefly addressed below.

1. “Scenarios study on post-consumer plastic packaging waste recycling—Finger-print” [1] indicates that the packaging waste, plastic waste, and waste recycling on the Earth is the highest and various logistics systems are used to determine them. This research demonstrated that plastic trash is the most prevalent and is continually wreaking havoc on the ecosystem, but other wastes are also equally

to blame for the deteriorating environmental conditions. Hence, it justified the need for a framework to promote reverse supply chain and the generous use of resources, keeping in mind the main motive of sustainable development as well.

2. Researchers justify that companies are working with consumers to reduce waste [2]. In a survey of 54 of the world's leading corporations (including Apple, Google, Coca-Cola, Nike, and Unilever), 84% stated that they had already adopted sustainable business practices or were actively working on doing so. Furthermore, many of these companies have seen an increase in sales as a result of their efforts. For example, Starbucks has reported an annual growth rate in its green tea business of 5%.

In line with this, sustainable product design has become increasingly popular among manufacturers and designers, who are now developing products that last longer and use fewer resources than their predecessors. Some real-time examples are as follows:

- (A) Smartphones: Many smartphones now come with features such as touch-screen displays, which can be used without having to charge them. These devices also have built-in sensors that constantly monitor their battery usage and shut down functions when they are not needed.
- (B) Clothes: Clothing that has been made out of recyclable material is better for the environment because it reduces the number of natural resources required to make clothing. Most stores now provide customers with information about the environmental impact of the clothes they sell.
- (C) Refrigerators: Manufacturers have started making refrigerators that run on solar power or require less electricity, and therefore consume fewer natural resources to operate.
- (D) Tablets and Computers: Devices such as tablets and computers are designed to be more durable, so they don't break easily. They are also easier to repair than previous models hence reducing the need for resources.

These examples showcase that there is an increasing desire among consumers for products that have a lower impact on the environment. As a result, manufacturers have begun designing products that are of higher quality, cheaper to manufacture, and sustainable.

3. The need for sustainable development is to stabilize the impact of constant urbanization [7]. Urban regions have traditionally been the centre of attraction, as evidenced by the fact that they account for around 80% of economic activity. The process of establishing and growing urban areas is known as urbanization. According to researchers, there will be around 300 new cities by 2030. The rate of urbanization has been a problem for all environmental operations since it is directly related to the amount of garbage produced. The growing rate of urbanization motivates individuals to seek a strategy for sustainable development. The implementation of a reverse supply chain might be one of those successful methods of sustainable development that not only reduces waste but also tends to boost profit for businesses.

4. BACK TO MAC

MAC, a well-known cosmetic business, promotes sustainable development by collecting empty containers from customers and awarding them with a complimentary lipstick in exchange for six empty containers [9]. The collected plastic containers are transported to subcontractors for recycling, making them environmentally beneficial. MAC created its takeback programme thirty years ago intending to recycle or recover as much of its black (Acrylonitrile Butadiene Styrene) plastic packages as possible. As a part of the programme, the company works with leading recycling partners around the world to recycle its Acrylonitrile Butadiene Styrene plastic into new plastic, which goes into the production of new materials for things like coffee machines, televisions, office supplies, or electronic devices.

5. Kabadiwala

It is a waste management firm established in Bhopal that collects rubbish from users and sends it for recycling, much to traditional scrap collectors. The company's mission is to simplify the unstructured sector of garbage collection, and it is now one of the largest organizations operating in this field [10].

They are working tirelessly to make recycling accessible to individuals and institutions alike.

6. Trash to Cash

This model works on the principle of collecting trash also referred to as waste from households and giving them compensation in the form of cash. At home during the lockdown period, the founder of the company considered assisting elderly parents, housewives, and others who had been struck in lockdown for months at a time, accumulating massive amounts of trash as people were afraid of the deadly virus attack and were unable to go out to clear the trash, which motivated the entrepreneur to start a social venture named "TRASH TO CASH" [11] to assist those in need by collecting garbage from door to door. More importantly, he thought about how women could become independent, make money, and avoid danger by sorting household rubbish.

7. Flipkart's 2019 project announcement about the collection of their plastic packaging from customers increases the scope of this project for a collection of the packaging from the consumer and delivering it to the organization for reuse [12].
8. "Myntra eliminates 100% single-use plastic packaging delivering a plastic free-edition" [13].

Myntra has pioneered scalable and sustainable packaging and delivery alternatives such as replacing poly pouches with recycled paper bags, bubble wraps with carton waste shredded material, and plastic adhesive tapes with paper tapes, replacing poly bags used for attaching customer invoices with recycled Kraft paper pouches, and RFID-tagged multi-use polyester bags piloted for reverse shipments as packaging materials. As of September 2021, this action has already resulted in the diversion of 670 tonnes of plastic.

9. Apart from Gpay's beautiful UI and elegant design, the application is intriguing because of the rewards scheme, which allows users to earn cashback depending on a specific scratch card rather than a predetermined sum [14]. The reward is determined by the number of days since the last transaction, not by the amount of the transaction. It absolutely makes sense to reward users who engage more often with the application rather than just focusing on the amount of the transaction.

All of the works listed above demonstrate the need for an effective and efficient reserve supply chain framework to mitigate the effect on the environment and foster a culture of sustainable development among the masses. The conclusion can be drawn that plastic and other harmful packaging waste have been deteriorating our environment, and the manufacturers and consumers have accepted the fact that it is a collective responsibility to come forward with innovative plans, platforms, schemes, and strategies to save the environment from exploitation.

3 Existing System

The world is working towards sustainable development to ensure the utilization of resources in a way that even future generations benefit from them. To combat the increase in solid waste production, the 3R system, which is depicted in Fig. 1, has been conventionally in place for a long time. However, the system is incapable of catering to the recent exponential increase in waste production due to various factors such as use-and-throw tendencies, fast-evolving technology, change is inevitable, and many more.

First and foremost, reuse aims to store the product and use it differently in future. It focuses on the utilization of the products already available rather than acquiring new products from the marketplace. For example, packaging from different brands and empty containers that are generally considered useless are thrown away, but storing them and using them differently can help the environment.

Furthermore, the reduce strategy aims to minimize the generation of waste by cutting down the usage of products which can have a better alternative. For example, reducing the use of single-use plastic by opting for reusable water bottles over plastic bottles, jute bags for shopping over plastic or paper bags, and metal or glass straws over plastic straws.

Lastly, recycling aims to reincarnate products that can be remoulded. It claims that must put in some extra effort to separately store recyclable items which can be sent together to a recycling centre and at the same time people should prefer products made from recyclable material over new products.

For example, “Waste Warriors” is an organization that recycles all perishable items by streamlining them and transforming them into new useful products such as chairs and tables.

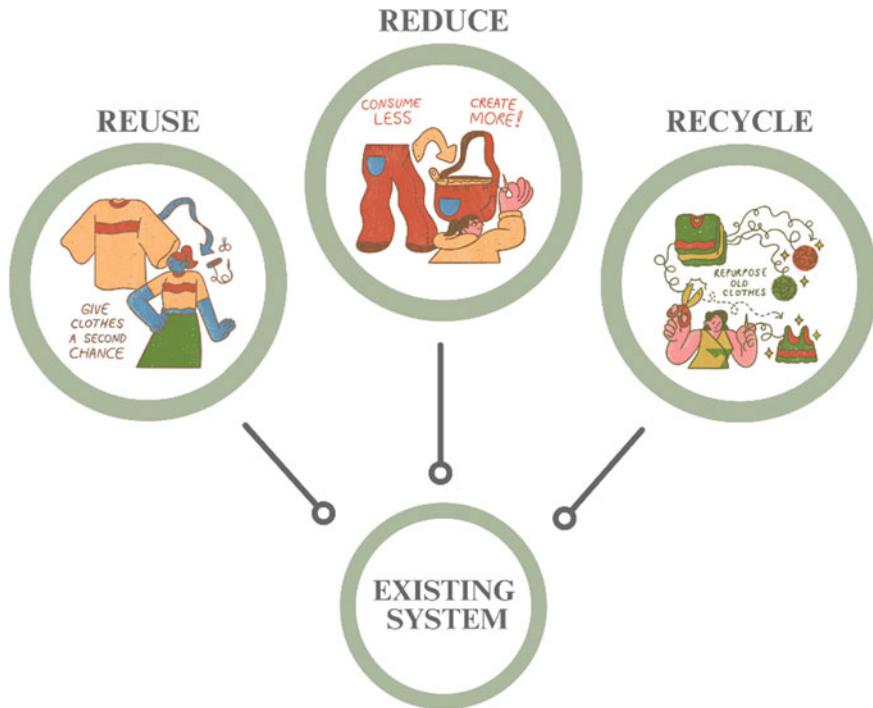


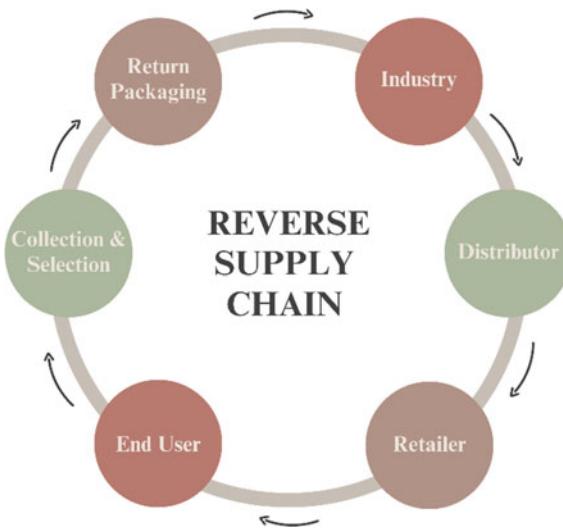
Fig. 1 Flow of the existing 3R system

4 Proposed System

In this era of industrialization, the use-and-throw tendency rules the world. This project aims to promote knowledge and awareness about sustainable development, along with a solution which shall help inculcate eco-friendly practices in the citizens. Products come in various packaging; after use, the packaging is discarded even if it is still intact. The primary purpose is to apply a reverse supply chain model to collect such packaging including e-commerce packaging, cosmetic packaging, home item packaging, and so on and further return the reusable collected packaging to the original manufacturer or the utilizing brand. The proposed reverse supply chain is shown in Fig. 2. This is a project that ensures commitment to a sustainable future.

This project intends to improve the environment, as it shall play the role of a catalyst to reduce the carbon footprint by motivating human beings to return the packaging that can be reused by the organization. The consumer returning the packaging would earn rewards and take pride in witnessing the amount of carbon footprint reduced by their efforts. This shall also benefit the organization as the packaging production cost would decrease.

Fig. 2 Proposed flow of reverse supply chain



The objective is to assist the environment on a broad scale, as well as consumers and producers by creating a platform for users to simply return the packaging they get daily, which can be repurposed by organizations.

5 Algorithm Used

Two machine learning algorithms [5] have been deployed in the backend corresponding to the key two capabilities, namely carbon footprint prediction and reward prediction [3]. Multiple linear regression is the machine learning approach [6] used for carbon footprint prediction, whereas multiple layer perceptron is used for optimal reward prediction. Both algorithms are critical to the operation of the portal based on the reverse supply chain model. The details of both algorithms are discussed below.

5.1 *Multiple Linear Regression for Carbon Footprint Prediction*

This study employs multiple linear regression to accurately forecast carbon footprint. The following are some notable features of multiple linear regression:

- It is a supervised machine learning [8] algorithm where a continuous value is expected as an output with multivariate inputs.
- The datasets for this model were split into training and testing datasets according to the standard split of 80% training data and 20% testing data.

- (c) A training dataset was used for training the model whereas the testing data was used to check the accuracy of training.
- (d) In multiple linear regression, the concept of multivariate inputs and a single output was used but it is necessary to analyze what features were necessary for training and which features can be dropped.
- (e) The basic equation for multiple linear regression is $y = b_0 + b_1x_1 + b_2x_2 \dots b_nx_n$.
- (f) The accuracy was then calculated and many different types of metrics to optimize the training.
- (g) For multiple linear regression, the adjusted R^2 proved to be the best metric to calculate the accuracy of the model.

5.2 *Multiple Layer Perceptron for Reward Prediction*

The following attributes of a multiple layer perceptron are used to forecast rewards:

- (a) An artificial neural network is a computational nonlinear model that is inspired by the brain cells, i.e. neurons.
- (b) Like people, ANN is trained/learned by example.
- (c) ANN consists of a large collection of artificial neurons or processing elements which operate in parallel.
- (d) Each connection link is connected with a weight which has information about the input signal.
- (e) Every neuron has weighted inputs (synapse), an activation function and one output.
- (f) Multiple layer perceptron is one of the complex and density-connected artificial neural networks where multiple layers of perceptron are present.
- (g) In the multiple layer perceptron model, each input is considered as a node for that particular layer and has a single output node which can have as many nodes in the hidden layer as shown in Fig. 3.

5.3 *Characteristics of Data*

An extensive dataset was employed in this study. The dataset was created manually in order to create a multiple linear regression model for carbon footprint prediction and a multiple layer perceptron for reward prediction. This dataset is critical in assisting with sustainable development efforts and guiding decision-making in carbon emission reduction plans.

(I) Dataset for Carbon Footprint Prediction—Multiple Linear Regression Model

Characteristics of the dataset used for the multiple linear regression model are as follows:

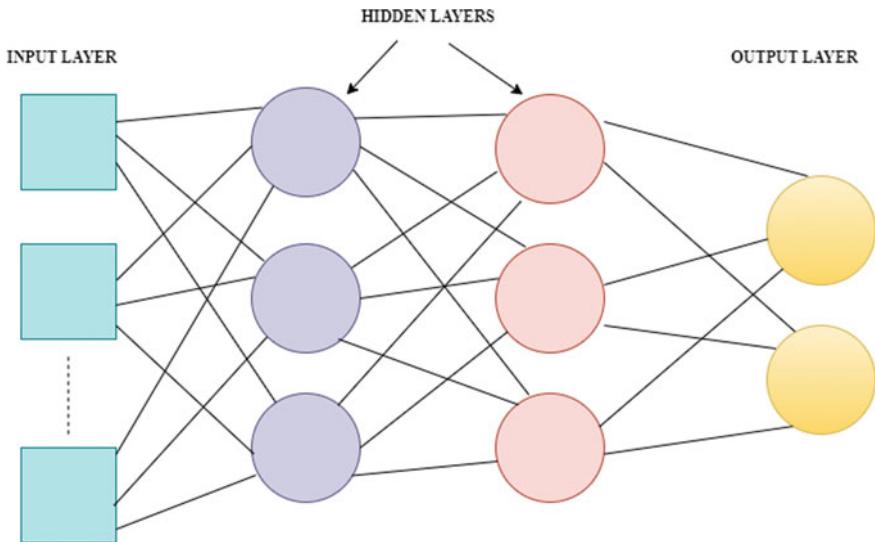


Fig. 3 Architecture of multiple layer perceptron

- (a) The dataset was manually created [4] by analyzing different types of features that were required for carbon footprint prediction.
 - (b) The dataset had three different input labels [Company Name, Product, Type of Packaging] and one output label which is carbon footprint (CO2eq).
 - (c) By analyzing the different labels, the inferences about the p-values for these labels were understood and Company Name was found to have the maximum p-value, i.e. 0.9807 which meant this label was not suitable for model creation.
 - (d) Since the dataset consists of characters it was important to convert them into numbers which can be done using one hot encoding.
 - (e) The input to the multiple linear regression model was an array where different products and types of packaging are present and the output was an array of values representing carbon footprint.
- (II) Dataset for Reward Prediction—Multiple Layer Perceptron

Characteristics of the dataset used for the multiple layer perceptron model are as follows:

- (a) The dataset was created manually by analyzing different types of features that were required for reward prediction.
- (b) The dataset had two different input labels [Item, Quantity] and one output label which was Reward Points.
- (c) A straightforward three-layer perceptron was built using this dataset by examining the various labels in the input set, and correspondingly output is determined on those labels.

6 Design

The unified modelling diagram in Fig. 4 depicts the operation of the portal named *PackDrop*, as well as the process of interaction between the user and the user interface. The user is expected to raise a query to return a particular packaging item. The query entails learning about the return packaging, reward system, and carbon footprint. The query will subsequently be processed, and the user will learn more about the portal. Based on the request, the portal will decide on the reward scheme and points to be shared with the user, as well as provide the estimated amount of carbon footprint that would have been discharged into the environment if those things had been discarded carelessly. Figure 4 seeks to provide a full visual explanation of how the reverse supply chain would function in terms of user interface and experience.

Furthermore, the portal's numerous functionalities sharpen user engagement while also delivering several instructive and beneficial aspects for the user to continue with and reconsider their non-sustainable actions. The following are some of the primary processes engaged on the portal:

A. Cart System

The cart system integrated into the graphical user interface enables the user to generate a return request for the packaging they would like to return.

B. Carbon Footprint Prediction

When consumers return some packaging, they shall receive information about the carbon footprint prevented from being emitted. This will instil in them a sense of

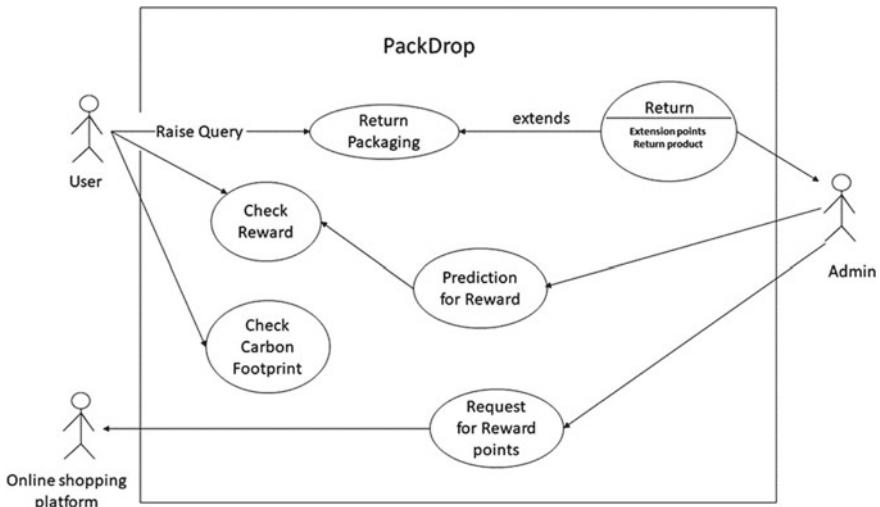


Fig. 4 Use case diagram of the portal

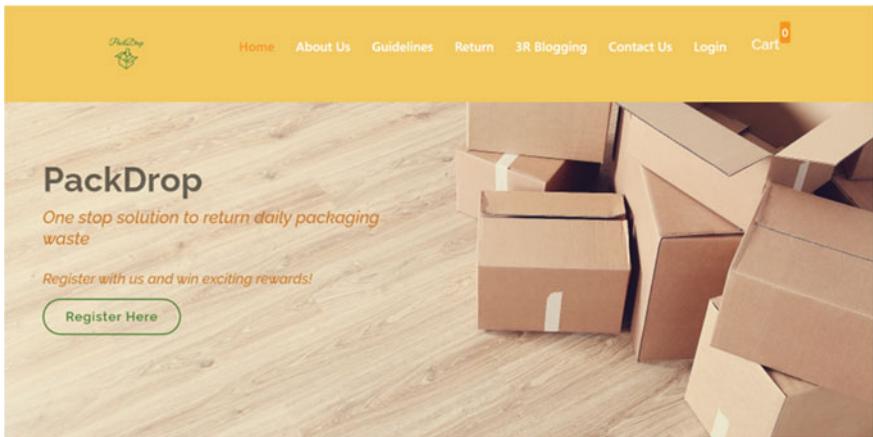


Fig. 5 Snapshot of the working user interface developed as the framework for the reverse supply chain model

pride which will give them an idea about the impact they are making in the field of sustainable development and further inspire them to contribute more.

C. Reward Prediction

The portal will inform customers about the expected reward for returning a collection of selected packaging.

A universal user class exists for the proposed portal as everyone will have access to it. People who frequently shop online and who are passionate about sustainable development will be the ones who are more likely to explore the information and tools offered on the platform.

Various interfaces were utilized to present the final look of the portal as displayed in Fig. 5. It includes the user interface, software interface, and database interface as detailed below:

A. User Interface

The project's front end was built using HTML, CSS, and JavaScript. HTML was utilized to create the basic outline of the webpage; CSS assisted in making the web page more pleasant, and JavaScript assisted in improving the user experience.

B. Software Interface

Portal has achieved the required functionality with Django being in the backend along with the frontend technologies, HTML, CSS, and JavaScript integrated with the machine learning models to provide efficient software.

C. Database Interface

Firebase was incorporated to handle the portal's data. The Firebase Realtime Database is a cloud-hosted database that stores data in JSON format. The Firebase

Realtime Database is a NoSQL database that allows us to store and sync data in real-time among the users. It's a large JSON object that developers may handle in real-time. The Firebase database provides the application with the current value of the data as well as modifications to that data via a single API. Consumers may access their data from any device, online or mobile, thanks to the real-time synchronization of Firebase.

Final Output

See Fig. 5.

7 Results and Discussion

The provided model fulfils the requirements of a user-friendly and efficient system for returns using a cart system. The system leverages the following results:

- (1) The product that the user opts to return using the cart system gets reflected in the cart: When a user uses the cart to return any product, it is made sure that the change is reflected in the cart properly that is, the addition of the product, or removal of the product from the cart is done properly and the reward policy is executed accurately.
- (2) Based on the items that the user wants to return they get an approximate number of units of carbon footprint they have prevented from being emitted into the environment, that is the model fulfils its sustainable development goal. The system first calculates the number of units of carbon footprints that have been prevented from emission and then the accurate results are presented in front of the user.
- (3) While returning items, users expect some brownie points in return; hence, the reward prediction system shall help them get an idea about how many reward points they can expect on the return as when a user raises a return product request, the model utilizes the multiple layer perceptron model deployed in the backend to predict the desired reward points. This prediction is purely based on the dataset used for the reward prediction model. The dataset considers features like the type of item and the quantity returned to further predict the reward points. This encourages the user to engage and contribute more.

Overall, the model seamlessly combines the cart system with the sustainable development agenda of the whole project. When a user returns a product, based on the type and quantity they get eligible for different rewards which are predicted using the reward prediction model of the project. Additionally, the user will receive an output from the carbon footprint model that shall display the number of units of carbon emissions they have avoided while returning the item. This holistic approach not only enhances the user experience but also encourages people to apply sustainable development motives more enthusiastically.

8 Conclusion

Witnessing a lot of packaging going to waste when it could be easily reused if returned to the company activated the need for a platform that can act as an intermediary to facilitate a reverse supply chain that allows consumers to return reusable packaging to the brand for sustainable development and earn brownie points in the form of various reward options such as free goodies, vouchers, cashback, and so on. This project promoted sustainable development and helped users understand how much carbon they can avoid and their contribution towards the environment.

The portal can be helpful for everyone as people keep on collecting bottles and containers that are of no use and at the same time increase waste in households. This project has the potential to generate employment as this process shall give rise to the need for delivery boys to collect waste from numerous homes and corporations. The reward prediction system will attract a humungous amount of people to contribute towards the environment by returning reusable products and hence reduce the emission of carbon significantly. By gathering recyclable packaging and containers, this initiative may prove useful for businesses in lowering their manufacturing costs, giving them an incentive to join forces on the project.

9 Future Scope

A framework to aid the reverse supply chain was highly anticipated in the field of sustainable development. This work successfully delivers a portal with the fundamental requirements for the reverse supply chain model. At the same time, recognizes future advancements that may help to present an improved user engagement of the platform developed in this project, as well as better sustainable planning of waste material produced. Some of the feasible extensions include the following:

1. Dataset optimization

The dataset being used for this work currently relies on quantity, while the ultimate goal is to optimize it by developing or figuring out an appropriate time-series dataset. For example, if the user returns three bottles of a plastic container, then the reward would be six points or so as per the current model. But in future, the dataset can be based on how frequently the user is placing the return request. The more someone returns the better rewards should be given.

2. RNN-LSTM model instead of ANN

The eventual plan of this project must be the utilization of RNN-LSTM model, i.e. recurrent neural network—long short-term memory model, in the reward prediction model instead of the simple artificial neural network as recurrent neural network takes the input based on the previous output. Hence, RNN would help extensively for a better reward prediction system, whereas LSTM shall help in learning the

previous records of the particular user as history would be stored in the memory. Therefore, the RNN-LSTM model for reward prediction would be easier, effective, and efficient. As of now the dataset for the reward prediction system only has a few constraints but when the recurrent neural network—long short-term memory (RNN-LSTM) model would be used then more constraints can be added as it will be based on the time-series dataset.

3. User model integration—enable multiple user usage

The portal is currently utilized to meet the needs of a single user, but it can be made to accommodate multiple users by incorporating the technology employed by other e-commerce websites like Myntra, Amazon, and Flipkart. Other websites adjust their content to the user's preferences. For the login or register options, which are available from Facebook, email IDs, and Google accounts, the portal is right now integrated with the Firebase database which can also be changed in future to cater to the advancements and the needs of the portal.

4. Introducing a recycling option for perishable item packaging

A recycling option can be given to the users to streamline the recycling of perishable items as well by collaborating with numerous organizations working in the same direction. They can collect all the waste materials such as packets of food items, plastic bottles, containers, etc., through the proposed portal and recycle them into useful items such as chairs, tables, pen stands, etc. Hence, by collaborating with such organizations, the portal can add an option of recycling waste materials, which can further help in promoting sustainable development.

Acknowledgements The generous backing for this research came in the form of a Research Support Fund (RSF) Grant from Symbiosis International (Deemed University), Pune, India.

References

1. Thoden van Velzen EU, Bos-Brouwers HEJ, Groot JJ, Bing Xiaoyun X, Jansen M, Luijsterburg B (2013) Scenarios study on post-consumer plastic packaging waste recycling. (Rapport/ Wageningen UR Food & Biobased Research; No. nr. 1408). Wageningen UR—Food & Biobased Research. <https://edepot.wur.nl/260432>
2. “Companies Are Working with Consumers to Reduce Waste.” Harvard Business Review, 7 June 2016, hbr.org/2016/06/companies-are-working-with-consumers-to-reduce-waste
3. Taneja S, Garg D, Tarun Kumar MV, Choudhury T (2018) The machine predicted market. In: 2018 international conference on computational techniques, electronics and mechanical systems (CTEMS), Belgaum, India, 2018, pp 256–260. <https://doi.org/10.1109/CTEMS.2018.8769306>
4. Gaikwad M, Ahirrao S, Phansalkar S, Kotecha K (2021) Online extremism detection: a systematic literature review with emphasis on datasets, classification techniques, validation methods, and tools. IEEE Access 9:48364–48404. <https://doi.org/10.1109/ACCESS.2021.3068313>
5. Skala V, Singh TP, Choudhury T, Tomar R, Bashar MA (2022) Machine intelligence and data science applications. Springer Nature

6. Singh S, Anand A, Mukherjee S, Choudhury T (2022) Machine learning applications in decision intelligence analytics. In: Jeyanthi PM, Choudhury T, Hack-Polay D, Singh TP, Abujar S (eds) Decision intelligence analytics and the implementation of strategic business management. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-82763-2_15
7. Gue IHV, Lopez NSA, Chiu ASF, Ubando AT, Tan RR (2022) Predicting waste management system performance from city and country attributes. *J Clean Prod* 366:132951, ISSN 0959-6526. <https://doi.org/10.1016/j.jclepro.2022.132951>. <https://www.sciencedirect.com/science/article/pii/S0959652622025434>
8. Harish V, Mansurali A, Choudhury T (2023) Data analytics in operation management. In: Ramdane-Cherif A, Singh TP, Tomar R, Choudhury T, Um JS (eds) Machine intelligence and data science applications. MIDAS 2022. Algorithms for intelligent systems. Springer, Singapore. https://doi.org/10.1007/978-981-99-1620-7_6
9. MAC Cosmetics Expands Packaging Takeback Program|Beauty Packaging. (2022). Beauty Packaging. Retrieved October 10, 2022, from https://www.beautypackaging.com/contents/view_breaking-news/2022-03-22/mac-cosmetics-expands-packaging-takeback-program
10. “The KabadiwalaSell Scrap Online.” The KabadiwalaSell Scrap Online, www.thekabadiwala.com
11. HOME|TRASH TO CASH (n.d.) TRASH TO CASH. Retrieved October 10, 2022, from <https://www.trashtocash.co.in/>
12. Flipkart announces new project to collect plastic packaging from customers|The News Minute. (2019). The News Minute. Retrieved October 10, 2022, from <https://www.thenewsminute.com/article/flipkart-announces-new-project-collect-plastic-packaging-customers-112309>
13. <https://www.apnnews.com/myntra-eliminates-100-single-use-plastic-packaging-delivering-a-plastic-free-edition-of-its-ongoing-big-fashion-festival/> (2022) Beauty packaging. Retrieved October 10, 2022
14. Google Pay Gamification (Product Management)—Startup Analytics (2021) Startup analytics. Retrieved October 10, 2022, from <https://startapanalytics.in/products-tez-aka-google-pay/>

Sentiment Analysis Survey Using Deep Learning Techniques



Neha Singh, Umesh Chandra Jaiswal, and Jyoti Srivastava

Abstract This study focuses on deep learning methodologies as it examines the spectrum of sentiment analysis techniques. Identifying the emotional tone or attitude reflected in textual information is called sentiment examination, an important component of natural language processing. A subset of machine learning called deep learning has completely changed the discipline by making it possible to build sophisticated models that can accurately capture complicated linguistic patterns. This study provides a thorough analysis of deep learning methods applied to opinion analysis. In this research, we examine and classify a variety of deep learning methods for sentiment analysis. We also examine current issues and potential developments in sentiment analysis. We also covered assessment metrics, which are frequently used for developing and assessing sentiment analysis models.

Keywords Natural language processing · Deep learning · Machine learning · Sentiment analysis

1 Introduction

The management of feelings, viewpoints, and subjective language is known as sentiment analysis (SA). Sentiment analysis, which examines various tweets and reviews, gives comprehensive information about public views. It is a tried-and-true method for forecasting a wide range of important occasions, including the success of films at the box office and presidential elections. On many websites like Yelp and Amazon, you

N. Singh (✉) · U. C. Jaiswal · J. Srivastava
Madan Mohan Malviya University of Technology, Gorakhpur, India
e-mail: nehaps2703@gmail.com

U. C. Jaiswal
e-mail: ucjitca@mmmut.ac.in

J. Srivastava
e-mail: sriv.jyoti1996@gmail.com

may find public feedback that is meant to assess a specific object, such as a person, thing, or location. The examination might be classified as either positive, negative, or neutral. The goal of SA is to automatically identify the expressive direction of user reviews. The demand for SA has increased due to the growing requirement to analyse and examine unstructured data obtained from social media [1].

The level of the textual target being studied, such as the document, sentence, or feature can be used to categorise the SA. The opinion polarity for big textual chunks, such as an entire paragraph or document, is provided by document SA. The polarity of each sentence depends on the sentence level. SA at the aspect level determines the polarity of a number of the sentence's provided aspects. The distinctive characteristics of the entity or situation are its aspects or objectives. To comprehend the numerous targets and aspects of the product or service, as well as the tweets, comments, posts, chats, and messages, aspect level SA is utilised. A nice example would be the ambiance, however the room service is subpar. The same subject is stated twice in this passage. A significantly more in-depth report is produced by aspect level analysis. With the help of these insights, business intelligence has a huge chance to gather data on how customers are reacting to a product or service. The conclusions drawn from the digital reviews are used by business organisations to make innovative decisions [2].

Many academics are utilising deep learning to address SA, and it significantly affects both supervised and unsupervised learning. It includes a number of well-known and useful models, and these models are used to solve a range of issues. Deep learning has great promise for high frequency stock market prediction since it can extract characteristics from a vast amount of raw data without depending on past predictor knowledge. The approaches differ significantly in terms of the structure of the network, activation function, and various other model variables, and the performance of the methods is strongly influenced by the data representation. The benefits and limitations of deep learning techniques for stock market prediction and analysis have both been studied by researchers. The ability to include a vast array of data into investment selection and portfolio creation is one of the main benefits of employing deep learning in finance. A smaller feature space can be created by compressing the information. Research has shown that deep learning techniques' nonlinear feature reduction is successful at predicting price trends. Natural language ambiguity and complexity may be addressed by deep learning in ways that classic text mining techniques cannot [3].

1.1 *Sentiment Analysis Procedure*

Sentiment analysis procedure is categorised into four parts: data collection, pre-processing, classification, and display result.

- *Data Collection:* The work of data collection comes first in the sentiment analysis procedure. Any website can provide information, as well as the numerous online databases of user thoughts and evaluations [4].
- *Preprocessing:* Data cleansing is being done at this point. The use of unnecessary words and symbols is eliminated. To make future processing easier, this is done. This stage includes removing hyperlinks, repetitious sentences, emoticons, and special characters. There is also stemming and lemmatization. Last but not least, the classifier is given a condensed set of features.
- *Classification:* The most crucial element of a framework for sentiment analysis is a classifier. Classification is done into negative, positive, or neutral categories. The training set for the classifier is typically one-third of the database. The training set has a significant impact on the classifier's accuracy. SVM, Bayes, maximum entropy, and other machine learning classifiers, among others, can be used for classification. Deep neural networks may also be used to classify the data, but for machine learning classifiers, feature extraction is done before the classifier is trained and graded. Classification can be done easily with a deep learning system.
- *Display Results:* The outcome is shown following the data's passage through the classifier. The kind of classifier used dictates the level of detail provided and the polarity of the emotions throughout the entire dataset [4].

The paper also addresses the research and its findings in part II. Section III of the essay will analyse the methods. Section IV discusses evaluation measures parameters and Section V discusses open challenges and future directions. Section VI concludes the paper.

2 Related Work

SA is an interacting range of research in the present era of business promotions. Different kinds of research have been conducted by various researchers too. We analysed many different research papers, which were based on deep learning approaches.

Yang et al. [5] propose a brand-new sentiment lexicon-based sentiment analysis model called SLCABG. Chinese sentiment analysis can make extensive use of the data due to its scale, which has surpassed 100,000 orders of magnitude. According to the experimental findings, the model can significantly enhance text sentiment analysis performance.

Alahmari et al. [6] classify the opinions in texts written in the Saudi dialect using deep learning. Two approaches were then used to perform SA on a dataset of 32,063 tweets that had been gathered. To identify the feelings of the gathered data for comparison, they also used the well-known support vector machine (SVM) technique.

Mittal et al. [7] focus on a few notable deep learning models as well as the applicability of their applications in image SA and their drawbacks. The report also covers the issues and prospects facing this developing profession.

Gouliaras et al. [8] compare various techniques applied to Twitter data SA. In this study, they assess and contrast CNN ensembles and combinations with LSTM networks, a subset of RNN. They also contrast various word embedding schemes. *Onan* et al. [9] offer an effective opinion categorization technique with great prediction accuracy in MOOC reviews. They aim to address a number of research problems on SA of educational data in this contribution. According to the empirical investigation, deep learning-based architectures perform better at SA on educational data mining than ensemble learning approaches and supervised learning methods.

Cen et al. [10] introduce three networks used for SA of IMDB movie feedbacks. 50% of the reviews were positive and 50% were negative in the dataset. While CNN is frequently employed in image recognition tasks, RNN and LSTM neural networks are the two primary varieties that are regularly used in NLP activities.

Wadawadagi et al. [11] describe opinion classification and their uses in an empirical investigation. The paper first conducts an analysis of many modern DNN models and their underlying theories. Additionally, experiments using sentiment datasets are used to determine how well various DNN models presented in the literature perform. Following this research, the performance of each model is also studied in relation to the impact of adjusting certain hyperparameters. A few straightforward data visualisation approaches have been used to aid in the better understanding of the empirical findings.

Tyagi et al. [12] to analyse feelings and classify evaluations or opinions into two polarities—positive or negative. On a benchmark dataset, their suggested methodology for implementing the results performs better. When compared with standard machine learning techniques, the CNN-LSTM-based deep learning method performed better.

Lin et al. [13] to forecast stock markets, analysts use unstructured data from social media with structured data. The performance of forecasting models is heavily influenced by parameter selection and has emerged as potent solution to forecasting issues.

Kapočiūtė-Dzikiienė et al. [14] discuss the SA experiments, incorporating characteristics into traditional machine learning algorithms. Although deep learning methods outperformed more conventional machine learning techniques, deep learning showed promising results when used on smaller datasets.

Chen et al. [15] create a framework and approaches for SA on social media in order to analyse the performance of several methods. The outcome findings demonstrate that the results integrate current opinion dictionaries.

Ali et al. [16] introduce comparative results of several deep learning networks and describe a constructed classification SA using deep learning networks. The hybrid model outperformed the single model, according to the results.

Shilpa et al. [17] identify an expression's sentiments as favourable or negative using deep learning algorithms. On three separate datasets, they tried and assessed the technique utilising RNN and LSTM.

Kottursamy et al. [18] highlight standard deep learning techniques for recognising emotions while employing the eXnet library to increase accuracy. Contrarily, memory and computation still pose challenges. A concern with large models is overfitting. Bringing down the generalisation error is one way to overcome this problem. They use a brand-new CNN called eXnet to build a new CNN model that makes use of parallel feature extraction. It makes use of efficient techniques to lessen overfitting while maintaining proper body weight.

Dash tipour et al. [19] provide a revolutionary Persian sentiment analysis method that is deep learning-driven and context-aware. According to the findings of the simulation, the LSTM algorithm performed better than other models.

Karasoy et al. [20] classify SMS messages based on their content in order to filter out undesired messages for Turkish speakers. Turkish SMS messages from a variety of age groups and geographic areas were collected to create the dataset.

Table 1 gives the details of past work comparisons. It consists of the reference number, purpose, techniques, results, and future work. From above Table 1, we analyzed that most of the authors worked on the various domain data of the sentiment analysis in order to enhance its features and accuracy. Various domain data on sentiment analysis can be predicted using word2vec, fast Text, and Glove, a novel multi-strategy, deep neural network, hybrid model, and many others. The effectiveness of various domain data is enhanced by employing these techniques. We found that the combining hybrid strategy improves the accuracy of the result and exceeds the most recent algorithms for sentiment classification.

3 Deep Learning Techniques

Below is a detailed description of the procedures employed in the present study.

- *CNN*: Artificial neural networks known as “convolutional neural networks” are capable of accurately detecting information in a variety of situations. Several issues with artificial natural language processing and picture processing, including question and answer answering, opinion analysis, and text summarization have been resolved by this methodology. A specific architecture is what makes it unique because it makes learning easier. As a multilayer network, a convolutional neural network uses the output of one layer as the input for the subsequent layer. A typical configuration consists of an input, one to multiple hidden layers, and an output [21].
- *RNN*: Sequential data are modelled using RNN, a form of Rec NN. Sequential data has various uses. The primary distinction between RNN and Rec NN is that RNN considers the processing time for each element in a sequence. As a result, the output of RNNs depends on both the output obtained from the network’s previously hidden state as well as the input that is currently being used. RNNs store the internal states of the inputs by iteratively processing each word in a

Table 1 Comparison of different domain datasets and methodologies

Ref. No.	Method	Result	Future scope
[5]	CNN, BiGRU, SLCABG, GRU	They discovered from the experiment that sentiment features produce little accuracy	By utilising Google's training strategy, they will create their own word2Vec that is appropriate for medical web forums in the future
[6]	LSTM, Bi-LSTM, SVM	The findings show that the approach used by SVM is unable to contend with deep learning techniques	In order to make the SDCT dataset bigger and a more reliable source of huge, clean, and annotated data for the study of the Saudi language, they intend to gather more tweets in the future
[7]	CNN, region-based CNN, DNN, and fast R-CNN	The advantages and disadvantages of deep learning as it relates to image SA	Furthermore, when this field of study is carried out, it is anticipated that soon researchers will present more effective approaches to obtain the best outcomes
[9]	CNN, RNN, LSTM, GRU	According to the empirical investigation, deep learning-based architectures outperform supervised and ensemble learning techniques	Additionally, they provide the first corpus of reviews for massive open online courses, which could be useful for future research
[10]	CNN, RNN, LSTM, RNN- AM	The outcomes demonstrated that the CNN network model can produce an effective classification result	In further study, they plan to create better models for tests, improve their own synthesis using the combined model technique, and increase the impact of film reviews or other sentimental assessments
[14]	NBM, SVM, LSTM, CNN	The NBM approach showed the greatest results, with an accuracy of 0.735	They want to narrow the disparity between conventional and deep learning methodologies through future study
[15]	LSTM, BiLSTM	The experimental findings demonstrate that the model's result is superior to those obtained by using existing sentiment dictionaries when it mixes them with the military's own self-developed sentiment dictionary	Additional research can be used to improve model efficacy when coupled with other beneficial learning components and to expand calibration possibilities utilising distinct models and parameters
[17]	RNN, LSTM	A comprehensive examination shows that the approach identifies feelings using the LSTM approach with a precision of 91.3%	In order to make the system more individualised, it will be important to examine the analysis of user personalities based on their tweets in future work

(continued)

Table 1 (continued)

Ref. No.	Method	Result	Future scope
[19]	CNN, LSTM, MLP, SVM, LR, Auto encoder	According to the findings of the simulation, the LSTM algorithm performed better than the MLP, auto encoder, SVM, LR, and CNN algorithms	The established corpus will be expanded to include multilingual product reviews in further work
[20]	RF, NBM, SVM, MLP, LR, KNN, LSTM, CNN, CNN + Word2Vec	As a consequence, the CNN algorithm, which has an accuracy classification rate of 99.86%, has been identified as the most successful method	Future studies can assess how well the method used in this study integrates with various online messaging platforms, such as Viber, WhatsApp, Instagram, and Messenger

phrase. In order to anticipate the next word in a phrase, RNN will therefore record all the words that came before it and their associations [22].

- **LSTM:** By adding a gating mechanism to the conventional RNN, Hochreiter et al. in 1997 proposed the LSTM to address the problem of vanishing gradient. They developed the forget gate, which enables the memory cell to store information for an extended period of time or discard the results of earlier computations. But LSTM is charged with having a complicated structure. The gated recurrent unit (GRU) was created as a more straightforward version of LSTM as a result [23].
- **GRU:** Another RNN variation that resembles LSTM is GRU. In contrast to LSTM, which has three layers, GRU has just two. The combination of fresh input and earlier computations is controlled by the first gate, called the reset gate. The second gate, known as the update gate, chooses which data to retain from the prior computations. In comparison with LTSM and plain RNN, GRU is thought of as a more efficient LSTM model in terms of processing power [23].
- **BiLSTM:** RNN, also known as bidirectional LSTM, is one type. It uses two hidden layers and handles data in two directions. This is where LSTM and other methods diverge most. The natural language processing capabilities of BiLSTM have been demonstrated [21].

4 Evaluation Measures

Most cutting-edge SA systems use accuracy, F1 score, precision, and recall as their primary performance metrics [22]. Here is an explanation of these actions:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{P + N} \\ &= \frac{TP + TN}{TP + TN + FP + FN} \end{aligned} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} = 1 - \frac{FN}{T} \quad (3)$$

$$F1 = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where TP is the true positive, FP is the false positive, TN is the true negative, and FN is the false negative [22].

5 Open Challenges and Future Directions

Research on SA has provided a number of reliable methodologies, however there are a few intriguing problems that still need to be explored:

- *Public Datasets and Models*: The majority of SA studies in education continue to use very tiny datasets that were not made available to the public and make it challenging to train a neural network. It is important to exchange more datasets and models [24].
- *Text Representation Modelling*: Current methods, like Word2Vec, have limitations that researchers can use to predict future trends. Large datasets in the intended context are needed for this, together with abundant storage and processing power. Finally, because word embeddings are often trained on neural networks with a small number of hidden layers, their semantic potential is constrained.
- *Sentiment Prediction Model Design*: FNN has been typically used as the foundational architecture for SA systems. Its highly connected layers, however, only have access to the input that is being processed at the moment and have no recall of any previous input. Recent studies have demonstrated that RNNs can deliver cutting-edge embeddings that largely correct the drawbacks of earlier methods.
- *Transfer Learning across Contexts*: The context that the underlying training data targets tends to affect existing models negatively. As a result, users may easily train pre-trained models using modest quantities of context-specific data to create new services on top of them.
- *Multi-aspect Sentiment Modelling*: The main application of SA in the realm of education has been to determine the overall polarity or sentiment score of a feedback. More research considering these many potentially connected characteristics presented within a single review is required in order to gain meaningful understanding on how people perceive each of them [24].

6 Conclusion

In conclusion, this survey of SA using deep learning models offers a thorough examination of the developments, difficulties, and probable future directions in the field. The survey demonstrated how deep learning approaches have developed over time, moving from basic neural networks to more complex structures models. The capabilities of sentiment analysis models have been considerably improved by these developments, enabling them to comprehend nuanced linguistic details and contextual information. The survey's results have a promising future scope. As sentiment analysis remains a crucial tool for comprehending human behaviour and preferences, there are many opportunities for additional research and development. For example, extending sentiment analysis to include multiple informational formats, such as images, videos, and audio may result in a more comprehensive understanding of sentiment expression.

References

1. Ain QT, Ali M, Riaz A, Noureen A, Kamran M, Hayat B, Rehman A (2017) Sentiment analysis using deep learning techniques: a review. *Int J Adv Comput Sci Appl* 8(6)
2. Prabha MI, Srikanth GU (2019) Survey of sentiment analysis using deep learning techniques. In: 2019 1st International conference on innovations in information and communication technology (ICIICT). IEEE, pp 1–9
3. Souma W, Vodenska I, Aoyama H (2019) Enhanced news sentiment analysis using deep learning methods. *J Comput Soc Sci* 2(1):33–46
4. Jain K, Kaushal S (2018) A comparative study of machine learning and deep learning techniques for sentiment analysis. In: 2018 7th International conference on reliability, infocom technologies and optimization (Trends and Future Directions) (ICRITO). IEEE, pp 483–487
5. Yang L, Li Y, Wang J, Sherratt RS (2020) Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* 8:23522–23530
6. Alahmary RM, Al-Dossari HZ, Emam AZ (2019) Sentiment analysis of Saudi dialect using deep learning techniques. In: 2019 International conference on electronics, information, and communication (ICEIC). IEEE, pp 1–6
7. Mittal N, Sharma D, Joshi ML (2018) Image sentiment analysis using deep learning. In: 2018 IEEE/WIC/ACM International conference on web intelligence (WI). IEEE, pp 684–687
8. Goularas D, Kamis S (2019) Evaluation of deep learning techniques in sentiment analysis from twitter data. In: 2019 International conference on deep learning and machine learning in emerging applications (Deep-ML). IEEE, pp 12–17
9. Onan A (2021) Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach. *Comput Appl Eng Educ* 29(3):572–589
10. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Alkwai LM, Kumar S (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Elect Eng* 107:108617, ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
11. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerging Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>

12. Sharma R, Arya R (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Elect Eng* 108:108715, ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
13. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun*, ISSN 0140-3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
14. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM, An efficient hybrid deep learning model for denial of service detection in cyber physical systems. In: IEEE transactions on network science and engineering. <https://doi.org/10.1109/TNSE.2023.3273301>
15. Gupta U, Sharma R (2023) Analysis of criminal spatial events in India using exploratory data analysis and regression. *Comput Electr Eng* 109, Part A, 108761, ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108761>
16. Goyal B et al, Detection of fake accounts on social media using multimodal data with deep learning. In: IEEE transactions on computational social systems. <https://doi.org/10.1109/TCSS.2023.3296837>
17. Sneha, Malik P, Sharma R, Ghosh U, Alnumay WS (2023) Internet of things and long-range antenna's; challenges, solutions and comparison in next generation systems. *Microprocess Microsyst* 104934, ISSN 0141-9331. <https://doi.org/10.1016/j.micpro.2023.104934>
18. Vohnout R et al, Living lab long-term sustainability in hybrid access positive energy districts-A Prosumager smart fog computing perspective. *IEEE Internet of Things J.* <https://doi.org/10.1109/JIOT.2023.3280594>
19. Yu X, Li W, Zhou X et al (2023) Deep learning personalized recommendation-based construction method of hybrid blockchain model. *Sci Rep* 13:17915. <https://doi.org/10.1038/s41598-023-39564-x>
20. Yadav S et al, Video object detection from compressed formats for modern lightweight consumer electronics. In: IEEE transactions on consumer electronics. <https://doi.org/10.1109/TCE.2023.3325480>
21. Rhanoui M, Mikram M, Yousfi S, Barzali S (2019) A CNN-BiLSTM model for document-level sentiment analysis. *Machine Learn Knowl Extr* 1(3):832–847
22. Yadav A, Vishwakarma DK (2020) Sentiment analysis using deep learning architectures: a review. *Artif Intell Rev* 53(6):4335–4385
23. Habimana O, Li Y, Li R, Gu X, Yu G (2020) Sentiment analysis using deep learning approaches: an overview. *Sci China Inf Sci* 63:1–36
24. Pathak AR, Agarwal B, Pandey M, Rautaray S (2020) Application of deep learning approaches for sentiment analysis. *Deep learning-based approaches for sentiment analysis*, pp1–31

Identifying Multiple Diseases on a Single Citrus Leaf Using Deep Learning Techniques



Ayushi Gupta , Anuradha Chug, and Amit Prakash Singh

Abstract Deep learning techniques for classifying images into multiple classes have made significant strides in the past few years. Nevertheless, allocating multiple classes to a single image, as seen in object detection scenarios, remains a relatively unexplored avenue of research. This study is dedicated to harnessing the potential of deep learning methods to discern various diseases within a singular leaf image of a citrus fruit. The investigation focuses on citrus leaves affected by none, one, two or all three severe diseases—anthracnose, melanose and bacterial brown spot. These leaves serve as the training dataset for ten distinct deep learning models. The performance of these models is meticulously examined, gauging their accuracy in classifying the diseases. The outcomes highlight the superiority of the DenseNet121 architecture in terms of accuracy and training duration, as it accurately identifies overlapping disease classes present in citrus leaves. Following suit, the MobileNetV2 architecture showcases comparable accuracy and a noteworthy reduction of 66% in training time.

Keywords Deep learning · Leaf disease classification · Multi-label images

1 Introduction

Citrus fruits, being excellent sources of vital nutrients, have proven beneficial for human health. The diversity of the fruit is well-known in terms of oranges, tangerines, limes and lemons, grapefruits and many other varieties, with India being the leading supplier of limes and lemons in 2021 [1]. Citrus fruits are among the world's top two most widely produced fruits, with a production of 161.8 million tonnes in 2021. However, the fruits are significantly affected by pests and pathogens, leading to degraded fruit quality, disease outbreaks, and, therefore, major yield losses [2]. Anthracnose or *Collectotrichum gloeosporioides* is a fungus characterized by yellow

A. Gupta  · A. Chug · A. P. Singh

University School of Information, Communication and Technology, GGSIPU, New Delhi, India
e-mail: ayushi.20616490021@ipu.ac.in

tips on the leaves and tiny brown speckles on the citrus fruit [3]. Melanose, caused by the fungus *Diaporthe citri*, affects the citrus leaves and fruits in the form of small spots or scab-like lesions on the surface [4]. Bacterial brown spot caused by the bacterium *Burkholderia andropogonis* leads to flat circular lesions on leaves with raised margins and may lead to early dropping of severely infected leaves [5]. With the increase in global warming and changing climatic conditions, the growth of such diseases in plants is accelerated. Therefore, their timely identification is necessary so that appropriate steps can be implemented to mitigate substantial crop reduction [6]. In recent times, numerous studies have been conducted to classify multiple diseases in plant images. Table 1 mentions some of the latest studies. However, classifying a plant image with more than one disease is still an open area of research. A few studies have employed object detection methods for bounding box predictions of multiple diseases in leaf images, which are also listed in Table 1. In [7], the authors worked on bounding box predictions in citrus leaves for identifying multiple diseases in a single image. Nonetheless, manual detection of bounding boxes in images is an exhaustive task and requires professionals for accurate measures; therefore, limited data is available regarding the same. The novelty of this study lies in the fact that it attempts to classify an image with multiple labels without the availability of bounding box measures. Ten deep learning models—DenseNet121, InceptionResNetV2, MobileNetV2, NASNetMobile, ResNetRS101, Xception, VGG16, InceptionV3, VGG19 and RegNetX002 have been employed to detect overlapping disease classes in citrus leaf images. This will benefit agricultural workers in identifying multiple diseases in plants and taking the necessary actions.

The further sections describe in detail the background methodology (Sect. 2), the results achieved (Sect. 3), followed with the conclusion and future works (Sect. 4).

2 Methodology

This section describes the approach and techniques used in this study. The dataset utilized is presented in Sect. 2.1. A brief overview of the classifiers is given in Sect. 2.2, and the research methodology is detailed in Sect. 2.3.

2.1 Dataset

The study utilizes the Conghua Citrus Leaf 2020 (CCL'20) dataset publicly available at <https://github.com/researchzkhu/CCL-20>. The dataset consists of citrus leaf images infected with one, two or all of the three diseases—anthracnose (A), melanose (M) and bacterial brown spot (B). The dataset also contains healthy leaf images. If $S = \{A, B, M\}$ is the set of distinct diseases in the dataset, the possible overlapping classes will be the power set P of S , and the total number of overlapping classes will be the cardinality of the P . Hence, $P = \{\text{Healthy}, A, M, B, AM, AB, BM, ABM}\},$

Table 1 Literature survey on disease identification in leaf images

Year	Model	Crop	Disease	Accuracy (%)
2023	Convolutional neural network (CNN) [8]	Paddy	Brown spot, leaf blast, sheath rot, false smut, bacterial blight	91.45
2023	Class activation maps, ResNet18 [9]	Segmentation dataset	Lesion	Success rate—0.45
2023	EfficientNetB3, ResNet50, MobiNetV2, InceptionV3 [10]	Citrus	Black spot, canker, greening, melanose	99.58
2023	YOLOv3-tiny, YOLOv4, YOLOv5s, YOLOv7s, YOLOv8n [11]	Maize	Blight, sugarcane mosaic virus, leaf spot	mAP—99.04
2023	DeepDream [12]	Tomato	Multiple	96
2023	YOLOv7 [13]	Tea	Red spiders, mosquito bugs, black rot, brown blight, leaf rust	97.3
2023	YOLOv3 [14]	Bell pepper	Bacterial spot	mAP—90
2022	CentreNet, YOLOv4, Faster-RCNN, DetectoRS, Cascade-RCNN, Foveabox, Deformable Detr [7]	Citrus	Anthracnose, melanose, bacterial brown spot	—
2021	2 stage CNN [15]	Citrus	Black spot, canker, greening	94.37
2021	CNN [16]	Citrus	Black spot, canker, scab, greening, melanose	94.55

which results in eight overlapping classes. However, the data contains no citrus leaves affected with both A and B, resulting in seven class labels. The dataset has been preprocessed with data augmentation techniques—resizing, brightness, contrast and vertical and horizontal flips, to attain a higher quantity of diseased images. The total image count after augmentation for each class is shown in Fig. 1.

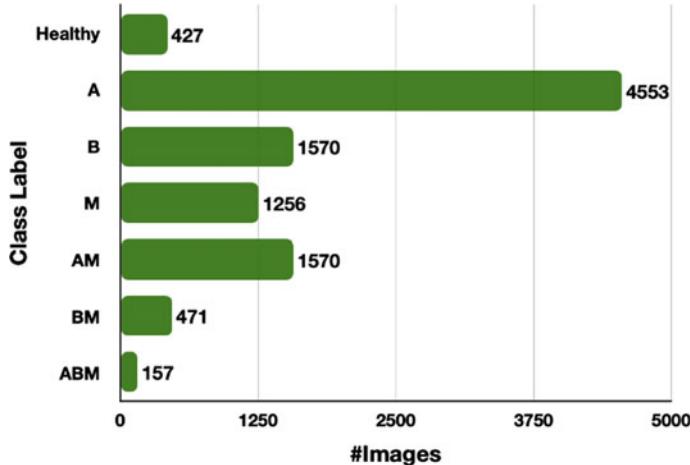


Fig. 1 Class distribution in CCL'20

2.2 Intelligent Classifiers

Convolutional neural networks (CNN) are well-known and most widely used for image processing tasks. The current research uses ten CNNs pre-trained on the ImageNet dataset consisting of 1000 classes. DenseNet121 [17] consists of 58 bottleneck layers followed by 3×3 convolutions composed in dense blocks. Each layer receives the features of all the previous layers concatenated with each other. InceptionV3 [18] is a 48-layer deep network that performs factorization of large convolutions for dimensionality reduction and replaces $n \times n$ convolutions with $1 \times n$ and $n \times 1$ convolutions for computation speedup. InceptionResNetV2 [19] is a 164-layer deep network that uses reduction blocks along with residual connections between the inception modules that add the output of the previous module to the next. MobileNetV2 [20] is a 53-layer deep network that uses depth-wise separable convolutions to reduce computational costs followed by bottleneck layers and residual blocks. NASNetMobile [21] or neural architecture search model searches for the best architecture and employs performance evaluation strategies to better the model architectures without actually training them. ResNetRS101 [22] uses improvised training and scaling strategies rather than architectural changes for faster and better performance. Xception [23] or extreme inception first applies filters and then uses 1×1 convolutions across the depth. VGG16 and VGG19 [24] are 16- and 19-layer deep networks which employ 3×3 convolutions uniformly throughout the network, making the models more straightforward. Lastly, RegNetX002 [25] is a model residing in RegNetX design space that uses quantized linear functions for searching good networks.

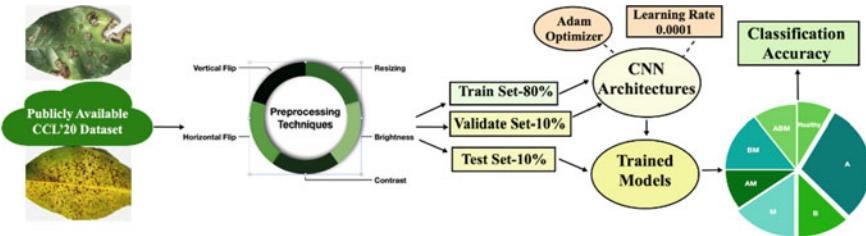


Fig. 2 Research methodology

2.3 Methodology Diagram

Figure 2 details the methodology employed in the current research. After preprocessing the dataset, the data is divided into 80-10-10 train-test-val ratio. The ten CNN architectures are trained using the train and validation sets. For training the models, all the top layers have been frozen to use the weights learned on the ImageNet dataset. Two fully connected layers are added to the flattened output of the models, each with 512 neurons and ReLu activation, followed by dropout layers with a value of 0.5. The training has been done using Adam optimizer and a learning rate of 0.0001. Finally, a Softmax layer has been added with the number of classes set to 7, and the input images' class labels are predicted. The classification accuracy of the models on the test set is analyzed, and the best-performing model is reported.

The utilization of Python's Tensorflow library [26] was employed for the implementation. Additionally, we displayed the training and validation accuracy plots while the model was being trained. These implementations were conducted on a system equipped with an Intel(R) Core(TM) i3-5005U CPU running at 2 GHz and 12 GB of RAM.

3 Results

This section presents the results achieved by the classifiers in terms of classification accuracy. Figure 3 shows the performance of the models along with their training duration. The following observations can be made:

- The performance bars for ResNetRS101 and RegNetX002 are absent from Fig. 3 due to their accuracy falling below the 90% threshold, specifically 70.92% for ResNetRS101 and 42.25% for RegNetX002. This performance efficiency is attributed to the models' inability to effectively retain information about classes that overlap.
- DenseNet121 and Xception emerged as the top performers, achieving an impressive accuracy score of 98.4% each. It is noteworthy that DenseNet121 achieved this exceptional performance while also requiring less training time compared with

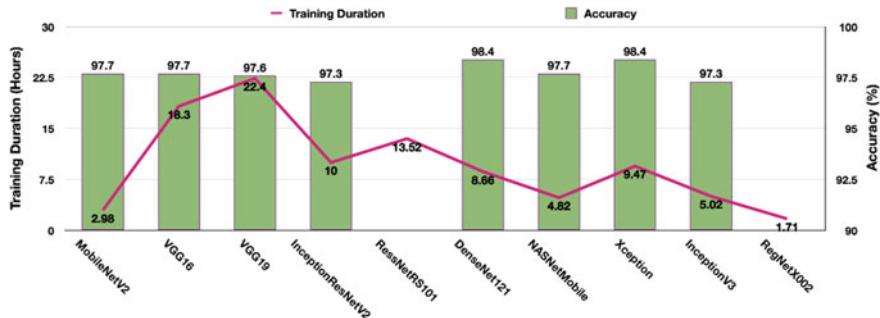


Fig. 3 Model performance

Xception, with training durations of 8.66 and 9.47 h, respectively. This equates to a reduction of approximately 9% in training time for DenseNet121 compared with Xception

- Four model architectures—MobileNetV2, VGG16, VGG19 and NASNetMobile attained similar accuracy levels of 97.7%. Notably, MobileNetV2 exhibited the shortest training time among them all, clocking in at just 2.98 h. This represents a substantial reduction in training time compared with the other models, specifically 87% less time than VGG19, 84% less than VGG16 and 38% less than NASNetMobile.

To deeply understand the behaviour of the models in predicting overlapping classes, the confusion matrix of the top two performers—DenseNet121 and Xception are shown in Fig. 4. The observations are:

- Among the 423 images labelled as A, both models successfully classify 422 images accurately, but they both classify one image as B incorrectly. This misclassification raises suspicions of potential incorrect class labelling, as both models yield identical outcomes in this particular instance.
- Both models effectively and accurately classify the healthy images and those labelled as ABM (exhibiting infections from all three diseases).
- Among the 176 images labelled as AM, DenseNet121 accurately classifies 171 of them. However, it falls short in identifying the presence of disease A in four of these images. In contrast, Xception successfully classifies 168 images correctly, yet it also struggles to detect disease A in eight other images.
- Out of 166 images labelled B, DenseNet121 misclassifies one image as BM, erroneously identifying both B and M in a single instance. On the contrary, Xception incorrectly classifies one of the images as M, demonstrating a distinct misclassification pattern.
- DenseNet121 correctly classifies all the images labelled as BM, while Xception fails to identify the disease M in four of the images.
- The majority of images labelled as M are accurately classified by both models. Nevertheless, there are instances where both models mistakenly identify disease

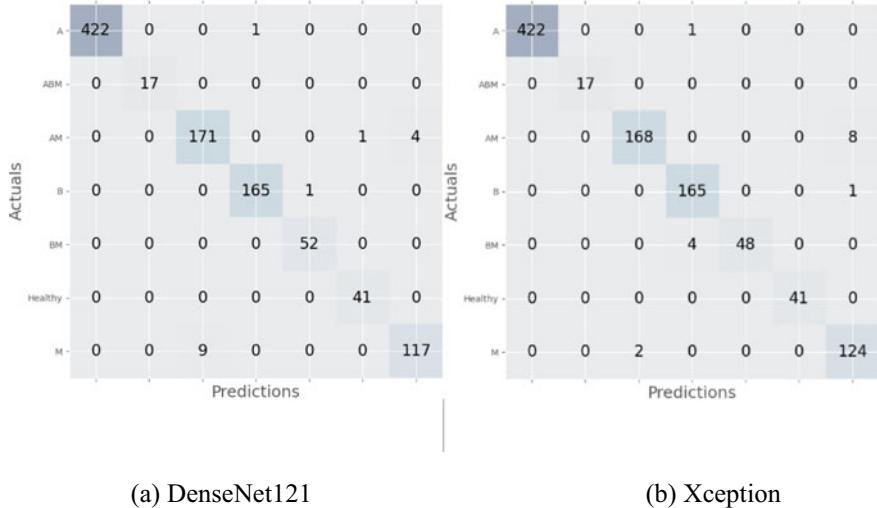


Fig. 4 Confusion matrices of top-performing models

A alongside M in a subset of the images—DenseNet121 in nine images while Xception in two images.

The observations indicate that the DenseNet121 architecture demonstrates a higher level of accuracy in detecting overlapping classes than the Xception architecture. Conclusively, DenseNet121 exhibits effective capability in detecting multiple diseases within a single citrus leaf, making it a viable option for real-time disease prediction applications.

Figure 5 illustrates the training and validation curves, presenting the accuracy and loss trends of the DenseNet121 architecture over the course of 20 epochs. The figure demonstrates that the model exhibits similar validation accuracy and loss patterns in comparison with its training accuracy and loss. This further indicates that the model is not overfitting and is able to recognize overlapping classes.

Considering the training duration of the models, MobileNetV2 and NASNet-Mobile architectures give comparable accuracy values—both 97.7%, corresponding to only 0.7% loss in accuracy, while providing a substantial reduction in training time—66% and 45%, respectively. The confusion matrices of the two architectures are shown in Fig. 6. The following insights are apparent from the figure.

- Among the 423 images labelled as A, MobileNetV2 correctly classifies 416 of them, while NASNetMobile is able to correctly classify 420 of them. Both models classify three images incorrectly as B. MobileNetV2 also misclassifies three of the diseased images as healthy.
- Both models correctly classify all the images labelled as ABM. Further, MobileNetV2 architecture is also able to correctly classify all the images labelled

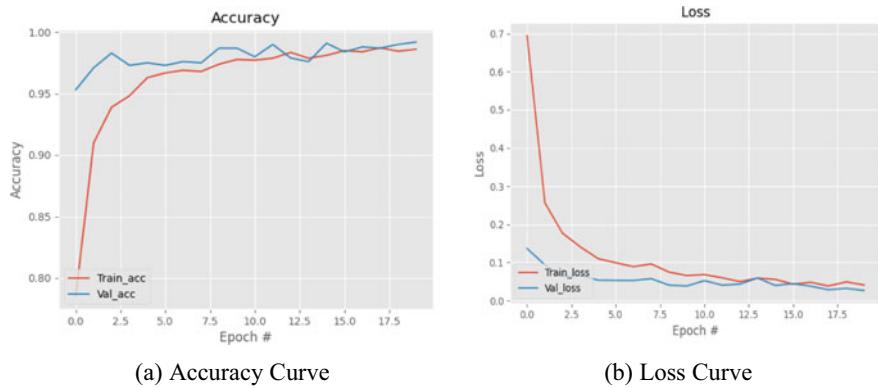


Fig. 5 Training and validation curves for DenseNet121



Fig. 6 Confusion matrices of MobileNetV2 and NASNetMobile

B as well as healthy, while NASNetMobile misclassifies one of the images in both cases.

- Among the 176 images labelled as AM, MobileNetV2 accurately classifies 169 of them. However, it misses the presence of disease A in six of these images. In contrast, NASNetMobile successfully classifies 166 images correctly, yet it also struggles to detect disease A in ten other images.
- Both models fail to identify the disease M in two of the images labelled BM. Further, MobileNetV2 fails to recognize disease B's presence in two images.
- The majority of images labelled as M are accurately classified by both the models. Nevertheless, both the models mistakenly identify disease A alongside M in a subset of the images—MobileNetV2 in five images while NASNetMobile in six images.

The above observations state that the overall performance of both the models is reasonably equivalent in detecting overlapping classes. However, MobileNetV2 architecture requires much less training duration when compared with NASNetMobile architecture—nearly 38% less time. Hence, MobileNetV2 architecture turns out to be the second optimal choice after DenseNet121 architecture for detecting overlapping diseases in citrus leaves.

4 Conclusion and Future Scope

The study employs ten distinct deep learning techniques to effectively discern multiple diseases within a single citrus leaf. The findings highlight that the DenseNet121 architecture emerges as a standout performer, achieving an exceptional classification accuracy of 98.4% while requiring a reasonable training duration of 8.66 h. Alternatively, for situations demanding quicker training, the MobileNetV2 architecture offers a comparable accuracy level of 97.7% while substantially reducing training time by 66%.

In the future, there is potential for enhancement by incorporating a sigmoid activation function in the final layer to predict individual class probabilities. This approach could involve setting a threshold to determine the ultimate disease classes. Furthermore, the exploration of ensemble techniques, combining models proficient in detecting different overlapping classes, could yield a more robust classifier capable of identifying all overlapping classes collectively. The prospects extend to the development of a real-time mobile application, catering to users seeking efficient detection of multiple overlapping diseases not only in citrus but also across various fruits. This application could provide valuable assistance in disease management and prevention.

Acknowledgements This research is endorsed by the Department of Science and Technology (DST) under a project with reference number “DST/Reference.No.T-319/2018-19”. We thank them for their support.

References

1. FAO, “Food and agriculture organizations of the united nations.” <http://www.fao.org/faostat/en/#data,2023>
2. Khamsaw P, Sangta J, Chaiwan P, Rachtanapun P, Sirilun S, Sringarm K, Thanakkasaraneey S, Sommano SR (2022) Bio-circular perspective of citrus fruitloss caused by pathogens: Occurrences, active ingredient recovery and applications. Horticulturae 8(8):748
3. U. of California Statewide Integrated Pest Management Program(UC IPM), “Anthracnose of citrus—collectotrichum gloeosporioides.” <https://ipm.ucanr.edu/PMG/GARDEN/FRUIT/DISEASE/citanthracnose.html#:text=Anthracnose%20of%20citrusCollectotrichum%20gloeosporiooides,Lemons%20are%20occasionally%20infected.,,2023>
4. Nelson S (2008) Citrus melanose

5. Duan Y, Sun X, Zhou L, Gabriel D, Benyon L, Gottwald T (2009) Bacterial brown leaf spot of citrus, a new disease caused by *Burkholderia andropogonis*. *Plant Dis* 93(6):607–614
6. Elame F, Chebli Y, Jamal H, Hayat L (2023) Climate change impact modeling on citrus yield. In: Strategizing agricultural management for climate change mitigation and adaptation. Springer, pp 233–245
7. Dananjayan S, Tang Y, Zhuang J, Hou C, Luo S, Assessment of state-of-the-art deep learning based citrus disease detection techniques using annotated optical leaf images. *Comput Electronics Agric* 193:106658
8. Haridasan A, Thomas J, Raj ED (2023) Deep learning system for paddy plant disease detection and classification. *Environ Monit Assess* 195(1):120
9. Uysal ES, Sen D, Ornek AH, Yetkin AE (2023) Lesion detection on leaves using class activation maps. arXiv preprint [arXiv:2306.13366](https://arxiv.org/abs/2306.13366)
10. Faisal S, Javed K, Ali S, Alasiry A, Marzougui M, Khan MA, Cha JH (2023) Deep transfer learning based detection and classification of citrus plant diseases. *Comput Mater Continua* 76(1)
11. Khan F, Zafar N, Tahir MN, Aqib M, Waheed H, Haroon Z (2023) A mobile-based system for maize plant leaf disease detection and classification using deeplearning. *Front Plant Sci* 14:1079366
12. Sahu P, Chug A, Singh AP, Singh D (2023) Classification of crop leaf diseases using image to image translation with deep-dream. *Multimedia Tools Appl* 1–35
13. Soeb MJA, Jubayer MF, Tarin TA, Al Mamun MR, Ruhad FM, Parven A, Mubarak NM, Karri SL, Meftaul IM (2023) Tea leaf disease detection and identification based on yolov7 (yolo-t). *Sci Rep* 13(1):6078
14. Mahesh TY, Mathew MP (2023) Detection of bacterial spot disease in bell pepper plant using yolov3. *IETE J Res*:1–8
15. Syed-Ab-Rahman SF, Hesamian MH, Prasad M (2022) Citrus disease detection and classification using end-to-end anchor-based deep learning model. *Appl Intell* 52(1):927–938
16. Khattak A, Asghar MU, Batool U, Asghar MZ, Ullah H, Al-Rakhami M, Gumaei A (2021) Automatic detection of citrus fruit and leaves diseases using deep neural network model. *IEEE Access* 9:112942–112954
17. Huang G, Liu Z, Van Der Maaten L, Weinberger KY (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
18. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
19. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence 31
20. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
21. Zoph B, Vasudevan V, Shlens J, Le QV (2018) Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 8697–8710
22. Bello I, Fedus W, Du X, Cubuk ED, Srinivas A, Lin TY, Shlens J, Zoph B (2021) Revisiting resnets: improved training and scaling strategies. *Adv Neural Inf Process Syst* 34:22614–22627
23. Chollet F (2017) Xception: deep learning with depth wise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258
24. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)

25. Radosavovic I, Kosaraju RP, Girshick R, He K, Dollár P (2020) Designing network design spaces. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10428–10436
26. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M et al (2016) Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint [arXiv:1603.04467](https://arxiv.org/abs/1603.04467)

IoT-Based Health Monitoring System for Heartbeat—Analysis



B. Mary Havilah Haque and K. Martin Sagayam

Abstract The health of human is the functional ability to face everyday in life. The human living has changed a lot and in a better way in this modern days. The Internet of Things (IoT) technology melding with healthcare sector affirms every individual good efficiency. IoT in healthcare sector assures a very improved and better treatment as it supports remote monitoring. In remote monitoring, the human body is monitored remotely with less human involvement. This paper includes the work done with MAX30100, AD8232 sensors for heart monitoring. The average error % is 4.04%, and the average accuracy is 96.12% which is the result of analysis.

Keywords Internet of Things · Health monitoring · Raspberry Pi

1 Introduction

1.1 *Internet of Things is Everywhere*

The things like sensors or any other thing could be connected to each other and exchange (communicate) a specific required data over Internet, and this is to be done without or with less human involvement that is related to define IoT. The IoT technology coverage is so vast that it almost includes everything, namely all the communicating devices, software, algorithms and with artificial intelligence (AI) for computing, classification, decision making, etc. are done. Having those previously mentioned in IoT, its applications are extended to industrial areas, organizations, consumer sector and many. Among these the remote monitoring has highlighted the requirement of IoT.

B. Mary Havilah Haque · K. Martin Sagayam ()

Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu 641114, India

e-mail: martinsagayam.k@gmail.com

1.2 *Remote Monitoring—An Essence of IoT*

Earlier the name explained that monitoring is done in remote site. Nowadays, in almost all possible monitoring areas, the remote monitoring concept has occupied a standard place. The sensors and devices connected together is placed in the site planned for supervision. The remote monitoring raises an alert notification on any adverse situation to avoid any mishaps. The alerting is done with the help of buzzer, alarm sound, sending message to mobile devices, computers and so on.

Healthcare Sector. Healthcare sector aids in caring the patient's health, and the person is supervised in routine whenever required. During the pandemic situation, the regular and direct checkup of health has become uncertain. The health of human is monitored with the concept of remote monitoring. This healthcare sector ensures medical services, modeling and producing the equipment meant for medical. Having IoT in healthcare sector has put all to good use and everything has become smart, which also uplifted the quality of the healthcare service. The customized regular checkup done at hospitals is avoided as it could be done right at the comfortable place of the patient, and the data is sent to doctors or advisors for suggestions to the patient. The figure shown in Fig. 1 is exemplifying the IoT in healthcare sector showing the typical remote monitoring.

2 Literature Survey

The following work is the literature survey related to the digital heart monitoring. The manual study of heartbeat is time-consuming and so there is much research on the topic of smart heart monitoring. These smart heart monitoring systems and modules presented by various authors and researchers have worked with IoT technology. The architecture of IoT and its application in many sectors is given in [1–3].

It is said in [4] that sensors connected to Raspberry Pi (RPi) and RPi connected to Internet act like a server. Using machine learning, the health condition of old people is classified. Here AD8232 is connected to Arduino, and then the data is shared with RPi. This work along with AD8232 sensor, and three other sensors are also used. In [5], the proposed model with machine learning is an automation system and has used predictive method on observation of different metrics. This work also included a heartbeat sensor. The major units covered are AD8232, RPi, and NodeMCU in the proposed model shown in [6]. AD8232 connected to NodeMCU. As per the authors, the model analyzes the report of electrocardiogram (ECG) about heartbeat. In [7], Arduino and Raspberry Pi are connected together so that the Arduino is in turn connected to AD8232.

A noticeable research work carried out in the paper [8] regarding the mountaineers' lives. The heartbeat, ECG and temperature of trekker body are accounted and connected in wireless sensor network (WSN). This model as per the authors does not require WiFi. In [9], in the patient's health-checking parameters, a total of six

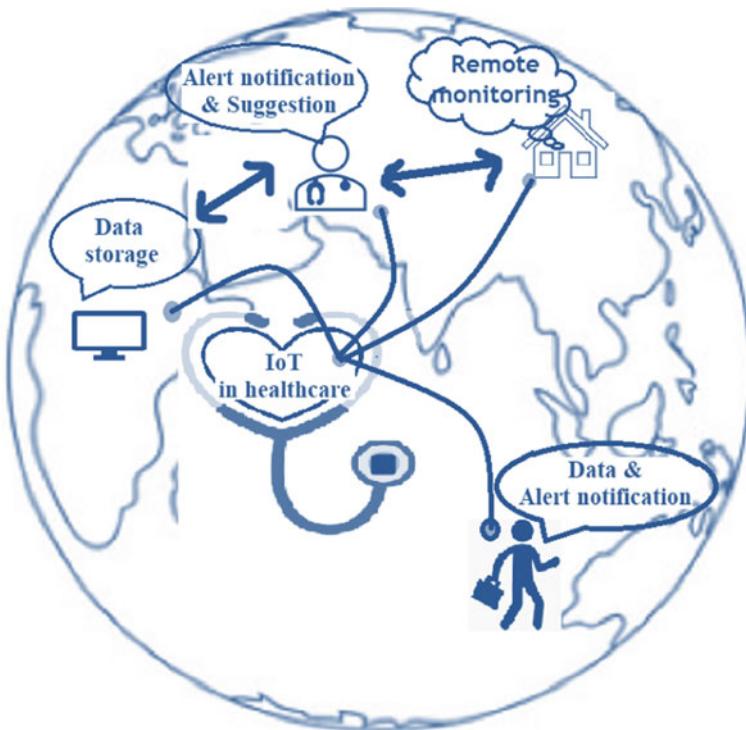


Fig. 1 An exemplar of remote monitoring in IoT of healthcare sector

tests are done at the same location of the patient, considering the situation of patients' exhaustion for moving for different tests from one place to another. In [10], the ECG is used to diagnose the heart rhythm for a disease also in [11], and ECG signal is monitored while the data is taken from MIT-BIH. A heartbeat sensor, Arduino and Bluetooth module are used to carry the work shown in [12] and likewise as of the research work done in this particular area of heart totally in healthcare sector, and the work carried out by most of the authors is by using the sensors to calculate heart rate.

Our proposed model work is very effective and advantageous compared to the paper [13] that carried out work using AD8232, pulse sensor, ESP32, and the sensed data is viewed on Ubidots and ThingSpeak.

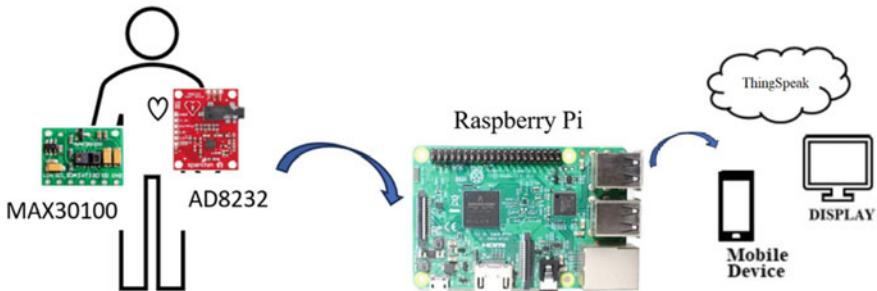


Fig. 2 Proposed model representation in diagram form

3 Methodology

3.1 Internet of Things is Everywhere

The model presented in this paper is helpful to all people who wants to do the heartbeat check remotely. The sensors MAX30100 and AD8232 are connected to Raspberry Pi. The output of the sensors is read on monitor screen, and the data is also made available on cloud. This representation is shown in the Fig. 2. MAX30100 data is obtained by placing the human finger on the sensor, and the sensor gives the beats per minute (BPM) and oxygen saturation (SPO₂) value displayed on the screen. The AD8232 sensor works with 3.3 V, and the MAX30100 sensor works with 3.3 V. AD8232 is ECG sensor. The electrodes of the AD8232 sensor are connected to the human body, and the output is displayed on the screen.

The work done in this paper is that heartbeat values obtained from the output of sensors AD8232 and MAX30100 is compared to that of the heartbeat values obtained by manual procedure, and the error percentage value is calculated. The error % formula is

$$\text{Error \%} = \frac{\text{Calculated value} - \text{Actual value}}{\text{Actual value}} * 100 \quad (1)$$

Equation (1) is used in Table 1. The absolute error value is presented in Table 1. The heart rate found through manual procedure that is the actual value data is taken from [14].

Table 1 Heartbeat from sensor as output and error % calculation

Gender	Actual/Manual heartbeat	Heartrate using MAX30100	Error %	Accuracy %
M	85	89	4.70	95.29
M	86	91	5.81	94.18
M	75	76	1.33	98.66
M	76	80	5.26	94.73
M	71	75	5.63	94.36
F	89	92	3.37	96.62
F	90	86	4.44	95.55
F	83	87	4.81	95.18
F	85	87	2.35	97.64
F	73	71	2.73	97.26

The average error % is 4.04%. The average accuracy is 96.128 %.

4 Results

As the device is power on, the person's finger is to be placed on the sensor, the sensed data that is the data from MAX30100 sensor, and the bpm and SpO₂ is sent to a device with WhatsApp installed in the device previously. The sensor connected to Raspberry Pi is shown in Fig. 3. The received data screenshot is shown in Fig. 5.

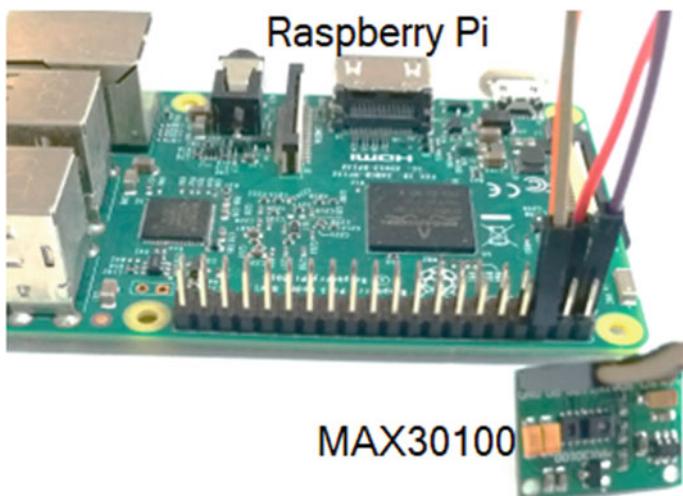


Fig. 3 Raspberry Pi connected to MAX30100

The Raspberry Pi is connected to AD8232 sensor and also to MCP3008. Here, MCP3008 is analog-to-digital converter. The sensors connected to Raspberry Pi are shown in Fig. 4. The AD8232 sensor ECG output plot is seen on ThingSpeak shown in Fig. 6. Upon seeing the ECG plot, the doctors and health advisors may provide suggestions to the patient (Figs. 5 and 6).

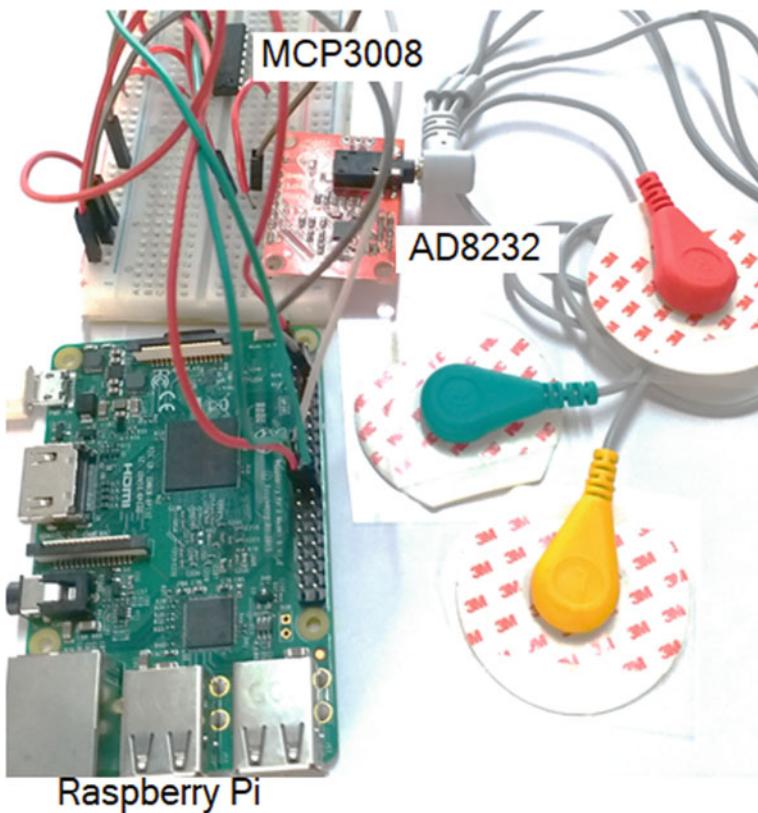


Fig. 4 Raspberry Pi connected to AD8232 and MCP3008



Fig. 5 WhatsApp screenshot of bpm and SpO₂

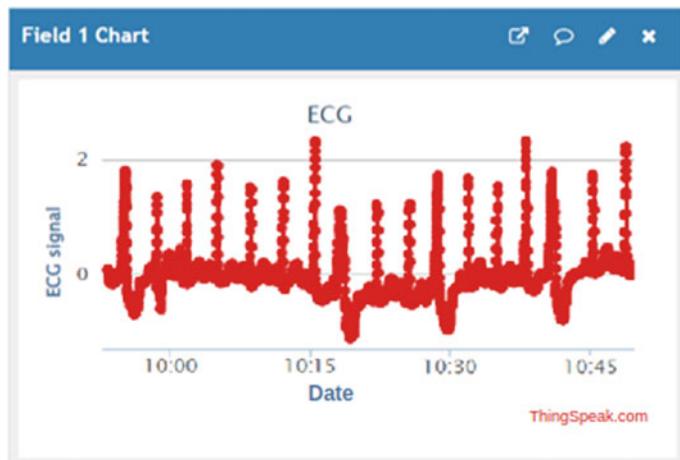


Fig. 6 ECG plot on ThingSpeak

5 Conclusion

The model is very handy and can be carried along. This model avoids the time-consuming procedures for checking the patient. The heartrate sensor data received on WhatsApp is the advantage point in this model because it avoids any additional health-related apps in the devices as new apps occupy extra memory space in the device. It also avoids the problem of SMS alerts. To the existing WhatsApp in the device, the sensor data is received in this model. The ECG plot is accessed by doctors and advisors for suggestions. The average accuracy is 96.12%. The work shown in this paper is tremendously helpful for the researchers.

References

1. Hammi B, Khatoun R, Zeadally S, Fayad A, Khoukhi L (2018) IoT technologies for smart cities. *IET Networks* 7(1):1–13. <https://doi.org/10.1049/iet-net.2017.0163>
2. Choudhary N (2018) A review on IOT-based smart cities. *IRE Journals* 1(10):202–209
3. Agaj M, Borate G, Gharat P, Mhatre V (2020) Smart city using IOT. *IJIRT* 6(12):382–389. <https://doi.org/10.22214/ijraset.2021.33627>
4. Nandyal S, Kulkarni RU, Metre RP (2019) Old age people health monitoring system using IoT and ML. *Int J Innov Sci Res Technol* 4(5):160–166
5. Godi B, Viswanadham S, Muttipati AS, Prakash Samantray O, Gadiraju SR (2020) E-healthcare monitoring system using IoT with machine learning approaches. *Int Conf Comput Sci Eng Appl ICCSEA* 1–5. <https://doi.org/10.1109/ICCSEA49143.2020.9132937>
6. M.-I.-A. M, Kabir MH (2021) A healthcare system for Internet of Things (IoT) application : machine learning based approach. *J Comput Commun* 21–30. <https://doi.org/10.4236/jcc.2021.97003>

7. Jinan UA, Rahman A, Us-Salehin Z (2020) Fog assisted and IoT based real-time health monitoring system implementation 2:99–106
8. Garg RK, Bhola J, Soni SK (2021) Healthcare monitoring of mountaineers by low power wireless sensor networks. *Inf Med Unlocked* 27. <https://doi.org/10.1016/j imu.2021.100775>
9. Akash MRR, Yousuf, Shikder K (2020) IoT based real time health monitoring system. *Proc Int Conf Res Innov Knowl Manag Technol Appl Bus Sustain INBUSH 2020* 167–171. <https://doi.org/10.1109/INBUSH46973.2020.9392163>
10. Bhoi AK, Mishra P, Sarkar S, Manimegalai P (2012) A significant approach to detect heart rate in ECG signal more. *IJAEEE* 1(1):1–4
11. Muankid A, Ketcham M (2019) The real-time electrocardiogram signal monitoring system in wireless sensor network. *Int J Online Biomed Eng* 15(2):4–20. <https://doi.org/10.3991/ijoe.v15i02.9422>
12. Khamitkar SS, Rafi M (2020) IoT based system for heart rate monitoring. *Int J Eng Res Technol* 9(7):1563–1571. <https://doi.org/10.17577/ijertv9is070673>
13. Rahman MA, Li Y, Nabeed T, Rahman MT (2021) Remote monitoring of heart rate and ECG signal using ESP32. In: 2021 4th international conference on advanced electronic materials, computers and software engineering AEMCSE, pp 604–610. <https://doi.org/10.1109/AEMCSE51986.2021.00127>
14. Sari NN, Gani MN, Aprilia Regina YM, Firmando R (2021) Telemedicine for silent hypoxia : improving the reliability and accuracy of Max30100-based system 22(3):1419–1426. <https://doi.org/10.11591/ijeecs.v22.i3.pp1419-1426>

A Study and Comparison of Cryptographic Mechanisms on Data Communication in Internet of Things (IoT) Network and Devices



Abhinav Vidwans and Manoj Ramaiya

Abstract A trending topic that has grown in prominence over the past several years is the Internet of Things (IoT). With the steadily increasing adoption rate of Internet-enabled devices in applications like smart homes, smart cities, smart grids, and health-care applications, the demand of IoT becomes increasing. Now this needs to guarantee the safety of data and communications privacy among these IoT devices and their supporting infrastructure. IoT involves numerous low-resource devices, and most of these devices often need to secure interaction with their network administrators, which are IoT network nodes with more resources. Even though more services and applications are being connected through wireless network, security is still a major concern. IoT system security is a subject that is actively being researched. In this research paper, we discussed and contrasted the various network-based IoT authentication and communication security techniques. Few key points are also discussed regarding the IoT challenges and popular IoT cryptography approaches along with lightweight cryptography and trends in IoT.

Keywords Internet of things (IoT) · Cryptography · Session · Encryption/decryption · Keys

1 Introduction

Research shows that IoT devices (such RFID readers, low-powered IEEE 802.15.4 readers, embedded systems, and wearable computers) are gathering and sensing data and information [1]. The Internet of Things (IoT) business will continue to grow and incorporate more applications into daily life [2]. In the Internet of Things (IoT) world, there are billions of connected devices, giving hackers a significant chance to take over the IoT system. The “IoT” is the concept of data exchange between objects with sensing, processing, or both capabilities across a network (IoT). It allows for the

A. Vidwans (✉) · M. Ramaiya

Institute of Advanced Computing, Sage University, Indore, MP, India
e-mail: vidwans.abhinav@gmail.com

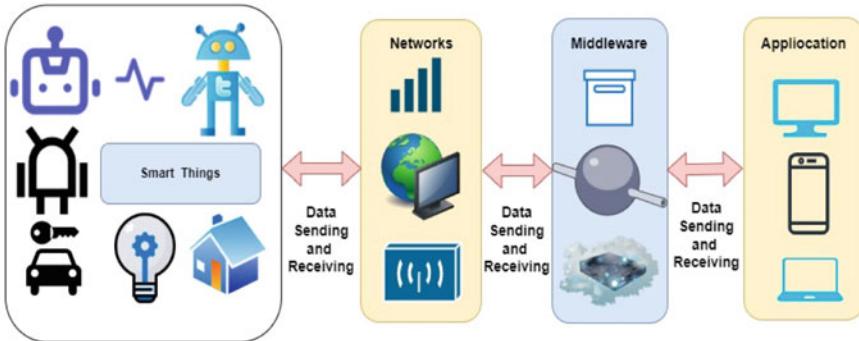


Fig. 1 IoT building blocks

online communication and control of devices with the aforementioned features that are generally integrated into everyday objects (such as an air conditioner, lighting, garage doors, waste management, traffic) [3]. Some of the primary security concerns with IoT are authentication, authorization, privacy, and data confidentiality [4, 5]. IoT device attacks may occur at one or more of the following layers: Hardware, network, and cloud layers are listed in that order [6] (Fig. 1).

Any security settings or keys that are kept inside the IoT device can be retrieved by an attacker once they have physical access to the IoT hardware at the hardware layer. The hacker could build a duplicate or virtual IoT device using the security specifications they've got. The fake Internet of Things device may join a network, send false data to a server, and receive secure user data from a server. There are numerous side channel hacks that let an attacker access the security settings without having physical access to the IoT device. Researchers have created electromagnetic-based side channel attacks to recover the encryption keys for RSA and ECC [7, 8]. IoT devices can have their AES encryption keys stolen through side channel attacks [9–11]. IoT devices' Internet access makes them susceptible to network threats. Uncountable Internet of Things (IoT) devices were externally targeted by the MIRAI malware and turned into zombie nodes that attacked other websites and services on the Internet [12].

However, because technology is being implemented in a wider variety of methods for use in daily life, there is a growing concern about the risk to the privacy of individuals' data. Any required information exchange over a wireless link is expensive to distribute and process, is subject to review, and could even be used against you [13]. These devices require low-power security solutions (Fig. 2).

Sensitive data is routinely protected using encryption to prevent access by unauthorized persons. Asymmetric encryption is the most expensive and most key size-dependent, whereas symmetric encryption is less expensive and less key size-dependent. Additionally, they made note of the possibility that a gadget could sacrifice security in order to save more energy. A significant amount of Internet data is encrypted using symmetric and asymmetric encryption methods, some examples of

Fig. 2 IoT layered architecture



which include TCP/IP, TLS, and IPsec [14]. However, the local area is where wireless IoT device data privacy concerns are the greatest. The entire network may be in danger if adjacent attackers are successful in intercepting confidential data while it is being transmitted from these IoT devices [15]. As a result, the IoT system's structure must have the proper encryption techniques. However, the security of various encryption techniques depends on the keys employed for encoding and decoding.

In order to upsurge the resistance of the resource-constrained devices that make up the majority of IoT systems against different cyberattacks, it is required to evaluate the veracity and integrity of the conveyed information. Attackers can seize control of a server or link phoney servers to draw in network traffic, or they can distort information to lead decision-making authorities astray and harm system functionality [11]. Prior to the commencement of the real data transmission, it is crucial that servers and sensors successfully complete mutual authentication. The authentication process must be quick and secure against intrusions. Elliptic Curve Cryptography (ECC), Pre-Shared Keys (PSK) distribution, One-Way Hash Function, and other lightweight authentication and key establishment procedures have been introduced

in the literature [16, 17], and [18]. This paper's main contributions are a review and highlighting of the various strategies that can be used to provide a basic security framework for IoT.

Only specific situations, such as those involving RFID devices, sensors, contactless smart cards, need for the usage of lightweight cryptography. Personal information leakage and risky actuation jobs can be stopped by peer authentication and secure data transmission. We offer a special delegation architecture for one-way certificate-based authentication [19].

This document is organized as follows for the remaining portions: Sect. 2 provides a review of recent IoT lightweight cryptography technologies to safeguard data and communication between IoT devices and servers. The topic of cryptographic and its methods, with popular IoT encryption methods and challenges in IoT cryptography is covered in Sect. 3. IoT security, lightweight cryptography approaches and trends were presented in Sect. 4. Finally, the paper is concluded, and future work is mentioned in Sect. 5.

2 Literature Survey

In order to provide a safe route, Bhattacharyya et al. [20] proposed a revolutionary cross-layer lightweight solution in 2015. The application and transport layers each have a secure channel of communication. Through the Constrained Application Protocol, secure session establishment is carried out utilizing a payload-contained challenge response technique (CoAP). Application layer data exchanges employ Datagram Transport Layer Security (DTLS) record encryption and Pre-Shared Key (PSK).

In 2016, Yang et al. [21] used several types of user data, such as user ratings and user behaviours, to create a framework for the CF recommender system. We summarize the distinguishing characteristics of these two types of information. Popular CF algorithms are also categorized into memory-based and model-based categories so that their respective strengths and weaknesses may be better understood and compared. The suggested paradigm is shown with the help of two specific cases.

To address the problem of safe and reliable transmission of the large volume data creation and the practical approaches to cope with the on-board data explosion in LTE-Advanced (LTE-A) networks, Qinglei Kong et al. [22] devised a secure handover session key management strategy employing mobile relay. This was done to fix the problem of unsafe and unreliable data transfer during periods of large data output.

Using symmetric and asymmetric cryptography, Henriques et al. [3] proposed a solution in 2017 to secure communication between devices in an IoT system. Combining symmetric and asymmetric cryptography speeds up encryption than of only using an asymmetric cryptographic algorithm. The problem of session key distribution is resolved, and the symmetric encryption method is strengthened by using random keys each time.

In 2017, Sridhar et al. [23] suggested a method that uses lattice-based cryptography to secure broker devices/gateways and cloud services while using lightweight asymmetric cryptography to secure end-to-end devices that safeguard IoT service gateways and low-power sensor nodes. This paradigm addresses network hazards, IoT hardware and software, as well as security issues including confidentiality and privacy. A session key is transmitted between nodes and used for message transmission in the proposed architecture using asymmetric key encryption. The architecture protects against distributed denial of service attacks, quantum algorithm attacks, and eavesdropping. In order to create a key pair for mutual authentication between devices and services, the suggested methodology builds on the distinctive Device IDs of the sensors.

In 2018, Trusit Shah et al. [24] developed a mutual authentication technique that uses a set of keys instead of a single one. (Also used as passphrases). The encrypted vault's contents are sent between the server and the IoT device during the first successful communication session and are subsequently updated during subsequent successful sessions.

Mohammad et al. [25] propose the development of a decentralized authentication infrastructure by the year 2020. The public key certificates are kept in a decentralized manner, while the private keys are kept within the devices themselves. The proposed platform is made up of a client-side module, a server-side wallet management function, and an Ethereum Blockchain network-stored smart contract. Applications can utilize the suggested platform for end device and/or intermediate device authentication as well as secure machine-to-machine (M2M) communication. The creation of secure sessions between various Internet of Things devices demonstrates the practicality of the proposed platform.

3 Background

3.1 *Cryptography*

Cryptography is one of the oldest and most used techniques for safeguarding IT assets. Cryptography is used by almost all businesses to protect sensitive data and IT infrastructure. A collection of rule-based calculations known as algorithms is utilized in the practice of cryptography.

In order to secure data, there are two main methodologies or methods of cryptography:

Encryption: Data encryption is the process of transforming binary data from one form to another using an algorithm and a unique key. In order for encryption to function, an algorithm transforms plaintext into ciphertext, a form that is hard to discern and can only be converted back to plaintext using a cryptographic key. The development of sophisticated encryption methods will improve data transmission security and reduce the likelihood of data compromise.

Decryption: In essence, decryption is the opposite of encryption. A user can decode sensitive data, whether it is at rest or in transit, using a cryptographic key that matches the encryption technique. Encryption and decryption in cryptography both contribute to the improvement of your security posture and the protection of sensitive data, depending on the complexity and resilience of the algorithms you apply.

3.2 IoT and the Popular IoT Encryption Methods

The “Internet of Things” refers to a network of interoperable, self-sufficient computing devices that are capable of communicating with one another and interface with one another through the Internet without the requirement of human involvement. Wireless networks, cloud databases for communication, sensors, data processing software, and networked smart devices make up the architecture of IoT systems in most cases.

Through a variety of methods and procedures, the physical objects, technology, operations, and networks that comprise an IoT ecosystem are shielded against a range of IoT security intrusions. IoT security’s two main objectives are to:

1. Verify the safety of data over its entire lifecycle, from collection to processing to storage to transmission.
2. Identify and address vulnerabilities in IoT component parts.

Technologies related to the Internet of Things are utilized in a range of business sectors, some of which are manufacturing, agriculture, transportation, logistics, and health care. However, security issues are also brought up by the growth. The ability to communicate vast volumes of data, frequently sensitive data, every second is a feature of new IoT devices.

Data and IoT devices are secure with cryptography

If the data is encrypted from beginning to finish, only the sender and the recipient to whom it is addressed will be able to access it as it is transferred from one device to another, even if the data has to travel over the Internet to get there. Therefore, if the right encryption standards are in place, even the creator of a particular gadget won’t have access to that information. Businesses that handle highly sensitive data and want to keep that data secure must use this type of protection (Fig. 3).

Over the next few years, it’s likely that the number of IoT devices in use will increase quickly. Sensors and other devices will probably become even more useful as Industry 4.0 technology advances, which will encourage businesses to adopt the technology. Cryptography will probably become much more crucial as a result of this trend. If end-to-end encryption is not employed, the data that is sent between the different Internet of Things devices will not be protected, leaving it vulnerable to eavesdropping and manipulation.

- **Use of cryptography in Internet of Things devices**

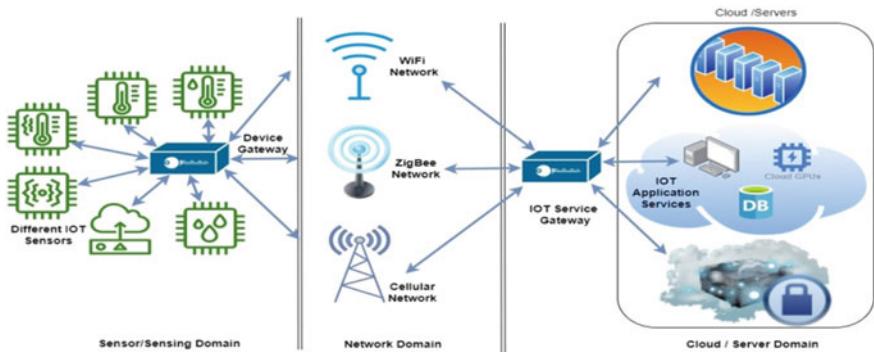


Fig. 3 Intelligent security lightweight framework for IoT

Manufacturers can pick from a few different cryptographic standards. The Advanced Encryption Standard is the industry standard for encryption, and it is used by the majority of data protection tools, governments, and security agencies. However, not all manufacturers are satisfied that well-known encryption methods, like the AES, are suitable for use with IoT devices. There are two distinct keys that are mathematically connected to one another. These keys are referred to as the “public key” and the “private key.” The provision of a higher level of safety is one of the benefits that come with using asymmetric encryption.

- **Advanced Encryption Standard (AES):** The AES encryption technique for the Internet of Things has grown in popularity. This is as a result of its simple implementation in hardware- and space-constrained settings. When this symmetric encryption technique was declassified, it was stated that it would “be capable of protecting critical government information well into the next century.” When implemented in 128-bit form, AES is incredibly effective.
- **Data Encryption Standard (DES):** One of the first encryption methods is the Data Encryption Standard (DES). This symmetric encryption technique was first proposed by IBM in 1976 to safeguard confidential and unclassified data. The DES is no longer in use despite being regarded as the foundation of cryptography. This is because numerous security researchers were able to break it. As a result, AES took its place. Due to DES’s short encryption key length, which left it open to brute-force cryptographic attacks, it failed.
- **Triple DES:** The Triple DES followed the DES. The original DES algorithm was intended to be replaced with this one. It uses three separate keys, each with a length of 56 bits. The key length is 168 bits long in total. Experts disagree with this, claiming that a key strength of 112 bits is more accurate.
- **RSA Algorithm:** In 1977, the Rivest–Shamir–Adleman (RSA) algorithm was created. Additionally, it is regarded as the most popular asymmetric encryption algorithm. Users can send encrypted messages without having to provide the recipient’s code first. It is therefore incredibly secure. The RSA technique has obvious security advantages, but it is also very scalable. It is available with different

encryption key lengths, such as 768-bit, 1024-bit, 2048-bit, and 4096-bit (and more) [33].

- **DSA:** Another asymmetric encryption scheme is the Digital Signature Algorithm (DSA). The National Institute of Standards and Technology (NIST) first suggested the DSA in 1991. The DSA is quicker at decryption and verification than the RSA, even though it may be slower at encryption and signature.
- **Blowfish and two fish:** Another symmetric encryption method created to take the role of DES is Blowfish and Two fish. Both are the most widely used, open-source symmetric encryption algorithms, with Two fish being the replacement for Blowfish. Plain text is broken down into 64-bit data blocks by Blowfish and 128-bit data blocks by Two fish (in 16 rounds irrespective of key size), and encrypts each block separately.
- **ECC:** Elliptical curve cryptography ECC is a substitute for Rivest–Shamir–Adleman (RSA) and uses algebraic operations to build actual security between “key pairs” (public and private keys). ECC creates keys for encryption and decryption that are shorter, faster, and more effective by utilizing the elliptic curve theory. For IoT devices, mobile applications, and those with constrained computing (CPU) capabilities, it is the ideal option.

3.3 The Challenges of IoT Cryptography

- **Software and firmware flaws:** Because many smart devices have constrained resources and computing power, it might be challenging to secure the security of IoT systems. As a result, they are less able to perform robust, resource-intensive security operations and are more susceptible than non-IoT devices.
- **Vulnerable transmissions:** The vast majority of the currently available security solutions are challenging to install on IoT devices owing to the limited resources available on these devices. As a result, traditional security techniques are less effective in protecting IoT device communication.
- **Information breaches through IoT systems:** We’ve already shown that hackers are able to intercept unencrypted messages and steal the data that your IoT system analyses. It’s possible that this will also include information about your location, finances, and medical history. Even though this is not the sole method, attackers can get important information by utilizing communications that are not adequately protected.
- **Perils caused by malicious software:** Set-top boxes, smart TVs, and smart-watches pose the biggest risk to users from malware attacks, according to a recent Zscaler analysis. If criminal actors are successful in injecting malware into the system, an Internet of Things system may have its functionality changed, personal data may be gathered, and other assaults may be carried out. If the makers of particular gadgets don’t adhere to the appropriate software security requirements, some of the devices can even come pre-loaded with viruses. Several companies have previously successfully fought off the most well-known IoT-targeted malware.

- **Security breaches:** Additional to the above-mentioned malware and MITM attacks, other cyberattacks may target IoT systems.

4 IoT Networking and Security (Lightweight Security LWS)

Lightweight encryption may be required for IoT devices with very limited computational capacity, such as Internet-connected microcontrollers in large machinery. New, lightweight encryption standards for Internet of Things devices have not yet been implemented on a large scale since, for the most part, in the past there hasn't really been a demand for that type of security protection. While programmers, cryptographers, and cybersecurity specialists are familiar with the advantages and disadvantages of AES, they won't be as familiar with a new encryption standard. This might increase the vulnerability of IoT devices that use these new standards. It may be very difficult to wait for portable encryption technology to become accessible. It may be difficult to adapt current Internet of Things technology with new security standards on a broad scale, which may necessitate the replacement of obsolete equipment or the acceptance by owners of an adequate level of data protection.

Sr. No.	Network	Connectivity	Pros and cons	Popular use cases
1	Ethernet	Wireless, small range	Pros—security, high speed Cons—range limited to wire length, limited mobility	Stationary IoT: Video cameras, games consoles, fixed equipment
2	WIFI	Wireless, limited range	Pros—high speed, great compatibility Cons—limited range, high power consumption	Smart home, devices that can be easily recharged
3	NFC	Wireless, ultra-short range	Pros—reliability, low-power consumption Cons—limited range, lack of availability	Payment systems, smart home
4	Bluetooth low-energy	Wireless, short range	Pros—high speed, low-power consumption Cons—limited range, low bandwidth	Small home devices, wearable, beacons
5	LPWAN	Wireless, long range	Pros—low range, low-power consumption Cons—low bandwidth, high latency	Smart home, smart city, smart agriculture (field monitoring)

(continued)

(continued)

Sr. No.	Network	Connectivity	Pros and cons	Popular use cases
6	ZigBee	Cellular, limited range, wireless	Pros—scalability, low-power consumption Cons—limited range, compliance issues	Home automation healthcare and industrial sites
7	Cellular network	Wireless, extensive range	Pros—high speed, nearly global coverage, reliability Cons—high power consumption, high cost	Drones sending video and images

- Lightweight cryptography requirements

For lightweight cryptography to work, the implementation must have the following elements.

- Dimensions (ROM/RAM sizes, circuit sizes).
- Power.
- Power usage.
- Processing Speed (throughput, delay).

The component's size is the most important consideration when determining whether or not a certain device can utilize it. For RFID and energy-harvesting devices, power consumption is just as crucial as it is for battery-powered devices. A high throughput is critical for devices that transmit a great deal of data, such as cameras or vibration sensors, yet a low latency is required for the real-time control processing of a vehicle management system, etc. Because the strength of the encryption method is so dependent on the hardware, such as the size of the circuits or CPU, the size serves as a baseline for both the strength and weight of the encryption methodology.

Lightweight cryptography trends

The National Institute of Standards and Technology (NIST), which is in charge of publishing standards for cryptographic technology, was the organization that initially presented the idea of lightweight cryptography. The block cypher known as “PRESENT,” which was published in 2007, is acknowledged as the originator of “lightweight cryptography.” The small circuit size of it allows its implementation in RFID tags, which is not achievable with standard AES encryption. The lightweight block cypher SIMON/SPECK was made public by the National Security Agency (NSA) in 2013 with the intention of achieving international standards. It has a very small ROM capacity suited for a restricted microprocessor.

The term “authenticated encryption” is used to describe a mode of operation for block ciphers that allows for both the encryption and authentication of messages. Given the importance of incorrect data detection in the Internet of Things, it is projected that encryption will eventually relate to verified encryption. Even when

the same block cypher is used, the efficacy and security vary significantly depending on how it is implemented as an authorized encryption. There were 60 submissions given the circumstances.

For authenticated encryptions like AES-CCM/GCM, which are recommended by NIST, the calculation required is equivalent to that for secrecy-preserving encryption but twice as massive. The theoretical cap for authenticated encryptions is the computation required to achieve encryption alone because authenticated encryption takes more computation than encryption alone. Although OCB is a potentially limit-clearing authenticated encryption, it requires a block cypher decryption process to decrypt data. OTR2, the first authenticated encryption ever created, reaches the theoretical computation limit just by using block cypher encryption methods.

5 Conclusion and Future Work

The Internet of Things (IoT) is a cutting-edge innovation with a noble goal: to enhance human well-being through the seamless interconnection of various intelligent systems, appliances, and programs. It is based on the idea of a connected world where everything is connected (a person, a thing, or a device). It does, however, also have a number of security and privacy concerns. This study provided a high-level review of this technology's foundations, as well as its data and communication security in IoT. We reviewed the literature to find out what safeguards are currently in place to protect IoT infrastructure, and we summarized these security measures to show how they meet security concerns in the IoT. From a security standpoint, the article will be helpful to both researchers and those creating IoT applications. In our upcoming efforts, we will effectively integrate blockchain technology, artificial intelligence approaches, and lightweight cryptosystems to deliver a smart IoT security model.

References

1. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of things (IoT): a vision, architectural elements, and future directions. *Fut Gener Comput Syst* 29(7):1645–1660
2. Vasseur J-P, Dunkels A (2010) Interconnecting smart objects with IP: the next internet. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA
3. Henriques MS, Vernekar NK (2017) Using symmetric and asymmetric cryptography to secure communication between devices in IoT. *Int Conferen IoT Appl (ICIOT) 2017:1–4.* <https://doi.org/10.1109/ICIOTA.2017.8073643>
4. Dheresh S, Atish M, Satyendra ST, Nishant C (2011) Applying frequent pattern mining in cloud computing environment. *Int J Adv Comput Res (IJACR)* 1(2):84–88
5. Riahi A, Challal Y, Natalizio E, Chtourou Z, Bouabdallah A (2013) A systemic approach for IoT security. In: *Distributed computing in sensor systems (DCOSS), 2013 ieee international conference on IEEE*, pp 351–355

6. Jing Q, Vasilakos AV, Wan J, Lu J, Qiu D (2014) Security of the internet of things: perspectives and challenges. *Wireless Netw* 20(8):2481–2501
7. Genkin D, Pachmanov L, Pipman I, Tromer E (2015) Stealing keys from PCs using a radio: cheap electromagnetic attacks on windowed exponentiation. In: Proceedings of the workshop on cryptographic hardware and embedded systems (CHES 2015). Springer, pp 207–228
8. Genkin D, Pachmanov L, Pipman I, Tromer E (2016) ECDH key-extraction via low-bandwidth electromagnetic attacks on PCs. In: Proceedings of the cryptographers' track of the RSA conference (CT-RSA 2016), Springer, pp 219–235
9. Craig R, Jasper L (2015) TEMPEST attacks against AES
10. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Lulwah MA, Sachin K (2023) Survivability of industrial internet of things using machine learning and smart contracts. *Comput Electr Eng* 107:108617, ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
11. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. *Trans Emerg Tel Tech* e4758. <https://doi.org/10.1002/ett.4758>
12. Rohit S, Rajeev A (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. *Comput Electric Eng* 108:108715, ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
13. Deng H, Hu J, Sharma R, Mo M, Ren Y (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. *Comput Commun* ISSN 0140–3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
14. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM (2023) An efficient hybrid deep learning model for denial of service detection in cyber physical systems. In: IEEE transactions on network science and engineering. <https://doi.org/10.1109/TNSE.2023.3273301>
15. Gupta U, Sharma R (2023) Analysis of criminal spatial events in India using exploratory data analysis and regression. *Comput Electr Eng* 109:Part A108761, ISSN 0045-7906. <https://doi.org/10.1016/j.compeleceng.2023.108761>
16. Goyal B et al (2023) Detection of fake accounts on social media using multimodal data with deep learning. In: IEEE transactions on computational social systems. <https://doi.org/10.1109/TCSS.2023.3296837>
17. Praveen Malik S, Sharma R, Ghosh U, Alnumay WS (2023) Internet of things and long-range antenna's; challenges, solutions and comparison in next generation systems, microprocessors and Microsystems. 104934:ISSN 0141-9331. <https://doi.org/10.1016/j.micpro.2023.104934>
18. Vohnout R et al (2023) Living lab long-term sustainability in hybrid access positive energy districts -a prosumager smart fog computing perspective. In: IEEE internet of things journal. <https://doi.org/10.1109/JIOT.2023.3280594>
19. Soni D, Sharma V, Srivastava D (2019) Optimization of security issues in adoption of cloud ecosystem. In: 2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU), pp 1–5. <https://doi.org/10.1109/IoT-SIU.2019.8777670>
20. Bhattacharyya A, Bose T, Bandyopadhyay S, Ukil A, Pal A (2015) LESS: lightweight establishment of secure session: a cross-layer approach using CoAP and DTLS-PSK channel encryption. In: 2015 IEEE 29th international conference on advanced information networking and applications workshops, pp 682–687. <https://doi.org/10.1109/WAINA.2015.52>
21. Yang Z, Wu B, Zheng K, Wang X, Lei L (2016) A survey of collaborative filtering-based recommender systems for mobile internet applications. *IEEE Access* 4:3273–3287. <https://doi.org/10.1109/ACCESS.2016.2573314>
22. Kong Q, Lu R, Chen S, Zhu H (2017) Achieve secure handover session key management via mobile relay in LTE-advanced networks. *IEEE Int Things J* 4(1):29–39. <https://doi.org/10.1109/JIOT.2016.2614976>
23. Sridhar S, Smys S (2017) Intelligent security framework for IOT devices cryptography based end-to-end security architecture. *Int Conferen Invent Syst Control (ICISC)* 2017:1–5. <https://doi.org/10.1109/ICISC.2017.8068718>

24. Kumari A, Kumar V, Yahya Abbasi M, Alam M (2018) The cryptanalysis of a secure authentication scheme based on elliptic curve cryptography for IOT and cloud servers. In: 2018 international conference on advances in computing, communication control and networking (ICACCCN), pp 321–325. <https://doi.org/10.1109/ICACCCN.2018.8748591>
25. El-Hajj M, Fadlallah A, Chamoun M, Serhrouchni A (2019) Ethereum for secure authentication of IoT using pre-shared keys (PSKs). Int Confereen Wireless Netw Mobile Commun (WINCOM) 2019:1–7. <https://doi.org/10.1109/WINCOM47513.2019.8942487>

Fake News Detection Using Data Science Approaches



Lina Shugaa Abdulzahra and Ahmed J. Obaid

Abstract In today's internet-driven world, fake news is a problem that is only becoming worse. Given the ease of exchanging information online, separating false information from reliable information is a crucial endeavor. Using bag-of-words and consecutive mining approaches, we provide a data mining solution in this work to categorize articles as genuine or fake. We also compare the accuracy of the solution for identifying fake news across different datasets. Our method first purifies the input information by normalizing words and eliminating "filler" words. The cleansed data is then vectorized using sequential mining techniques. After that, it uses vectorized data to train the classification models and categorizes unknown news as authentic or fake. Assessment of our technology to mine and categorize bogus news using actual data demonstrates its viability. The classification algorithms are then trained using vectorized data to categorize unreported news as real or bogus. The effectiveness of our technology in identifying and categorize bogus news has been evaluated using real-world data.

Keywords Fake news · Detection · Data science · Analysis

1 Introduction

People are increasingly turning to social media for news as online communication grows. Social media is replacing traditional media as a news source. Social media's simplicity, accessibility, and interconnectivity caused this change. Social media makes staying informed easy with quick access to news, videos, and discussions. Social media offers real-time updates, user-generated material, and customized news feeds. This shift toward social media as a news source mirrors our digital age,

L. S. Abdulzahra · A. J. Obaid (✉)

Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq

e-mail: ahmedj.aljanaby@uokufa.edu.iq

L. S. Abdulzahra

e-mail: linas.alabdali@student.uokufa.edu.iq

when social media is part of our daily life [1–3]. These social media platforms' basic characteristics provide the following explanations for this change in consumption habits: News on social media is often more timely and cheaper to consume than in more traditional forms of journalism like newspapers or television. Sharing, talking about, and debating the data with coworkers or other users on social media are also much easier [4–7]. For instance, 62% of American adults received news via social media in 2016, compared to only 49 percent in 2012.

Furthermore, social media has surpassed television as the primary source of news for many individuals. Despite the advantages offered by social media platforms, the quality of news reports shared on these networks tends to be lower compared to traditional news agencies. This discrepancy arises due to the proliferation of fake news, which involves intentionally spreading false information for various reasons, such as financial gain or political influence. Generating and promoting news online are cost-effective, rapid, and easily facilitated through social media channels. A notable example is the Pizzagate incident, where approximately 1 million tweets were connected to the dissemination of fake news after the conclusion of the presidential election [7–9].

The prevalence of fake news has become a widespread phenomenon, to the extent that it was even awarded the 2016 Macquarie Dictionary Word of the Year. The extensive spread of false information can have severe detrimental effects on individuals and society as a whole. Firstly, it disrupts the delicate balance of authenticity within the information ecosystem. For instance, during the 2016 U.S. presidential election, false news stories were found to have circulated more widely on Facebook compared to legitimate mainstream news sources. Secondly, deceptive information aims to intentionally influence readers and lead them to adopt biased or erroneous viewpoints [8, 9]. The impact of fake news poses significant challenges to fostering an informed society and maintaining the integrity of information dissemination. Fake news is often employed by propagandists as a means to disseminate misleading information or exert political influence. There have been reports suggesting that Russia, for instance, has utilized fabricated accounts and social media bots to propagate false narratives. The impact of fake news extends beyond its immediate dissemination; it also influences how people interpret and respond to genuine news. Some instances of fake news are intentionally created to sow mistrust and confusion, making it challenging for individuals to distinguish truth from falsehood [10]. To mitigate the harm caused by false information and safeguard the general public and the news ecosystem, it is crucial to develop tools that can swiftly detect and identify fraudulent news articles shared on social media platforms.

2 Related Study

Coronavirus illness of 2019 (COVID-19), as stated by Gupta et al. [11], is currently known as COVID-19. It is crucial to understand the makeup and features of COVID-19 as it not only results in modified individual beliefs and behavior shifts, including

such irrational prevention strategies actions, but also poses a potential threat to everyone's safety and well-being. To better understand COVID-19, we employ First Amendment jurisprudence, text and efficacy, and a split-method approach. Latent Dirichlet Allocation (LDA) is used in the first pillar to classify COVID-19 news articles as either fake or real. The final component compares and contrasts how fake and real news affects people differently. According to the research by Raza et al. [12], fake media is a serious issue today since it is both pervasive and hard to spot. The challenge of spotting fake news when it is still in its infancy is pressing. Another issue with identifying fake news is the lack of labeled data to train detection methods. Jarrahi et al. [13] present a novel method for identifying fake news that gets around these problems. Our proposed method employs metadata for both media pieces and social contexts to identify fake news. The proposed model is constructed on a Modular framework, which is itself split into two parts: an encode, which derives useful abstractions from erroneous sources of information, and a receiver, which predicts behavioral intents based on historical data.

3 Proposed Model

This article clarifies the system, which has been created in three sections. The first part is static and makes use of a machine learning classifier. We did our investigation and made the most effective classifier which was chosen for the last run using a model consisting of four different classifiers. The second part is dynamic and searches the internet for data on the probability that the news is true using the user's keyword or text. The final paragraph attests to the reliability of the consumer URL.

For this study, we have utilized Python and its Sci-Kit packages. Python provides a wide range of modules and add-ons that can be used to automate machine learnings tasks. As practically all machine learning algorithms are freely available for Python, the Sci-Kit Learn module is their best source, making it easy and quick to evaluate ML methods. We chose Django for the model's web-based implementation because it provides client-side implementation using HTML, CSS, and Java string.

3.1 *Design and Architecture of the Model*

(i) Static Search Framework

The static part of the architecture of the fake news detection system is quite simple and was built with the standard machine learning workflow in mind. The below diagram of the system's layout explains everything. The following are the main techniques employed during the design process.

(ii) Dynamic Analysis

The second search area on the website asks for particular keywords to be looked up online, and it then returns an appropriate result with the likelihood

of that term being in an article or an article with similar content that uses those keywords.

(iii) URL Analysis

Users can enter the domain name of a certain website into the website's third search form, after which the implementation searches for that website in either the database of prohibited sites or our database of actual sites. The databases of actual websites keep a record of the domain names of websites that frequently provide truthful and dependable news reports and vice versa. If neither database contains the website, the code will only notice that the news aggregator has been removed rather than classifying the domains.

Many websites, including social media platforms, search engines, news agency homepages, and fact-checking internet sites, allow us to access online news.

There really are a few readily viewable databases for identifying fake stories on the web, including those from Buzzfeed News, BS Detector, and others. For assessing the truthfulness of news, these datasets have been extensively used in numerous research studies. I have included a quick overview of the origins of the information utilized in this research in the parts that follow. Search results, websites for social networks, and the homepages of news organizations are just a selection of web-based media resources. Annotators with domain expertise who scrutinize statements and supplementary evidence are needed for the arduous work of manually determining whether news is accurate or incorrect. The model's overall layout is depicted in Fig. 1.

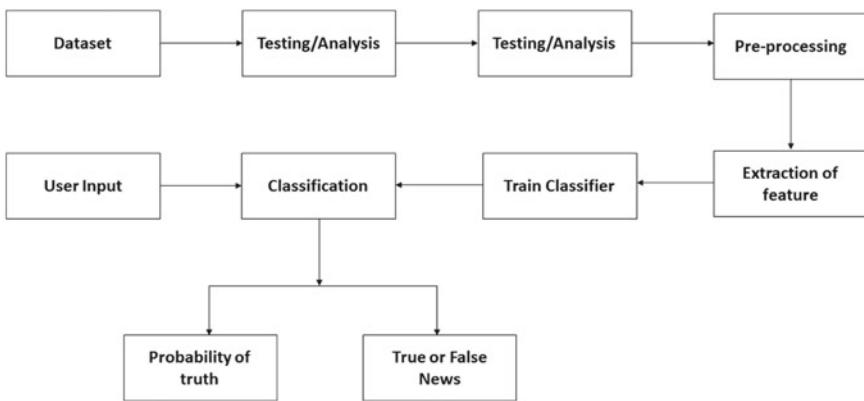


Fig. 1 Architecture of the proposed model

4 Result Analysis

We used natural language processing in the initial search box to obtain an exceptional answer. As a result, we created a classifier that can spot fake news based on the phrases used in newspaper articles. Our framework provides techniques such as Count Vectorization and High-frequency sub-data compression before putting the data into a Condensing Classifier to determine the legitimacy of content as a percent chance.

The correct search field of the platform makes a request that certain terms can be located on the internet. After that, it offers an appropriate result for the percentage likelihood of those phrases actually being included in an item or a linked post that has those search connections in it [14, 15].

Because the execution searches for the sites in both the authentic sites' directory and the banned sites' directory, the third search field on both sites requires a specific website domain name to be entered. The genuine sites' database comprises domains that consistently deliver reliable and genuine news and vice versa. If the website is not discovered for either index, the application does not somehow classify it; instead, it simply asserts that the searchable archive does not function. Table 1 shows the confusion matrix of the statement which gets the probability of finding whether the statement is false or true.

By using the above values, precision, accuracy, recall, and F1 score have been calculated.

$$\text{Accuracy} = \frac{Tp + Tn}{Tp + Fp + Tn + Fn},$$

$$\text{Precision} = \frac{Tp}{Tp + Fp},$$

$$\text{Recall} = \frac{Tp}{Tp + Fn},$$

$$F1 \text{ score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

The experimental dataset has been taken from MATLAB and stated in Table 2.

The values of precision, recall, F1-score, and accuracy have been calculated in Table 3.

The graphical analysis of confusion matrix is shown in Fig. 2. The experimental outcome is briefly explained in Table 3 and Fig. 3.

Table 1 Confusion matrix

Sl. No.	Predication class 1	Predication class 2
Accurate class 1	True positive (Tp)	False negative (Fn)
Actual class 2	False positive (Fp)	True negative (Tn)

Table 2 Confusion matrix for experimental dataset

Sl. No.	Predication class 1	Predication class 2
Accurate class 1	1588	2801
Actual class 2	1062	4800

Table 3 Experimental observation

Parameter	Outcome
Precision	0.67
Accuracy	0.59
Recall	0.79
F1 score	0.71

Confusion matrix

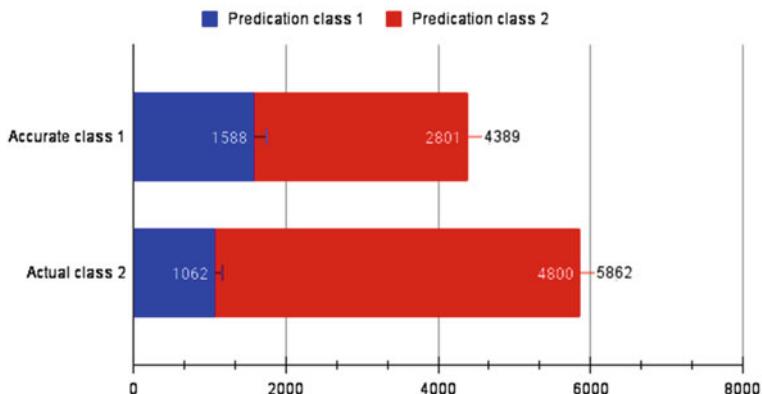


Fig. 2. Graphical analysis of confusion matrix

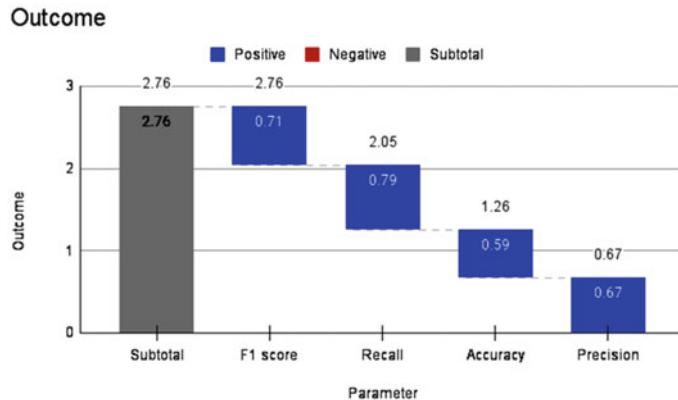


Fig. 3 Experiment outcome analysis using graphical methods

5 Conclusion

The preponderance of tasks is completed online within the twenty-first century. Publications that were formerly favored as online media articles and programs like Facebook and Twitter are gradually replacing backup copies. The forwards on WhatsApp are another important source. Complicating matters further, fake news aims to influence or affect people's views toward using digital technologies, and this problem is only becoming worse. When genuine news misleads a person, one of two things may occur: First, they may begin to believe that their impressions of a certain subject are accurate. Therefore, we created our Fake Media Detection mechanism, which accepts user input and categorizes it as true, to stop the situation. Our best system was linear in the parameters, which had an effectiveness of 76%, as shown for the above dynamic search. To enhance the efficacy of regression analysis, we employed grid search optimization, which provided us with an efficiency of 72%. So, one can state that there are 72% possibilities that a person would successfully classify a given news story or its header according to its real character if they submit it to this algorithm. The viewer can research news articles, search terms, and the legitimacy of websites online. With each cycle, the dynamic program's accuracy rises to 81%. We intend to create our individual data, that will be updated in accordance with the most recent news.

References

1. Amahan PA (2023) The perspective of data mining: the study of fake news on social media. *Dyn J Pure Appl Sci* Ozamiz city, Philippines
2. Allein L, Moens M-F, Perrotta D (2023) Preventing profiling for ethical fake news detection. *Inf Process Manage* 60(2):103206

3. Prasee A, Rodrigues J, Santhi Thilagam P (2023) Hindi fake news detection using transformer ensembles. *Eng Appl Artific Intell* 119:105731
4. Tyrankiewicz A, Jahankhani H (2023) The role of blockchain to reduce the dissemination of fake news on social media and messaging platforms. In: *Cybersecurity in the age of smart societies: proceedings of the 14th international conference on global security, safety and sustainability*, London, September 2022, Cham, Springer International Publishing
5. Adrian G (2023) Towards detecting fake news using natural language understanding and reasoning in description logics. In: *Measuring ontologies for value enhancement: aligning computing productivity with human creativity for societal adaptation: first international workshop, move 2020, virtual event, October 17–18, 2020, revised selected papers*. Cham, Springer Nature Switzerland
6. Oliver B et al (2023) Automation on twitter: Measuring the effectiveness of approaches to bot detection. *Soc Sci Comput Rev* 41.1:181–200
7. Andrea P, Harris E, Van Bavel JJ (2023) Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news. *Group Processes Intergroup Relat* 26.1:24–47
8. Daniil B et al (2023) Using open government data to facilitate the design of voting advice applications. In: *Electronic participation: 14th IFIP WG 8.5 international conference, ePart 2022, Linköping, Sweden, September 6–8, 2022, proceedings*. Cham, Springer Nature Switzerland
9. Sandeep S, Brown S, Brandimarte L (2023) Understanding the role of nonverbal tokens in the spread of online information
10. Salsabila G, Latif MFF (2023) Filtering communication media as an antidote to the spread of hoax news: Study of Takhrij and Syarah Hadith. *J Takhrij Al-Hadith* 2(1):11–20
11. Ashish G et al (2022) Understanding patterns of COVID Infodemic: a systematic and pragmatic approach to curb fake news. *J Bus Res* 140:670–683
12. Raza S, Ding C (2022) Fake news detection based on news content and social contexts: a transformer-based approach. *Int J Data Sci Anal* 13(4):335–362
13. Jarrahi A, Safari L (2023) Evaluating the effectiveness of publishers' features in fake news detection on social media. *Multimedia Tools Appl* 82(2):2913–2939
14. Jawad ZA, Obaid AJ (2022) Combination of convolution neural networks and deep neural networks for fake news detection. arXiv preprint [arXiv:2210.08331](https://arxiv.org/abs/2210.08331)
15. Naaz S, Rao H, Aggarwal P, Obaid AJ (2023) 5G as a new phase of wireless network technology. In: Sharma DK, Peng SL, Sharma R, Jeon G (eds) *Micro-electronics and telecommunication engineering . lecture notes in networks and systems, vol 617*. Springer, Singapore. https://doi.org/10.1007/978-981-19-9512-5_31

Reversible Data-Hiding Scheme Using Color Coding for Ownership Authentication



Anuj Kumar Singh, Sandeep Kumar, and Vineet Kumar Singh

Abstract The concept of reversible data hiding (RDH) enables the full restoration of the cover image while simultaneously recovering the concealed data from a previously obscured image. Hence, it is the favored choice when complete restoration of the cover image is required in situations where the concealment of critical data is important. This work introduces a system that utilizes interpolation-based color coding method (CCM) to temporarily conceal data without permanent deletion. The refinement of the expanded form of the original image is achieved by the utilization of two distinct methods, namely enhanced neighbor mean interpolation (ENMI) and modified neighbor mean interpolation (MNMI). This procedure is conducted prior to the inclusion of any confidential data. The experimental findings indicate that the suggested approach has the potential to be implemented and exhibits superior performance compared to standard methods in relations of highest signal-to-noise ratio and data whacking size.

Keywords Reversible data hiding (RDH) · Color coding method (CCM) · Interpolation

1 Introduction

Data concealing is the practice of covering sensitive information with another digital media, so that no one would have any cause to suspect the data's existence [1]. Possible uses for this technique include authentication [2], copyright protection [3], verification of content ownership [4], and the transmission of patient information [5]. There are two broad categories for data concealment techniques: irreparable

A. K. Singh (✉)

School of Computing Science and Engineering, Galgotias University, Greater Noida, India
e-mail: mail2anujji@gmail.com

S. Kumar · V. K. Singh

Department of CSE (AI), ABES Institute of Technology Ghaziabad, Ghaziabad, India
e-mail: vineet.singh@abesit.edu.in

and rescindable. In the case of irreparable data hiding, the recipient determination is being bright to regain the unique secret message but not the unique cover media. As an added bonus, RDH facilitates the precise recovery of both with zero distortion. Some fields, including health, the military, and politics, require a lossless restoration of the original medium after the embedded data has been extracted. In this kind of predicaments, RDH is invaluable. The goals of many picture-based RDH systems include increasing the total quantity of embedded data, reducing the amount of distortion induced into the steganographic image, and, in some circumstances, reducing the complexity of the processing needs of the proposed scheme. The information is encoded within a compressed version of the image using compression domain techniques. Using the stego-image to recover the compressed version of the image rather than the original image is necessary due to the lossy nature of various compression methods. Subsequently, the output coefficients are subjected to an inversion process, resulting in the generation of a steganographic depiction that represents the original image. Two frequently used transformations are the distinct cos transmute [6] and the distinct ripple transform [7]. The embedding procedure, similar to spatial domain methods, functions directly on the individual pixels of the image. The spatial domain-based reversible data-hiding (RDH) approaches can be categorized into four separate types: pixel-value-ordering (PVO) schemes [8], difference expansion (DE) schemes [9], histogram shifting (HS) schemes [10], and interpolation-based schemes. The objective of this study is to propose an Information Redundancy and Data-Hiding (IRDH) technique that can effectively improve the quality of images while simultaneously improving the capacity for embedding data. In order to accomplish this, we can integrate the most efficacious interpolation technique and embedding framework as documented in the scholarly literature.

2 Literature Review

In this article, we survey both existing interpolation methods and those that have been proposed in the academic literature, which are dependent on interpolation to accomplish RDH. The overall procedures of IRDH schemes are depicted in Fig. 1. In order to create a low-resolution image, these methods commonly begin by lowering the size of the high-resolution input image by a factor of four [11, 12]. First, we take the original image, also known as the source image, and interpolate it to make it four times larger. The resulting picture might be referred to as a cover image or an interpolated image. There are two types of pixels in this interpolated image: the original, unaltered pixels, and the interpolated pixels. The initial intention was to utilize the NMI interpolation method which was proposed [11] for the implementation of the innovative IRDH scheme. The dimensions of each image block were increased by a factor of 2 using the nearest neighbor interpolation method (NMI), resulting in a transformation from the inventive 2×2 pixel size to a new size of 3×3 pixels. The interpolation method known as INP was introduced by Lee and Huang [13] as a means of improving the NMI approach. The worth of the inserted

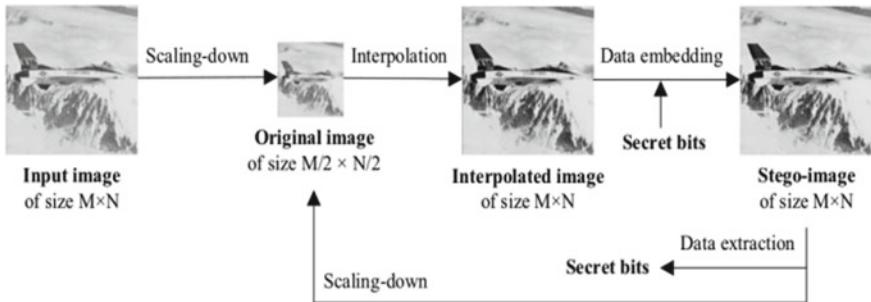


Fig. 1 Block diagram of IRDH schemes

pixel is resolute by calculating a weighted mean of the original pixels, where one of the pixels is given greater attention. The value assigned to each interpolated pixel was determined by calculating the mean of the values assigned to it in both the horizontal and vertical interpolation directions. The INP method demonstrates superior image quality compared to the NMI method. Researchers introduced the ENMI outburst method as a means to improve upon the NMI interpolation methodology. The ENMI technology resulted in an enhancement of the image quality. In 2017, Malik et al. [14] proposed the modified neighbor mean outburst (MNMI) scheme as a means to enhance the image quality of the NMI approach. In this study, we assigned greater significance to the estimation of the horizontal and vertical interpolated pixels compared to previous approaches. In the year coinciding with [14], they introduced the technique of parabolic interpolation. The implementation of the aforementioned strategy was envisaged as an integral component of their Information Resource Development and Management (IRDH) strategy. The interpolated picture was partitioned into blocks measuring 1 by 5, with each block consisting of five contiguous pixels organized in a linear, columnar, or diagonal configuration. Each outburst block consisted of a total of five pixels, with two incorporated pixels sandwiching three original pixels. A parabolic equation was derived by utilizing the values of the three initial pixels and their respective positions within the block. In the year 2020, a novel interpolation approach was suggested by Malik et al. [15], which was shown to outperform the 1D [16] and 2D [17] parabolical outburst techniques. The term MNMI was utilized; however, to prevent any potential ambiguity with the existing MNMI, it will be referred to as MNMI2 [14].

3 Proposed Scheme

Most frequently, the procedure that has been proposed is carried out in two stages, known as (1) data hiding and (2) data extraction. The information is provided below.

3.1 Data-Hiding Process

The details embedding procedure has been illustrated as follows.

1. Initial consideration is given to the Cover Image (CI), Secret Information (SI), and Key (K) as inputs.
2. Generate interpolated cover image ICI using MNMI scheme.
3. Separate the color components of ICI and generate I^R_{CI} I^G_{CI} I^B_{CI} .
4. A bits stream (BS) is generated from the SI.
5. Generate authentication key (AC) by applying SHA-512 on BS.
6. Now consider 8 bits BS and store it to D_{BS} after converting to decimal format.
7. Now, D_{BS} will have three digits, namely α , β , and γ .
8. Store two secret bits in BS_1 and BS_2 after obtaining them from BS.
9. To generate a 512 bit key stream (κ), enter the common secret key (K) and run the SHA-512 hash algorithm.
10. Then embed SI into the interpolated pixels of ICI in a clockwise direction if κ is 0, otherwise embed in anti-clockwise direction.
11. Embed α into the two LSB bits of I^R_{CI} .
12. Embed β into the I^G_{CI} by resetting the last digit zero and then using the rule $I^G_{CI} + (10 - \beta)$
13. Embed γ into the I^B_{CI} by resetting the last digit zero and then using the rule $I^B_{CI} + (10 - \gamma)$.
14. Embed two bits AC into the middle portion of the interpolated pixel blocks
15. Hence new RGB pixels are generated
16. Thus SI is embedded into the CI and generates stego-image I_{SCI} .

3.2 Data Abstraction Process

The extraction of secret image is illustrated as follows.

1. At first, stego-image I_{SCI} and a shared secret key (K) are considered as an input.
2. From input shared secret key (K) and generate SHA-512 hash algorithm to generate 512 bit key stream (κ).
3. Separate the color components of I_{SCI} and generate I^R_{SCI} I^G_{SCI} I^B_{SCI} .
4. Then extract SI from the interpolated pixels of I_{SCI} in a clockwise direction if κ is 0 otherwise extract in anti-clockwise direction.
5. Now, collect three digits I^R_{SCI} I^G_{SCI} I^B_{SCI} , respectively, and store in into α' , β' , and γ'
6. Extract two LSB bits from I^R_{SCI} and store it into α'' after converting to its equivalent decimal code.
7. Now using the rule $(10 - \beta')$ and $(10 - \gamma')$, obtain two values $(10 - \beta'')$ and $(10 - \gamma'')$.
8. Extract two bits AC into the middle portion of the interpolated pixel blocks and recovered authentication code (AC')

9. Generate secret image (SI') by concatenating α'' , β'' , and γ'' .
10. Generate authentication key (AC'') by applying SHA-512 on SI' .
11. Image authentication is verified by comparing AC' and AC'' .
12. Hence, original cover image is also recovered by down scaling the image.

4 Results and Discussion

Numerous investigations were carried out to assess how the proposed embedding capacity parameter would affect the safety of data-hiding schemes. We used pictures from the USC-SIPI image database [19] for the experiment. The proposed techniques are tested in a simulated environment using MATLAB. The SSIM, NCC, and PSNR are the measures used to assess the perceptual excellence of the pictures. The BER is used to evaluate the stability of an image. Raising the PSNR and SSIM also improves the perceived quality. To show that the proposed method works, we used four color test photos with a dimension of 512 pixels by 512 pixels. Figure 2 displays the test photos along with their corresponding stego-images.

Mean Square Error (MSE)

The variance among two ($M \times N$) pictures is simply unrushed by the mean square error (MSE). The average MSE for the eight images is 0.059 which is shown in Table 1.

Peak Signal-to-Noise Ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) evaluates how well a picture has been rebuilt.

The PSNR for the Lena, Mandril, Goldhill, Peppers, Barbara, Boat, F16, and Tiffany is 43.51, 42.51, 40.93, 41.27, 43.25, 40.51, 42.12, and 43.5 dB. Average PSNR for the eight images is 42.20 which is shown in Table 1.

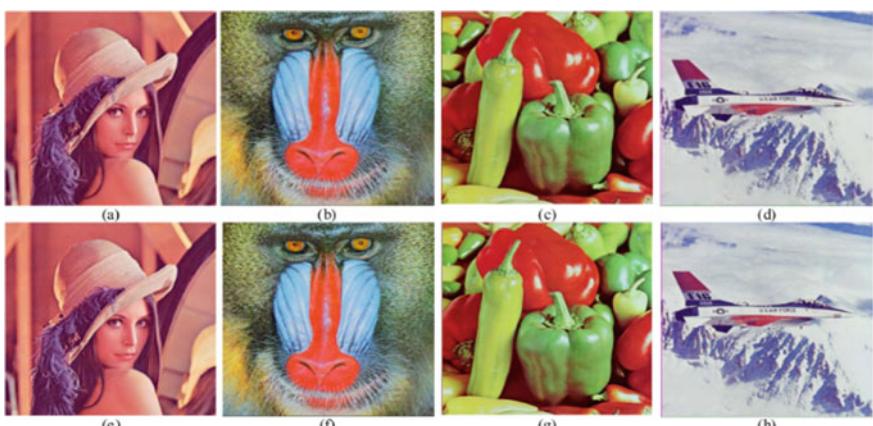


Fig. 2 Test images with size 512×512

Table 1 Experimental result for various metrics

	Capacity	MSE	PSNR	BER	SSIM	NCC
Lena	6,291,456	0.062	43.53	0.0132	0.9421	0.9321
Mandrill	6,291,456	0.071	42.52	0.0213	0.9322	0.8897
Goldhill	6,291,456	0.051	40.93	0.0212	0.9021	0.9321
Peppers	6,291,456	0.042	41.27	0.0312	0.9204	0.9125
Barbara	6,291,456	0.081	43.25	0.0214	0.9027	0.8957
Boat	6,291,456	0.045	40.51	0.0125	0.8954	0.9021
F16	6,291,456	0.083	42.12	0.0214	0.8971	0.9124
Tiffany	6,291,456	0.042	43.5	0.0254	0.9013	0.9254
Average	6,291,456	0.059	42.2	0.0201	0.9117	0.9128

Implanting Capacity (IC)

The amount of secret information that can be inserted into the cover image without causing visual distortions is referred to as the embedding capability, where MEB denotes the maximum number of bits which is entrenched in the cover image. L denotes the total number of bits in cover image.

Structural Similarity Index Measurement (SSIM)

The Structural Similarity Index (SSIM) is usually considered a milestone in the current history of Image Quality Assessment (IQA).

Normalized Correlation Coefficients (NCCs)

The heftiness of the technique is appraised using the NCC measure. It calculates the connection factors between the regained secret (W') and the initial one (W).

The NCC value is given on the interval $[0, 1]$ with unity as the ideal value.

Bit Error Rate (BER)

The bit error rate (BER) or perhaps more exactly as the bit error ratio is obtained by taking the total number of bits transferred and dividing that by the number of bits that were expected with an error. We can determine the BER by calculating the probability that a bit would be incorrectly received due to noise.

The average capacity, MSE, PSNR, BER, SSIM, and NCC are 6,291,456 bits, 0.059, 42.20 dB, 0.0201, 0.9117, and 0.9128, respectively, according to Table 1. According to the average PSNR result, the suggested method offers improved imperceptibility. Table 2 makes it evident that our method outperforms similar schemes in terms of results. Our proposed scheme provides 56.07, 49.42, 15.98, 52.39% better results with respect to PSNR and capacity.

It is shown in Table 2 that our scheme outperforms other analogous schemes in terms of results. Our proposed scheme provides 56.07, 49.42, 15.98, 52.39% better results with respect to PSNR and capacity.

Table 2 Comparison of the proposed scheme to the existing methods

Image	Metric	[13]	[11]	[12]	[18]	Proposed
Lena	PSNR	21.2	21.9	29.32	21.53	43.51
	Capacity	640,937	428,240	382,841	665,849	6,291,456
Airplane	PSNR	27.38	28.6	36.19	28.11	42.12
	Capacity	342,530	177,830	184,280	347,062	6,291,456
Boat	PSNR	27.41	28.73	36.88	28.13	40.51
	Capacity	384,669	216,258	219,442	398,200	6,291,456
House	PSNR	26.9	27.06	34.68	26.67	39.53
	Capacity	411,072	244,607	225,602	425,779	6,291,456
Peppers	PSNR	28.66	29.91	38.05	29.56	41.27
	Capacity	388,981	197,605	213,837	398,262	6,291,456
Goldhill	PSNR	28.35	29.44	37.13	28.9	40.93
	Capacity	443,245	251,003	240,232	462,914	6,291,456
Man	PSNR	25.66	26.94	34.91	26.12	38.87
	Capacity	488,169	285,449	261,444	506,671	6,291,456
Bridge	PSNR	23.75	25.13	33.41	24.46	38.64
	Capacity	516,058	332,834	537,345	537,345	6,291,456
Average	PSNR	26.06	27.22	35.07	26.69	40.67
	Capacity	451,956	266,728	249,352	467,760	6,291,456

5 Conclusion

This research proposes a fresh approach for reversible data concealment in digital photographs, which leverages the MNMI technique and incorporates the CCM. The HVS, or Human Visual System, has been seen to exhibit difficulty in distinguishing between a cover image and concealed images that include random data. This observation has been supported by both application and experimental studies, which have been conducted on widely recognized images such as Lena, Mandril, Goldhill, Peppers, Barbara, the Boat, F16, and Tiffany. The experimental findings demonstrate that the proposed methodology exhibits superior performance in terms of Peak Signal-to-Noise Ratio (PSNR) when compared to the current state-of-the-art techniques. Additionally, it possesses a greater capability for embedding confidential information in comparison to other approaches. In addition to enhancing the data embedding capability, it has also resulted in an augmentation of the Peak Signal-to-Noise Ratio (PSNR) values for widely recognized images. Our study demonstrates that the method we proposed offers a superior approach for concealing data, since it possesses high hidden data capacities and maintains generally steady visual quality of the covered images. It can be confidently stated that the Peak Signal-to-Noise Ratio (PSNR) will exceed 42.20 dB. The utilization of many applications of the MNMI

has the potential to enhance the data-hiding capacity. Consequently, this technique offers a convenient and efficient means of achieving the desired outcome.

References

1. AlKhodaidi T, Gutub A (2020) Trustworthy target key alteration helping counting-based secret sharing applicability. *Arab J Sci Eng* 45:3403–3423
2. Alotaibi M, Al-hendi D, Alroithy B, AlGhamdi M, Gutub A (2019) Secure mobile computing authentication utilizing hash, cryptography and steganography combination. *J Inf Secur Cybercrimes Res (JISCR)* 2(1):9–20
3. Gutub A, Al-Haidari F, Al-Kahsah K, Hamodi J (2010) E-text watermarking: Utilizing “Kashida” extensions in Arabic language electronic writing. *J Emerg Technol Web Intell (JETWI)* 2(1):48–55
4. Almazrooie M, Samsudin A, Gutub A, Salleh MS, Omar MA, Hassan SA (2020) Integrity verification for digital Holy Quran verses using cryptographic hash function and compression. *J King Saud Univ Comput Inf Sci* 32(1):24–34
5. Alassaf N, Gutub A (2019) Simulating light-weight-cryptography implementation for IoT Healthcare data security applications. *Int J E-Health Med Commun (IJEHMC)* 10(4):1–15
6. Lin YK (2012) High capacity reversible data hiding scheme based upon discrete cosine transformation. *J Syst Softw* 85:2395–2404
7. Agrawal S, Kumar M (2016) Reversible data hiding for medical images using integer-to-integer wavelet transform. *IEEE Stud Conf Electr Electron Comput Sci (SCEECS)* 18–22
8. Wu H, Li X, Zhao Y, Ni R (2019) Improved reversible data hiding based on PVO and adaptive pairwise embedding. *J Real-Time Image Process* 16:685–695
9. Tian J (2003) Reversible data embedding using a difference expansion. *IEEE Trans Circ Syst Video Technol* 13:890–896
10. Ni Z, Shi YQ, Ansari N, Su W (2006) Reversible data hiding. *IEEE Trans Circ Syst Video Technol* 16:354–361
11. Jung KH, Yoo KY (2009) Data hiding method using image interpolation. *Comput Stand Interfaces* 31:465–470
12. Chang YT, Huang CT, Lee CF, Wang SJ (2013) Image interpolating based data hiding in conjunction with pixel-shifting of histogram. *J Supercomput* 66:1093–1110
13. Lee CF, Huang YL (2012) An efficient image interpolation increasing payload in reversible data hiding. *Expert Syst Appl* 39:6712–6719
14. Malik A, Sikka G, Verma HK (2017) Image interpolation based high capacity reversible data hiding scheme. *Multimed Tools Appl* 76:24107–24123
15. Malik A, Sikka G, Verma HK (2020) A reversible data hiding scheme for interpolated images based on pixel intensity range. In press, *Multimedia Tools Applications*
16. Zhang X, Sun Z, Tang Z, Yu C, Wang X (2017) High capacity data hiding based on interpolated image. *Multimed Tools Appl* 76:9195–9218
17. Shaik A, Thanikaiselvan V (2019) High capacity reversible data hiding using 2D parabolic interpolation. *Multimed Tools Appl* 78:9717–9735
18. Yalman Y, Akar F, Erturk I (2010). An image interpolation based reversible data hiding method using R-weighted coding. In: 2010 13th IEEE international conference on computational science and engineering, IEEE, pp 346–350
19. Hou D, Wang H, Zhang W, Yu N (2018) Reversible data hiding in JPEG image based on DCT frequency and block selection. *Sign Process* 148:41–47
20. Malik A, Sikka G, Verma HK (2018) An AMBTC compression based data hiding scheme using pixel value adjusting strategy. *Multidimens Syst Sign Process* 29:1801–1818

Comprehensive Approach for Image Noise Analysis: Detection, Classification, Estimation, and Denoising



Rusul A. Al Mudhafar and Nidhal K. El Abbadi

Abstract Image noise is undesirable that can negatively affect the quality of digital images. It reduces the image quality and increases the processing failure ratio. It is highly recommended to remove the noise, and before removing the noise, we have to know the type of noise and estimate the parameters of noise for developing effective noise reduction techniques. This study introduces a method to effectively detect, recognize, and estimate image noise of various types (Gaussian, lognormal, Rayleigh, salt and pepper, and speckle). The proposed model consists of four stages: the first stage is detecting the noise in an image using a convolutional neural network. The second stage classifies the noisy images into one of five types of noise using a new method based on a combination of deep wavelets and support vector machines (SVM) classifier. The third stage involves estimating the parameters of the noise using maximum likelihood estimation (MLE). Finally, choosing the most suitable noise reduction technique for each type using linear and nonlinear filters and showing the capability of the suggested technique in estimating multiple noises commonly present in digital images. The proposed method utilizes a likelihood function derived from the MLE model for each noise type to estimate the noise parameters. Then used to select the most suitable noise reduction technique for each type. The quality of the denoised images is calculated utilizing the peak signal-to-noise ratio (PSNR) as the evaluation metric. The results show that the combination of wavelets with machine learning, specifically SVM, can highly enhance the results, where the accuracy was 93.043% through many experiments conducted to build a sturdy classification model. The MLE-based noise estimation method is also a reliable and accurate method for image noise estimation, especially for Gaussian, salt and pepper, lognormal, and Rayleigh noise. However, for highly noisy types such as speckle noise, further research is required to improve the estimation accuracy. This study contributes to the

R. A. Al Mudhafar (✉)

Computer Science Department, Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq

e-mail: Rusula.mudhafar@student.uokufa.edu.iq

N. K. E. Abbadi

Computer Techniques Engineering Department, Al-Mustaqbal University, Babylon, Iraq

e-mail: nidhal.abass@fulbrightmail.org

development of more effective noise estimation methods for improving the quality of digital images.

Keywords CNN · Deep wavelet · SVM · Maximum likelihood estimation · PSNR

1 Introduction

Image noise can occur during image acquisition, transmission, or processing and it is a frequent issue in image processing and computer vision that can have a considerable impact on the accuracy and quality of image analysis outcomes. Image noise can make it challenging to precisely identify and analyze image features, including edges and textures, and can also introduce errors in image classification and recognition tasks. Therefore, studying image noise and devising methods to mitigate or eliminate it is crucial to enhance the quality and dependability of image analysis results. Therefore, image noise estimation is an essential step in image processing, which aims to estimate the noise parameters accurately and remove or reduce the noise from the images while preserving the important image features [1]. Various types of image noise can occur in digital images, including Gaussian noise, Rayleigh noise, salt and pepper noise, and lognormal noise. Each type of noise has a different statistical distribution and physical origin, which requires different methods to estimate and remove it [2].

Accurate noise calculation is important in various image-processing methods such as denoising, image restoration, compression, and image analysis. Noise estimation helps to determine the appropriate filter parameters and methods that can effectively reduce or remove noise from the images without affecting the image quality or information content. Moreover, noise estimation can also provide useful information about the image acquisition system and help improve the imaging system's performance [3, 4].

Noise detection and classification can be challenging, as several potential issues can arise. Some common problems include a lack of labeled data, variability in noise sources, the complexity of noise signals, and adaptability to new noise types [5].

The noise estimation process can be challenging due to several limitations. First, it assumes that the noise is additive, stationary, and Gaussian, which may not always be the case in real-world scenarios [6]. Second, the accuracy of the estimated noise level is affected by the presence of image structures and textures that can be misinterpreted as noise [7]. Third, the estimation accuracy can be affected by the SNR of the image, and low-SNR images can lead to inaccurate noise estimation [8]. Finally, the estimation process may introduce additional artifacts in the image if not performed carefully [9].

In this article, we focus on analyzing various types of noise in digital pictures, such as, salt and pepper noise, Rayleigh noise, lognormal noise, and speckle noise. To classify the types of noise, we propose a method that utilizes deep wavelet and support vector machine (SVM). Estimating the noise parameters through maximum

likelihood estimation (MLE) is crucial for effectively removing or reducing noise from images while preserving important features.

The remainder of this essay is structured as follows: Part four provides a thorough description of the suggested approach, while part three introduces the research instruments and part two concentrates on related publications. Section five presents the achieved results and discusses them. Finally, the concluding section summarizes the paper's method and the results obtained.

2 Related Works

Many researchers work on image noise analysis including:

A CNN model was presented by Chuah to identify Gaussian noise in pictures and determine its intensity. The study utilized 12,000 and 3000 test images for training and testing, with varying levels of noise added to these images. The accuracy achieved was 74.7%. However, the research solely focused on Gaussian noise and may not apply to all types of noise encountered in real-world scenarios [10].

Ponomarenko and Gapon proposed a method to estimate the variance of white Gaussian noise in highly textured pictures by analyzing local statistics of image gradients. The method decomposes the image into patches and computes gradient statistics for each patch, fitting a local polynomial function to estimate the noise variance. The weightings are based on the assumption that gradients in textured regions are more informative. The suggested technique is compared with other state-of-the-art methods and shown to provide more accurate estimates, especially for highly textured images [11].

Fang and Yi suggested a method to estimate noise levels in natural images based on the analysis of flat patches and local statistics. The method divides the image into small patches and uses the distribution of pixel intensity values in each patch to estimate the noise level. Local statistics such as mean and variance are used to refine the estimate. The recommended method outperforms existing state-of-the-art methods in terms of accuracy and computational efficiency. However, the choice of patch size and local statistics may have an impact on the method's performance [12].

Jiang and Wang suggested a technique for evaluating the noise level in digital images based on the PCA of image texture. The algorithm decomposes an image into sub-bands using wavelet decomposition and applies PCA to each sub-band to extract the principal components. The noise level is estimated from the variance of the lowest principal component in each sub-band. The proposed method is efficient and accurate, even in images with non-uniform textures. However, it may not be suitable for certain types of images such as medical or scientific images with different noise characteristics [13].

Liu et al. introduced a new deep neural network-based method for classifying and denoising four types of image noise: Gaussian, Poisson, salt and pepper, and speckle. The method involves feature extraction, activation function, and network training.

The CNN model achieves an accuracy of 93.7% for noise classification with higher PSNR and SSIM. However, requires more time to train [14].

Hiremath suggested a method for classifying noise in images using transfer learning with a pretrained neural network. It considers two types of noise, electronic noise, and impulse noise, and evaluates the method using a dataset of 400 images, with 200 images containing each type of noise. The results show an average accuracy of 94.37%, with a higher accuracy rate for salt and pepper noise than for Gaussian noise. However, the paper does not explicitly mention how many types of noise were classified, and it only refers to “noise” in general [15].

Chin and Chow proposed a new method for estimating image distortion based on image gradients. The proposed method uses the properties of gradient magnitudes and directions to estimate the amount of distortion present in the image. The method is based on the assumption that distortion in images leads to changes in gradient magnitude and direction, and these changes can be used to estimate the amount of distortion. Distortion estimation gradient can estimate additive noise, impulsive noise, and multiplicative noise from single-source images accurately. The advantage of this method is it can accurately estimate multiple types of noise parameters from single-source images, making it a versatile tool for various image-processing tasks [16].

Objois and Okumuş suggested a technique for calculating noise parameters in images affected by Poisson-Gaussian noise by taking into account specific characteristics of the target in the image. The method uses a maximum likelihood estimation approach and outperforms existing techniques in terms of accuracy and computational complexity. The method has potential applications in image denoising, restoration, and enhancement. The advantages of the method include the consideration of different target types and the resulting accurate and robust estimation. However, a disadvantage is a need to specify target locations in the image, which can be time-consuming [17].

Our paper includes several contributions, such as the detection, classification, and estimation of noise parameters for five types of noise, as well as the suggestion of appropriate noise reduction techniques for each type. We also evaluated the effectiveness of these techniques using PSNR.

3 Research Tools

Many tools (algorithms and techniques) are used in this proposal, and in this section, a brief explanation will be presented.

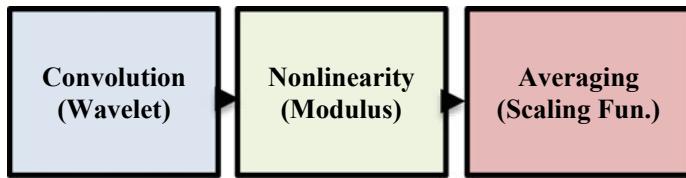


Fig. 1 Operations of wavelet scatter

3.1 Convolution Neural Network

ConvNet is a type of ANN utilized for visual image analysis. It has a unique shared-weight architecture of convolution filters that produce activation maps and comprises an input, hidden, and output layer. The hidden layers convolute input/output using activation functions, such as ReLU, and include convolutional, pooling, normalization, flattened, and fully connected layers. The output layer contains the final classification matrix [18].

3.2 Deep Wavelet Transforms

When it comes to obtaining accurate data representations and feature extractions that work with the majority of classification algorithms, wavelet techniques are useful tools. The wavelet transform allows the creation of reliable features that are locally stable to small deformations when used in combination with a deep neural network. A deep wavelet consists of many layers, where one layer's output serves as the next layer's input. Each layer consists of three operations, as shown in Fig. 1. [19].

3.3 SVM Classifier

An SVM is a supervised learning model with learning algorithms that look at data to classify it and predict what will happen next. SVMs are one of the best ways to make predictions because they use statistical learning frameworks or the theory of vector clustering (VC). The SVM training algorithm makes a model that gives new instances of a flag to tell one class from another [20].

3.4 Maximum Likelihood Estimation (MLE)

It is a method for evaluating the parameters of a statistical model by maximizing the likelihood function with regard to the unknown parameters. Given the parameter values, the likelihood function provides the probability of the observed data. The MLE approach finds parameter values that maximize the likelihood function [21]. MLE is extensively utilized in a broad range of domains, including statistics, machine learning, and computer vision. It has a solid theoretical foundation and is considered a consistent and efficient method for estimating model parameters [22].

One of the key advantages of MLE is that it provides unbiased estimates of the parameters, meaning that on average, the estimated parameter values are equal to the true parameter values. MLE has desirable asymptotic properties, meaning that as the sample size increases, the MLE estimates approach the true parameter values [23].

3.5 Denoising Techniques

Denoising methods can be either linear or nonlinear. Linear methods are generally faster but do not preserve image details, while nonlinear methods are better at preserving them. Denoising filters in the spatial domain can be broadly categorized into the following categories:

3.5.1 Median Filter

A nonlinear filter that substitutes the median value of each nearby pixel for the value of each individual pixel. It is often used for salt and pepper noise reduction [24].

3.5.2 Wiener Filter

A linear filter that minimizes the mean square error between the denoised image and the original image while preserving the important image features. It is often used for Gaussian reduction [25].

3.5.3 Block Matching and 3D (BM3D)

This algorithm works by first dividing the noisy image into overlapping blocks and then matching similar blocks to form groups. The denoising process is then performed on each group of similar blocks using a collaborative filtering technique that estimates the noise-free signal by averaging the overlapping blocks within each group [26].

3.5.4 Non-local Means Filter

Estimates the denoised value of each pixel based on the weighted average of its neighboring pixels, where the weights are based on the similarity between the image patches centered at the pixels. It is often used for various types of noise reduction [27].

3.5.5 Wavelet-Based Filter

Wavelet-based methods for denoising are a popular family of techniques that use the wavelet transform to decompose an image into different frequency bands. The idea behind this approach is that image noise often has a different frequency distribution than the image itself. By separating the image into different frequency bands, it is possible to apply denoising methods that are better-suited to the characteristics of the noise in each band [28].

4 Methodology

This section explains in detail the proposed model, which consists of four stages: (i) preprocessing, (ii) noise detection using CNN, (iii) noise classification using a combination of deep wavelet and SVM classifier, (iv) estimating the noise parameters using MLE, and (v) selection of the optimal denoising techniques using linear and nonlinear filters. Figure 2 shows the steps of a proposed model.

4.1 Preprocessing

In the preprocessing step, all training images are converted to grayscale, resized to 150×150 , and subjected to augmentation.

4.2 Noise Detection

This model aims to distinguish the noisy image from the clear image based on convolutional neural network. The CNN consists of five convolutional layers, each with 16, 32, 64, 128, and 256 filters and a kernel size of (3×3) , as shown in Fig. 3. ReLU is used as the activation function, and there are five max-pooling layers following each convolutional layer. The feature maps are then fed into a flattened layer to convert them from 2 to 1D, followed by two fully connected (FC) layers. The first FC layer has 512 channels, and the second FC layer has one channel. Sigmoid

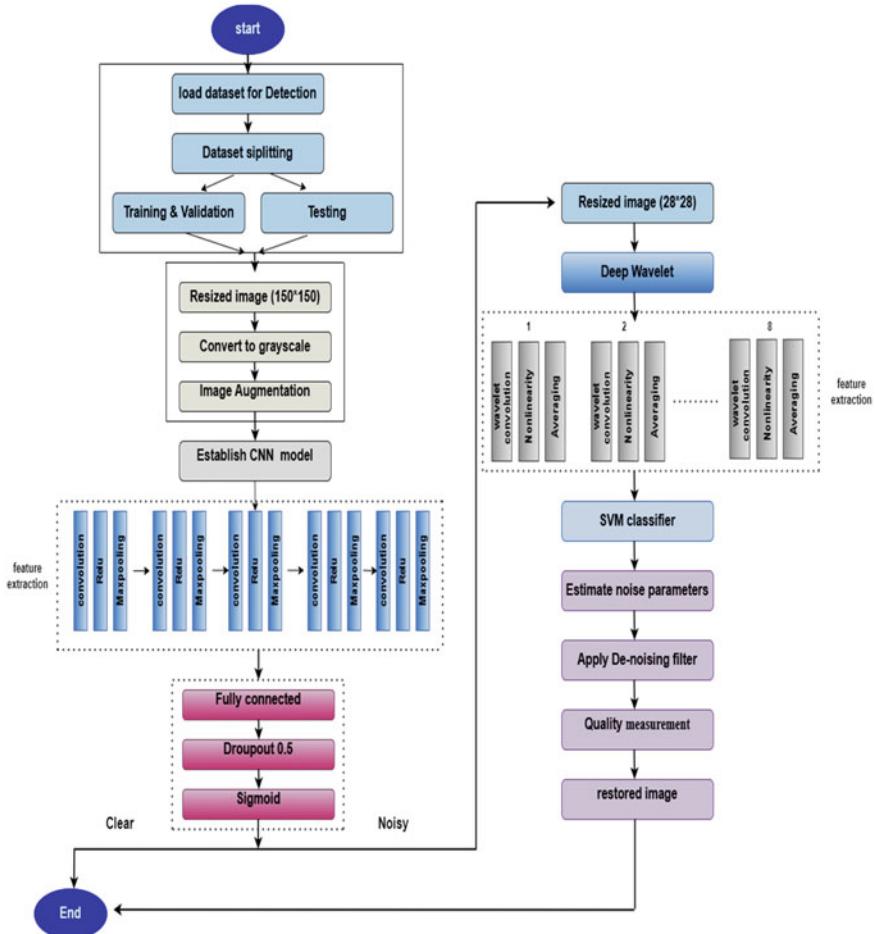


Fig. 2 Flowchart of the proposed method

is used as the activation function for the fully connected layers. The model is trained for 20 epochs.

4.3 Noise Classification

Once the noise detection is completed, the next step is to find the noise type, this is achieved by proposing a new method based on a combination of deep wavelet and machine learning classifiers SVM. The proposed method is capable of classifying five types of noise, including Gaussian, lognormal, Rayleigh, salt and pepper, and speckle. Deep wavelet refers to the use of deep learning techniques to analyze signals

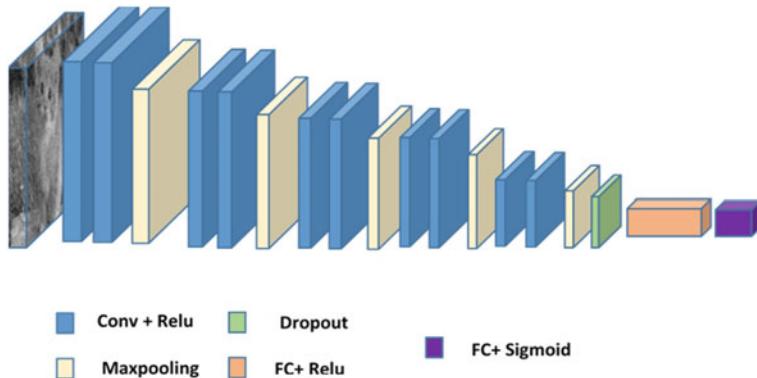


Fig. 3 Architecture of proposed CNN

using wavelets. Wavelets are mathematical functions that can be used to analyze signals in both the time and frequency domains. Deep learning algorithms can be used to extract features from signals using wavelets, which can then be used for classification.

SVM is a well-known ML technique for classification problems. SVM works by determining the optimal boundary (hyperplane) between data points from distinct classes. SVM is particularly useful for classification tasks where there are many features or dimensions.

The combination of deep wavelet and SVM for noise classification involves using deep wavelet to extract features from the signal and then used as input to the SVM algorithm, which learns to classify the noise type based on the extracted features.

4.4 Noise Estimation

The noise parameters (amount of noise) are estimated using a likelihood function derived from the MLE model for each noise type.

MLE is a statistical modeling technique that uses observable data to estimate the parameters of a probability distribution. Essentially, MLE seeks the parameter values that increase the chance of witnessing the provided data. In the context of estimating noise parameters, the likelihood function derived from MLE is used to estimate the amount of noise for a given image. This function takes into account the specific noise type and its characteristics, as well as the observed data. By fitting this function to the observed data, the maximum likelihood estimator can be utilized to estimate the parameters of the noise distribution that best fit the data.

Table 1 Noise reduction technique for each type of noise

Noise type	Denoising method
Gaussian	Wiener filter
Salt and pepper	Median filter
Speckle	Wavelet-based filter
Lognormal	BM3D
Rayleigh	Non-local means filter

4.5 Noise Reduction

The choice of the noise reduction technique depends on the type of noise present in the image and the characteristics of the image itself. Table 1 gives an appropriate noise reduction technique for each type of noise.

5 Result Analysis and Discussion

The dataset is the essential factor for any algorithm, model, or system. For noise detection, the dataset included 22,023 noisy and clear images of different sizes. These images are divided into training and testing images, the number of images for training and testing. For noise classification, we utilize the “9_classes_noisy_image_dataset” available on Kaggle, consisting of 1150 images with a size of 600×464 . The dataset is divided into two subsets for training and testing purposes.

5.1 Performance of the Proposed CNN

The first test is to measure the performance of the noise detection method. A noisy image is detected by using the CNN network before recognition of the noise type. The whole detection model’s confusion matrix is shown in Fig. 4, and the classification report for the performance outcomes, which include recall, precision, F1 score, and support, is given in Table 2 which is better than misclassifying noisy images as clear. However, this may be due to blur or low-quality images.

The detecting model’s overall effectiveness was 99.63% in training and 98.86% in testing. In this test, error detection may happen due to blur images or low-quality images.

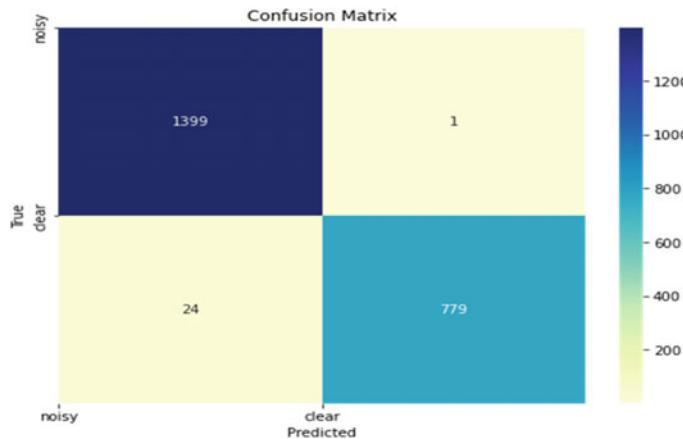


Fig. 4 Confusion matrix of a detection model

Table 2 Classification report of the proposed method for noise detection

Name of classes	Recall (%)	Precision (%)	F1 score (%)	Support
Noisy image	99.92	98.31	99.11	1400
Clear image	97.01	99.87	98.42	803
Accuracy	98.86			2203
Marco average	98.46	99.09	98.76	2203
Weighted average	98.86	98.88	98.86	2203

5.2 Performance of the Classification Model

The performance of the classification of the noise type is one of the necessary tests. The classification proposed method was tested by using 115 images with five different noise types. The results are shown in the confusion matrix in Fig. 5. Table 3 lists the classification report accuracy for every noise. The performance measurements are (recall, precision, F1 score, and support).

The overall accuracy for the proposed classification model for five noise types was 100% in training and 93.043% in testing with new images not included in the training. The significant similarity between the selected noise types, a challenge for most researchers, reduces the classification accuracy.

5.3 Performance of the Estimation Method

The method aimed to estimate the effect of different levels of noise on various image types using different noise parameters. The method examined the impact of speckle

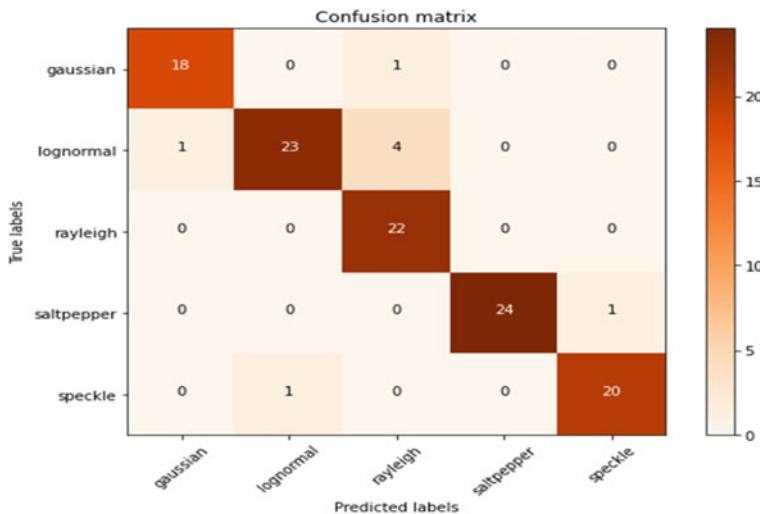


Fig. 5 Confusion matrix of a classification model

Table 3 Classification report of the proposed method for five noise types

Noise type	Precision (%)	Recall (%)	F1 score (%)	Support
Gaussian	94.73	94.73	94.73	19
Lognormal	95.83	82.14	88.46	28
Rayleigh	81.48	100	89.79	22
Salt and pepper	100	96	97.95	25
Speckle	95.23	95.23	95.23	21
Accuracy	93.043			115
Marco avg.	93.45	93.62	93.23	115
Weighted avg.	93.70	93.04	93.05	115

noise, salt and pepper noise, Gaussian noise, lognormal noise, Rayleigh noise, and combined noise on the images with varying standard deviations and noise amounts. Table 4 displays the study's findings.

This table demonstrates that the maximum likelihood estimation (MLE) algorithm is very good at estimating certain types of noise, such as Gaussian, salt and pepper, lognormal, and Rayleigh. However, the effect of speckle noise is relatively small compared with other types. Furthermore, the estimation of salt and pepper noise combined with other types of noise is more accurate with Gaussian noise than other noises.

Table 4 Parameter estimation of different types of noise

Type of noise	Truth parameter	Estimated parameter
Gaussian noise	5	4.99
	10	10.09
Salt and pepper noise	0.01	0.0108
	0.02	0.0207
Speckle noise	0.01	0.03
	0.02	0.0337
Lognormal noise	0.1	0.1
	0.2	0.201
Rayleigh noise	0.1	0.107
	0.2	0.215
Combined noise	Salt = 0.50, pep = 0.50, sigma = 3	Salt = 0.50, pep = 0.50, sigma = 3.06
	Salt = 0.40, pep = 0.60, sigma = 5	Salt = 0.40, pep = 0.60, sigma = 4.99
	Salt = 0.30, pep = 0.70, sigma = 10	Salt = 0.30, pep = 0.70, sigma = 7.78

5.4 Performance of the Denoising Method

Several metrics were used to evaluate the quality of the denoised image, including the peak signal to noise ratio (PSNR), as given in Table 5.

This table demonstrates the results of noise removal after adding various amounts of noise to the original image.

Table 5 Denoising results with different noise level parameters

Type of noise	Noise level	Method	PSNR
Gaussian noise	5	Wiener filter	36.571
	10		35.472
Salt and pepper noise	0.01	Median filter	38.419
	0.02		37.881
Speckle noise	0.01	Wavelet-based filter	42.552
	0.02		37.436
Lognormal noise	0.1	BM3D	74.090
	0.2		72.964
Rayleigh noise	0.1	Non-local means filter	72.896
	0.2		73.927

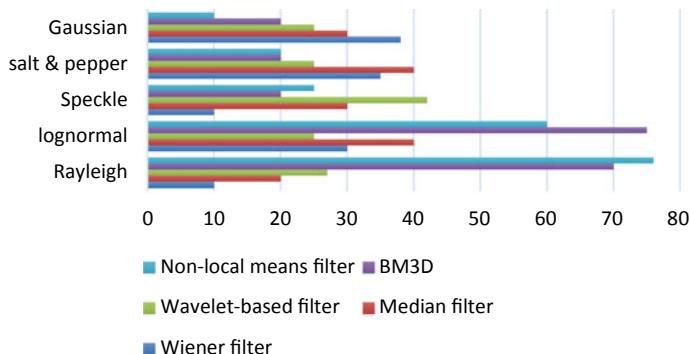


Fig. 6 Optimal algorithm for each type of noise

The choice of a suitable noise reduction technique depends on the type of noise present in the image and the PSNR of the algorithm. Figure 6 shows the optimal algorithm for each type of noise.

6 Conclusion

In this paper, we propose a technique for identify in and classifying five types of image noise, namely Gaussian, lognormal, Rayleigh, salt and pepper, and speckle, using a combination of deep wavelet and machine learning techniques. The proposed approach achieves high accuracy rates of 98 and 91.30% for noise detection and classification, respectively.

The combination of deep wavelet and SVM is a powerful technique for classifying the type of noise in a signal, as it allows for the extraction of relevant features and the use of a powerful classification algorithm.

The likelihood function derived from MLE is a powerful tool for estimating noise parameters, as it can be tailored to fit the specific characteristics of the noise type and can provide accurate estimates even in the presence of complex noise patterns.

References

1. Babaei-Mahani A, Sheikhzadeh H, Cheriet M (2014) A survey on image denoising using fuzzy logic. *J Electron Imaging* 23(1):1–26
2. Asano S, Kamata S, Miyatake T (2019) Survey of noise reduction methods for digital images. *J Signal Process* 23(1):1–10. <https://doi.org/10.2299/jsp.23.1>
3. Ismail M, Usman M, Daud A (2018) A comparative study of noise estimation methods for digital images. *EURASIP J Image Video Process* 2018(1):16. <https://doi.org/10.1186/s13640-018-0246-9>

4. Kirmizitas K, Besli N (2022) Image and texture independent deep learning noise estimation using multiple frames Elektron. ir Elektrotechnika 28:42–47
5. Patil S, Sherekar SS (2019) Noise detection and classification using machine learning: a review. J Ambient Intell Humaniz Comput 10(1):169–183. <https://doi.org/10.1007/s12652-018-0896-2>
6. Ponomaryov V, Egiazarian K (2012) Non-stationary noise estimation and filtering in the image and video processing. EURASIP J Adv Sign Process 2012(1):1–2
7. Egiazarian K, Foi A, Katkovnik V, Oksanen L (2017) Noise variance estimation in additive Gaussian noise: a review of methods and challenges. Digital Signal Process 70:1–19
8. Li C, Guo S, Porikli F (2013) No-reference quality assessment for natural videos via blind quality estimator in a spatiotemporal domain. IEEE Trans Circ Syst Video Technol 23(3):505–515. <https://doi.org/10.1109/TCSVT.2012.2231520>
9. Liu J, Wu W, Zuo W, Zhang D (2020) An efficient noise level estimation method for real images. IEEE Trans Image Process 29:559–570
10. Chuah JH, Khaw HY, Soon FC, Chow C (2017) Detection of gaussian noise and its level using deep convolutional neural network. In: Proceeding of the 2017 IEEE region 10 conference (TENCON), Malaysia, November 5–8, pp 2447–2450
11. Ponomarenko M, Gapon N, Voronin V, Egiazarian K (2018) Blind estimation of white Gaussian noise variance in highly textured images. IS T Int Symp Electron Imag Sci Technol. <https://doi.org/10.2352/ISSN.2470-1173.2018.13.IPAS-382>
12. Fang Z, Yi X (2019) A novel natural image noise level estimation based on flat patches and local statistics. Multimed Tools Appl 78(13):17337–17358. <https://doi.org/10.1007/s11042-018-7137-4>
13. Jiang P, Wang Q, Wu J (2020) Efficient noise-level estimation based on principal image texture. IEEE Trans Circuits Syst Video Technol 30(7):1987–1999. <https://doi.org/10.1109/TCSVT.2019.2912319>
14. Liu F, Song Q, Jin G (2020) The classification and denoising of image noise based on deep neural networks. IEEE Access 7:2194–2207. <https://doi.org/10.1107/s10489-019-01623-0>
15. Hiremath PS (2021) Identification of noise in an image using artificial neural network. Int J Eng Res Technol 10(2):345–348
16. Chin SC, Chow CO, Kanesan J, Chuah JH (2022) A study on distortion estimation based on image gradients. Sensors 22(2). <https://doi.org/10.3390/s22020639>
17. Objois É, Okumuş K, Bähler N (2022) Target aware poisson-gaussian noise parameters estimation from noisy images, pp 1–10. <http://arxiv.org/abs/2210.12142>
18. Sarigul M, Ozyildirim BM, Avci M (2019) Differential convolutional neural network. Neural Netw 116:279–287. <https://doi.org/10.1016/j.neunet.2019.04.013>
19. Soro B, Lee C (2019) A wavelet scattering feature extraction approach for deep neural network based indoor fingerprinting localization. Sensors (Switzerland) 19(80). <https://doi.org/10.3390/s19081790>
20. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Transact Intell Syst Technol 27(3):1–27. <https://doi.org/10.1145/1961189.1961199>
21. Bharadwaj Prakash KB, Kanagachidambaresan GR (2021) Pattern recognition and machine learning. EAI/Springer Innovations in Communication and Computing. https://doi.org/10.1007/978-3-030-57077-4_11
22. McLachlan GJ, Lee SX, Rathnayake SI (2019) Finite mixture models. Ann Rev Stat Appl 6:355–378
23. Myung IJ (2003) Tutorial on maximum likelihood estimation. J Math Psychol 47(1):90–100
24. Umbaugh SE (2018) Image processing and analysis. In: Digital image processing and analysis: applications with MATLAB and CVIPtools, 3rd ed., Boca Raton, FL, USA, CRC Press, pp 144–152
25. Zhang X (2016) Image denoising using local Wiener filter and its method noise. Optik 127(17):6821–6828
26. Sarjanova S, Boutellier J, Hannuksela H (2015) BM3D image denoising using heterogeneous computing platforms. In: 2015 conference on design and architectures for signal and image processing (DASIP), Krakow, pp 1–8. <https://doi.org/10.1109/DASIP.2015.7367217>

27. Sun Z, Meikle S, Calamante F, Cowin GJ (2022) CONN-NLM: a novel means filter for PET-MRI denoising. IEEE Access 16:1–14. <https://doi.org/10.1109/ACCESS.2022.3162002>
28. Jaiswal A, Upadhyay J, Somkuwar A (2014) "Image denoising and quality measurements by using filtering and wavelet-based techniques. AEU-Int J Electron Commun 68(8):699–705

Optimal Path Selection Algorithm for Energy and Lifetime Maximization in Mobile Ad Hoc Networks Using Deep Learning



Jyoti Srivastava and Jay Prakash

Abstract The energy-efficient path selection algorithm proposed in this paper balances the conflicting goals of maximizing network lifetime and minimizing energy usage routing in mobile ad hoc networks (MANETs). The proposed strategy maximizes lifetime energy efficiency, MANET, and deep learning. Produce the data after building the network by carrying out assaults and validating paths. Then sketch a neural network with capabilities for prediction and performance evaluation. Then nodes in a network that are negative by definition must be followed by choosing the optimum route. Employed in the current study to increase the energy efficiency as well as the kind of data handling on the network with the metrics of stolen time, total time, total energy, and packet delivery rate, predict the energy and lifetime maximization utilizing deep neural networks for deep learning, management, and lifetime energy efficiency maximization. Five hundred packets of data from a neural network were used to get the maximum value. The total energy used is 7570, packets are delivered at 74.60, time taken is 371.81, and the minimum theft rate for 500 packets is 6.8.

Keywords Deep learning · Neural network · Ad hoc network · Lifetime energy efficiency maximization · Optimal path · Secure routing

1 Introduction

To extend the network's lifespan, the sensor node's energy usage is decreased. Lifetime maximization in heterogeneous WSNs is the subject of very few studies. Mobile hosts connected by radio links form MANETs, dynamic networks without fixed infrastructure. Direct communication or router-like intermediate nodes connect these nodes [1]. When building routing protocols, it is necessary to consider several problems and challenges brought on by the network's dynamic nature. MANET's

J. Srivastava (✉) · J. Prakash

Madan Mohan Malviya University of Technology, Gorakhpur, India
e-mail: sriv.jyoti1996@gmail.com

real-time applications demand QoS support, including effective bandwidth, minimal delay, maximum throughput, maximum network lifetime, and routing protocols [2, 3]. The fast-paced nature of modern design challenges has spurred researchers to explore energy-aware QoS-based routing algorithms that can facilitate rapid and dependable communication within mobile ad hoc networks (MANETs). The advancement of mobile technology is bringing about a societal revolution [4]. The concept of pervasive computing, as proposed by Weiser in 1999, has led to a paradigm shift from personal computing to a more ubiquitous form of computing. With pervasive computing, users can access their data anytime and anywhere since mobile nodes can establish connections [5]. Today's 4G wireless evolution has as one of its main goals the creation of a pervasive environment that enables users to access information and share it with others at any time and from any location [6].

An infrastructure-less, self-configuring network that can be set up instantly is a fundamental innovation for a specially appointed organization in providing an unstoppable processing environment. For communication between the nodes, fixed backbone networks are not necessary [7, 8]. The organization allows any hub to join or exit at any time. The specially assigned network's hubs can all communicate with one another without the need for a tunnel. The bundle is advanced from one hub to the next by each hub in the organization, acting as a switch [9]. The impromptu organization's hubs are extremely movable, and communication between them is exceedingly challenging. Single-hop communication refers to direct communication between nodes, whereas multi-hop communication refers to indirect communication through intermediary nodes between nodes outside each other's communication range. The hubs in MANETs are independent of one another and do not rely on any base stations or earlier foundations [10]. Due to the network hubs' freedom of movement in MANETs, the organizational geography might vary fast and unexpectedly. The real hubs must carry out all organizational tasks individually or collectively, such as determining the geography and transferring information bundles [11]. Even in rural and military locations, ad hoc network development is quick due to a lack of infrastructure because no particular gear is needed for connections [12]. An organization of this type might operate alone or be connected to the Internet. Designing guiding conventions is extremely difficult because of factors such as multiple bounces, portability, and the scale of large companies, as well as changes in devices, gearbox capacity, and battery power requirements [13]. The performance of this adaptive network can be better by implementing good routing strategies [14].

Deep learning algorithms may be used in buffer-less systems to acquire input data as quickly as possible from various sources. Further study will also look at a cross-layer strategy to extend the network's lifetime and efficiency in the event of node failures [15]. Deep learning methods and meta-heuristics consider environmental conditions [16]. We also used several experiments to demonstrate the superiority of the suggested framework. Conventional energy-saving methods can change the structure [17].

2 Related Work

Bharti et al. have proposed an improved path routing with a buffer allocation (IPBA) system to enhance communication in mobile ad hoc networks (MANETs). This system is designed to prevent packet loss by utilizing a buffer temporarily storing information about nodes and data packets prepared for transmission and reception. A coupling node selection technique is also implemented to ensure that data packets are transmitted regularly in normal scenarios. This technique links two efficient nodes selected to carry out the communication. As a result, the lifespan of the network is extended while the packet loss rate is reduced. When determining how well the IPBA system performs, factors such as speed, network life, packet loss, end-to-end delay, and communication overhead are considered [12].

Misra et al. have proposed a strategy to improve the operation, extend the lifespan, and maximize the utility of wireless sensor networks (WSNs). To achieve these goals, the EAACOP protocol creates groups using the pillar K-means grouping technique and ensures that every node is covered as much as possible. It is also used to identify the best candidates for group leaders and to assess the energy consumption during TDMA-based MAC protocol data transmission. Simulations and comparisons with other current protocols have shown that the protocol works, and QoS performance standards have been used to conclude. Compared with traditional methods, the EAACOP protocol enhances throughput, residual energy, and packet distribution ratio while reducing packet loss, network latency, end-to-end delay, and battery power utilization. As a result, the network as a whole performs better [18].

Xue et al. the ideal solution to the EE maximization problem is using a two-level iterative technique based on Dinkelbach. The suggested algorithm's effectiveness and convergence are supported by numerical findings [11].

Yang et al. a method for assessing node trust using a cluster structure and a secure data transfer mechanism without a certificate authority (CA) has been proposed. To improve the effectiveness of node reliability evaluation, the suggested technique utilizes a hierarchical structure, with a trust management node maintaining the measured reliability of the nodes. While a key exchange technique between nodes without using a CA improves data transmission integrity, reliability considers the quantity and quality of packets. Checking for nodes that produce excessive traffic on the path when data is transmitted aids in identifying anomaly nodes and enhances routing efficiency [1].

Alkadhmi et al. have created a method for extending the life of nodes in a mobile ad hoc network (MANET) by staying in touch with the nearest base station. By enabling the quick creation of temporary connections with surviving nodes, this method reduces the need for significant infrastructure repairs. The load in a MANET should be spread out fairly among the nodes to extend the life of the nodes and keep those with low-energy reserves from running out of power. This type of even load sharing increases network life [19].

3 Proposed Methodology

The suggested approach for deep learning, MANET, and lifetime energy efficiency maximization is discussed in this section. By conducting assaults and validating paths, first, establish...000.0h1 the network, then produce the data. Then describe a neural network that includes performance evaluation and prediction. Then malicious nodes in a network definition. After this, the best path must be chosen, followed by time, energy, and PDR, as illustrated in Fig. 1.

A. Dataset Preparation

This work discusses preparing the data during the evaluation for lifetime energy efficiency that maximizes security using MANET with malicious node detection and new path generation schema [23].

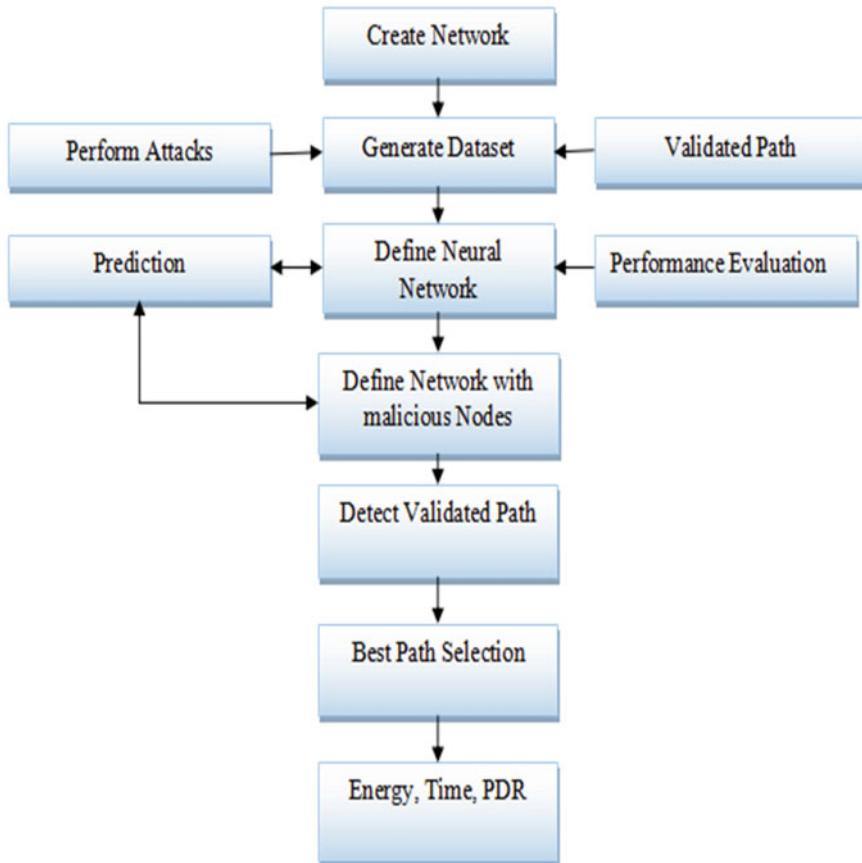


Fig. 1 Proposed flowchart

- *Define Network*—To implement this, use 50 nodes. Create the node first, specify the energy two times, and set the node's x and y positions to 200 and 800, respectively. The maximum meter range is $1000 * 1000$. where velocity is defined as shown in Eq. 1.

$$\text{Velocity(Node)} = \text{Random}(1, h) * 19\text{m/s} \quad (1)$$

While in the direction random source, 1 represents north, 2 represents south, 3 represents east, and 4 represents west [24]. Set the source in the destination node and each node's radio propagation range.

- *Apply Attack*—Applying an attack like a wormhole with two nodes in the entire network as attacker 1 and 2 with velocity 0.001 for both the nodes, then update connections after applying the attack.
- *Route Discovering*—Finding all possible paths between nodes can be done by defining nodes, creating route pools, creating connections, applying conditions (if empty == 1), breaking the connection, finding a random source node with a random destination, and considering that with a route pool, designating nodes 2 and 49 as attacker nodes, and repeating this process for all nodes with rout pools filled with real and defected nodes [25].
- *Feature Extraction*—To extract the path and features by extracting the maximum, minimum, mean, and standard deviation of speed, faster direction, and all possible directions. Use network injection with an attack-type condition [26]. Verify that an attack is contained, then confirm it using a verified path and a malicious node labeled zero. If not, confirm that the path is genuine and defined as one. Labels are equal to label + validation, while data is equal to data + vectors.
- The generated data has eight columns and 3977 rows.

Algorithm 1: Steps for Data Generation and Collection

1. **for** $i = 1$ to $i = 100$: **do**

 Create a network with 50 nodes

 Create attacker nodes and apply an attack

 Update connections with route discovery

 Extract features

for $k = 1$ to $k = \text{length of path}$: **do**

 generate vectors to store paths with information

end

 Packet inject

 reliability time = 5 in sec

 moving packets with network nodes

for $j = 1$ to $j = \text{path length}$: **do**

(continued)

(continued)

Algorithm 1: Steps for Data Generation and Collection

```

if it contains an attack node and an attack
then validation == 0
else path is valid
end
data == data + vectors
labels == labels + validation
2. end

```

B. Implement Deep Neural Network and Training

Use a feed-forward neural network using ten neurons for the input and output layers, with one neuron in each layer. Using a random function to split the data, the Levenberg-Marquardt training process, the MSE and accuracy performance metrics, and the Tan Sigmoid activation function [27].

$$\text{Accuracy} = \text{Sum}(\text{rand}(\text{net}(\text{data}))) == \frac{\text{labels}}{\text{Numel}(\text{labels})} * 100 \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (3)$$

Levenberg-Marquardt algorithm—“The Levenberg-Marquardt algorithm (LMA or LM), also known as the damped least-squares (DLS) approach, is a mathematical and computational tool used to solve non-linear least-squares problems, especially in curve fitting. The LMA algorithm interpolates between the gradient descent method and the Gauss-Newton algorithm (GNA) for solving minimization problems”.

Algorithm 2: Feed-Forward Neural Network to Detect Malicious Data

1. load data and labels
 2. create a neural network with ten neurons with a train
 3. create layer 2 with the activation function transit
 4. fit the model with training data and split data randomly
 5. save the model as net and evaluate the performance
-

C. Securing and Improving Lifetime Using Train Neural Network

Network setup with 100 nodes that are attacking. Apply the attack using the attacker node, load the trained neural network, and take the user’s response as either 1 for the prediction made by the neural network or 2 for no neural network [28]. Construct n-packets, define stolen packets, and count the remaining packets in networks if 500 simulation packets were also used, along with 100, 200, 300, and 400. Finding every

feasible path helps in establishing the source and destination nodes. Finding the optimum path with validation and checking the path using a trained neural network with the last vectors generated based on the path features [29]. Update the nodes' connection. Set and verify that the path contains the best path except for the first and second malicious nodes, then remove and add as a stolen value increases, creating final routing, choosing a path, and locating sources and destinations quickly [30].

4 Result and Discussion

This section discusses the simulation results for energy and lifetime maximization using a deep neural network for deep learning, MANET, and lifetime energy efficiency maximization with the metrics stolen time, total time, total energy, and packet delivery rate. Tables 2 and 3 give the performance evaluation of with/without neural network with packets, respectively.

Figures 2 and 3 show the theft and packet delivery rate plots with/without AI, respectively. Figures 4 and 5 show the time and energy comparison plot without/with AI, respectively.

The greatest value was attained by 500 packets of neural network with packets in Tables 1 and 2, representing the performance evaluation of without and with the neural network with packets. The total energy is 7570, the packet delivery rate is 74.60, the total time is 371.81, and the least amount of stolen goods were obtained using 500 packets, which is 6.8.

Research on SA has provided a number of reliable methodologies, however there are a few intriguing problems that still need to be explored.

Figures 6 and 7 display the energy in nodes and time in nodes where time (sec) and nodes calculate. The deep neural network's training states are shown in Fig. 8. For the evaluation, which includes a gradient, Mu, and validation check every 13 epochs.

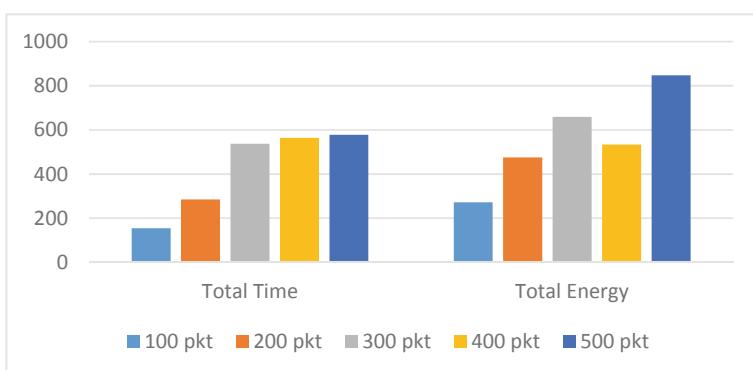


Fig. 2 Time and energy comparison plot without AI

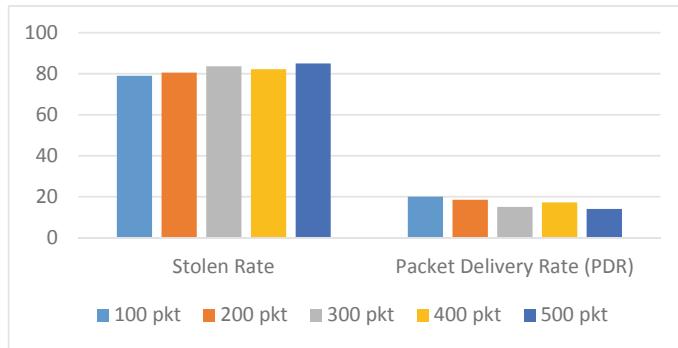


Fig. 3 Theft and packet delivery rate plot without AI

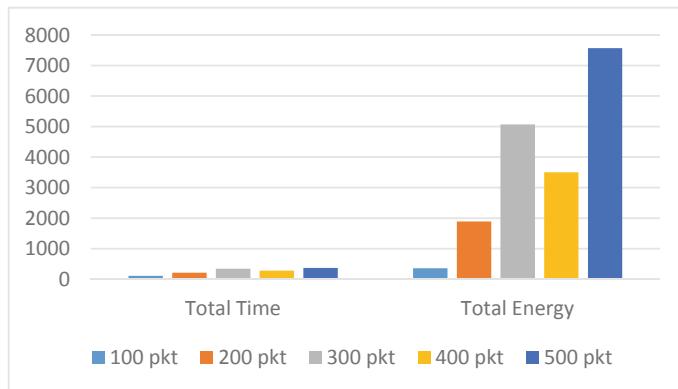


Fig. 4 Time and energy comparison plot without AI

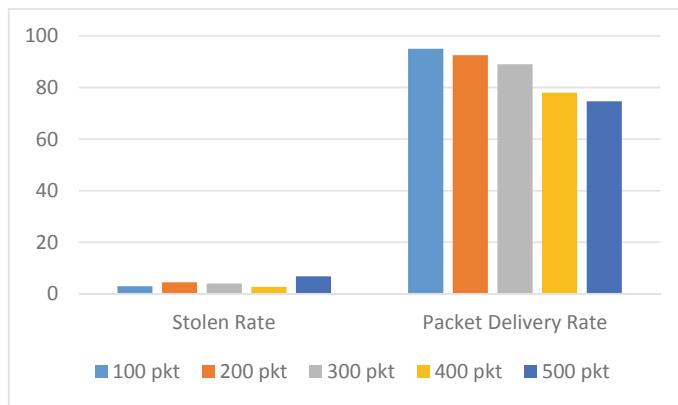


Fig. 5 Theft and packet delivery rate plot with AI

Table 1 Literature summary

Author/ year	Method	Ref. No.
Tomar/ 2022	Suggested protocols focus on one or two QoS factors while choosing a path	[20]
Bangotra/ 2022	In wireless sensor networks (WSNs), sensor nodes' processing capabilities are constrained and worries about data transmission security further restrict the transmission of WSN applications	[5]
Ding/2021	Machine learning (ML) technology has distinct advantages in various industries, as demonstrated by various applications, including picture and speech recognition, recommendation systems, and natural language processing	[13]
Natarajan/ 2019	Analysis of the many routing protocols that improve the performance of the flying wireless networks	[21]
Sharma/ 2015	Current routing algorithms for sensor networks, a taxonomy of the various tactics employed and a comparison of their advantages and disadvantages	[22]

Table 2 Performance evaluation of without neural network with packets

Metrics	100 pkt	200 pkt	300 pkt	400 pkt	500 pkt
Stolen rate	79	80.50	83.66	82.25	85
Total time	154.83	285.08	537.23	564.88	578.01
Total energy	272.09	475.73	659.88	534.34	847.35
Packet delivery rate	20	18.50	15	17.25	14.00

Table 3 Performance evaluation of neural network with packets

Metrics	100 pkt	200 pkt	300 pkt	400 pkt	500 pkt
Stolen rate	3	4.50	4	2.75	6.8
Total time	107.59	211.29	340.72	279.65	371.81
Total energy	354.61	1891	5067	3500	7570
Packet delivery rate	95	92.50	89	78	74.60

Figure 8 displays the error histogram with 20 bins and the best validation over seven epochs, where blue denotes training, green denotes validation, and red denotes testing.

Figure 9 displays the training, validation, test, and all graphs of a deep neural network, where blue, green, red, and black lines indicate fit, holes indicate data, and a dashed line indicates that the model is correct ($y = T$).

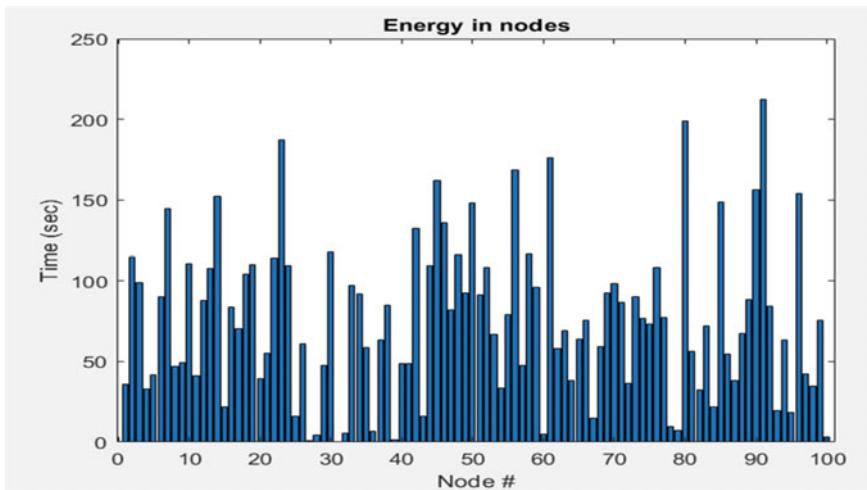


Fig. 6 Energy in nodes

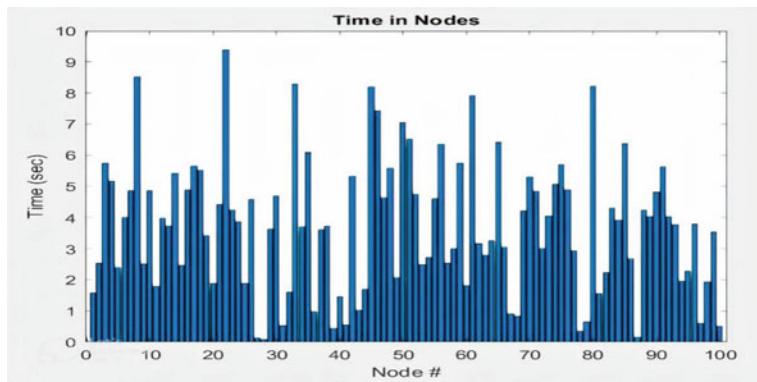


Fig. 7 Time in nodes

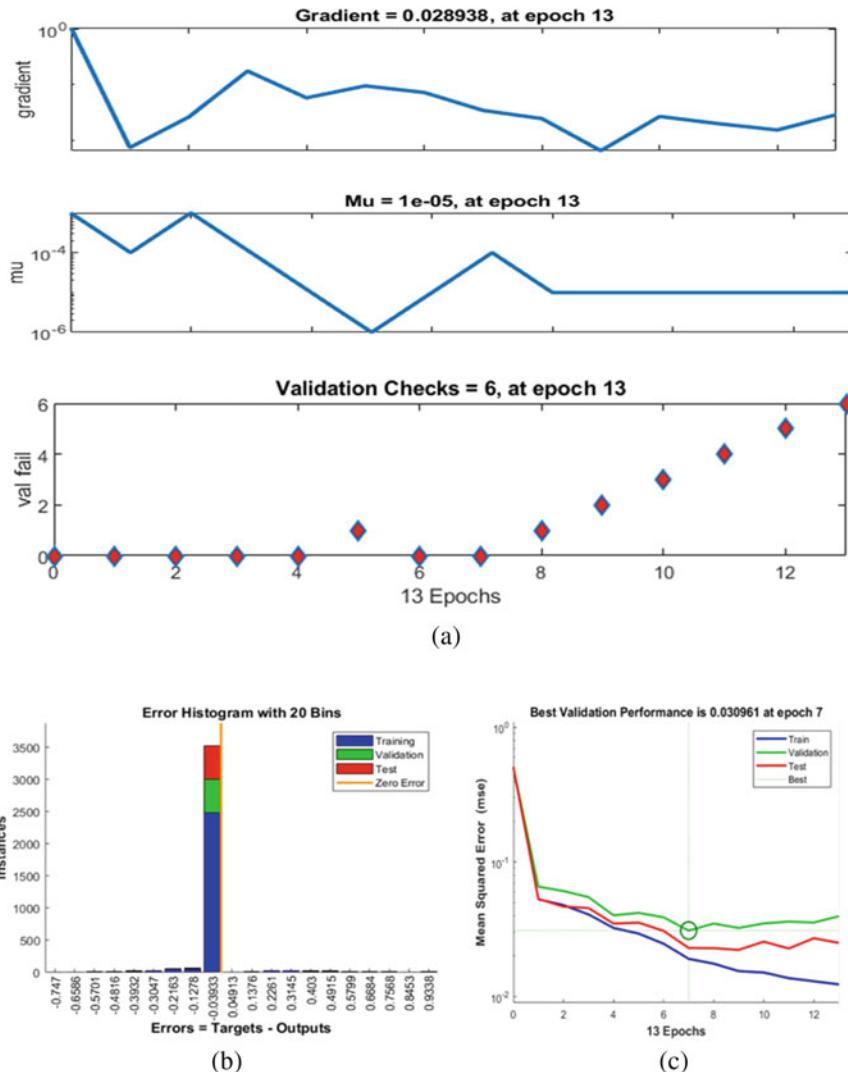


Fig. 8 **a** Training states, **b** Error histogram with 20 bins, **c** Best validation at seven epochs

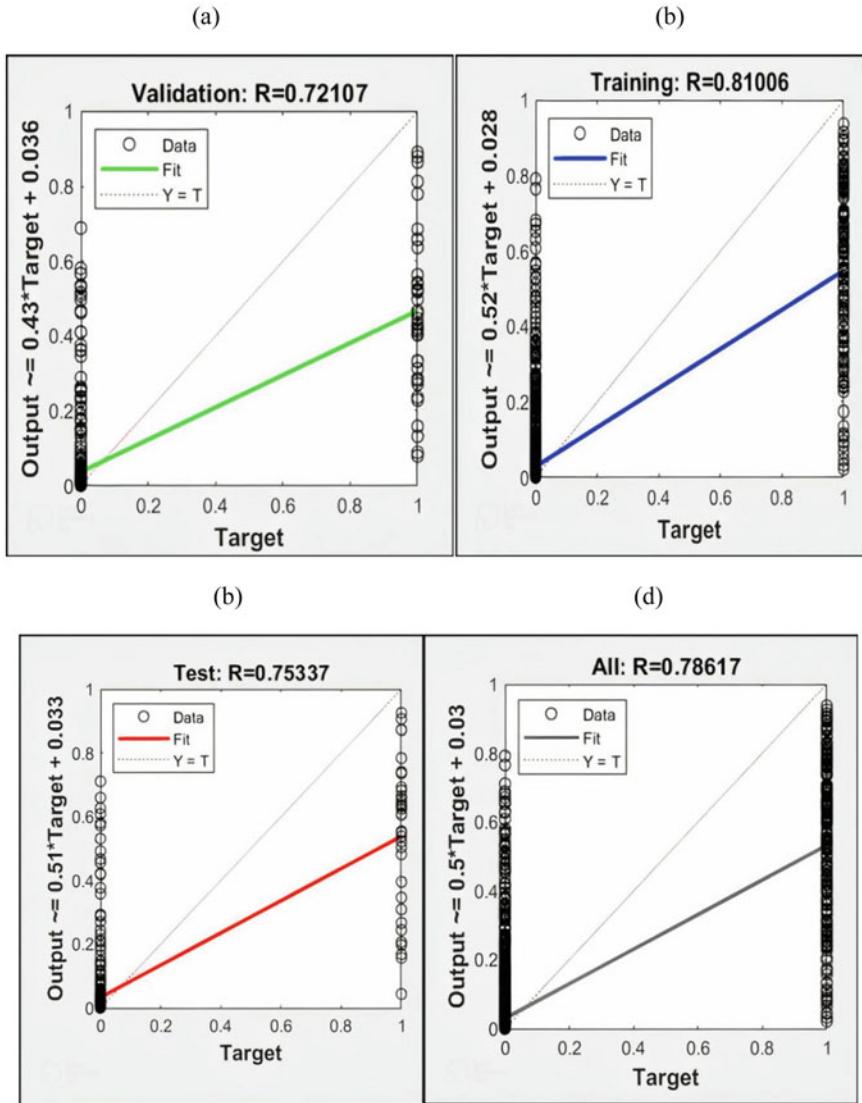


Fig. 9 **a** Training, **b** Validation, **c** Test, **d** All graphs

5 Conclusion

The energy-efficient path selection method is given in this research to balance the competing objectives of optimizing network lifetime and decreasing energy usage routing in mobile ad hoc networks (MANETs). The suggested approach combines deep learning, MANET, and lifetime energy efficiency. After creating the network

by conducting assaults and validating paths, produce the data. Then make a neural network that is capable of performance evaluation and prediction. Then nodes in a network that is, by definition, negative. The best route must then be chosen after this. Utilizing deep neural networks for deep learning, management, and lifetime energy efficiency maximization, anticipate energy and lifetime maximization using the metrics of stolen time, total time, total energy, and packet delivery rate. Predict energy and lifetime maximization using deep neural networks for deep learning, MANET, and lifetime energy efficiency maximization. The metrics are stolen time, total time, total energy, and packet delivery rate. Where the highest value is achieved by 500 pkt of neural network with packets. The total energy is 7570, the packet delivery rate is 74.60, the total duration is 371.81 mins, and the lowest theft rate obtained by 500 packets with packets is 6.8 with an improvement of 100, 200, 300, and 400.

References

1. Yang H (2020) A study on improving secure routing performance using trust model in MANET. *Mob Inf Syst* 2020. <https://doi.org/10.1155/2020/8819587>
2. Pan J, Sui T, Liu W, Wang J, Kong L, Zhao Y (2023) Secure control using homomorphic encryption and efficiency analysis
3. Doi R (2019) Maximizing the accuracy of continuous quantification measures using discrete packtest products with deep learning and pseudocolor imaging. *J Anal Methods Chem.* <https://doi.org/10.1155/2019/1685382>
4. Nair SKG, Soorya VU (2019) Energy efficiency and network lifetime improvement in MANET using AOMDV. *Int J Eng Res Technol* 8(8):10–14
5. Bangotra DY, Singh Y, Kumar N, Kumar Singh P, Ojeniyi A (2022) Energy-efficient and secure opportunistic routing protocol for WSN: performance analysis with nature-inspired algorithms and its application in biomedical applications. *Biomed Res Int.* <https://doi.org/10.1155/2022/1976694>
6. Zhou F et al (2022) A high-efficiency deep-learning-based antivibration hammer defect detection model for energy-efficient transmission line inspection systems. *Int J Antennas Propag.* <https://doi.org/10.1155/2022/3867581>.
7. Veeraiyah D, Joel Sunny Deol G, Ganiya RK, Nageswara Rao J, Bulla S, Alene A (2022) Energetic and valuable path compendium routing using frustration free communication dimension extension algorithm in MANET. *Wirel Commun Mob Comput.* <https://doi.org/10.1155/2022/3685419>
8. Yuan B (2022) A secure routing protocol for wireless sensor energy network based on trust management. *Wirel Commun Mob Comput.* <https://doi.org/10.1155/2022/5955543>
9. Pandey NK, Diwakar M, Shankar A, Singh P, Khosravi MR, Kumar V (2022) Energy efficiency strategy for big data in cloud environment using deep reinforcement learning. *Mob Inf Syst.* <https://doi.org/10.1155/2022/8716132>
10. Kasturi SB, Reddy PV, Venkata Nagendra K, Madhavi MR, Kumar Jha S (2022) An improved energy efficient solution for routing in IoT. *J Pharm Negative. Results* 13(6):1683–1691. <https://doi.org/10.47750/pnr.2022.13.S06.221>
11. Xue L, Ma Y, Zhang M, Qin W, Wang JL, Wu Y (2021) Energy efficiency maximization with optimal beamforming in secure MISO CRNs with SWIPT. *Int J Antennas Propag.* <https://doi.org/10.1155/2021/6378715>

12. Bharti RK et al (2022) Enhanced path routing with buffer allocation method using coupling node selection algorithm in MANET. *Wirel Commun Mob Comput.* <https://doi.org/10.1155/2022/1955290>
13. Ding Q, Zhu R, Liu H, Ma M (2021) An overview of machine learning-based energy-efficient routing algorithms in wireless sensor networks. *Electron* 10(13). <https://doi.org/10.3390/electronics10131539>
14. Yong J, Lin Z, Qian W, Ke B, Chen W, Ji-Fang L (2021) Tree-based multihop routing method for energy efficiency of wireless sensor networks. *J Sens.* <https://doi.org/10.1155/2021/6671978>
15. Khatoon N, Pranav P, Roy S, Amritanjali S (2021) FQ-MEC: fuzzy-based q-learning approach for mobility-aware energy-efficient clustering in MANET. *Wirel Commun Mob Comput.* <https://doi.org/10.1155/2021/8874632>
16. Mishra M, Sen Gupta G, Gui X (2021) Network lifetime improvement through energy-efficient hybrid routing protocol for iot applications. *Sensors* 21(22). <https://doi.org/10.3390/s21227439>
17. Aroulanandam VV, Latchoumi TP, Balamurugan K, Yookesh TL (2020) Improving the energy efficiency in mobile ad-hoc networks using learning-based routing. *Rev d'Intelligence Artif* 34(3):337–343. <https://doi.org/10.18280/ria.340312>
18. Misra Y et al (2022) Secure information collection and energy efficiency in heterogeneous sensor networks using machine learning with the internet of things. *Wirel Commun Mob Comput.* <https://doi.org/10.1155/2022/2874497>
19. Alkadhami MMA, Uçan ON, Ilyas M (2020) An efficient and reliable routing method for hybrid mobile ad hoc networks using deep reinforcement learning. *Appl Bionics Biomech.* <https://doi.org/10.1155/2020/8888904>
20. Tomar MS (2022) A survey on lifetime maximization in MANET. *IEEE Int Conf Curr Dev Eng Technol* 1–6. <https://doi.org/10.1109/CCET56606.2022.10080662>.
21. Natarajan K (2019) Analysis of routing protocols in flying wireless networks. *IRO J Sustain Wirel Syst* 01(03):148–160. <https://doi.org/10.36548/jsws.2019.3.002>
22. Sharma T, Singh H, Sharma A (2015) A comparative review on routing protocols in wireless sensor networks. *Int J Comput Appl* 123(14):28–33. <https://doi.org/10.5120/ijca2015905634>
23. Cheng G, Zhang Z, Li Q, Li Y, Jin W (2021) Energy theft detection in an edge data center using deep learning. *Math Probl Eng.* <https://doi.org/10.1155/2021/9938475>.
24. Revathi P (2020) Quality of service routing in manet using a hybrid intelligent algorithm inspired by ant colony optimization. *Int J Adv Sci Technol* 29(3):4033–4046
25. Singh VK, Sharma V (2014) Elist genetic algorithm based energy balanced routing strategy to prolong lifetime of wireless sensor networks. *Chinese J Eng* 2014:1–6. <https://doi.org/10.1155/2014/437625>
26. Chen Z, Yu H, Wen C (2014) An optimal control method for maximizing the efficiency of direct drive ocean wave energy extraction system. *Sci World J.* <https://doi.org/10.1155/2014/480916>
27. Zungeru AM, Seng KP, Ang LM, Chong Chia W (2013) Energy efficiency performance improvements for ant-based routing algorithm in wireless sensor networks. *J Sensors.* <https://doi.org/10.1155/2013/759654>
28. Nagdive M, Agrawal PA (2014) Maximizing the lifetime of wireless sensor networks using ERPMPT. *Int J Eng Trends Technol* 11(10):498–501. <https://doi.org/10.14445/22315381/ijett-v11p297>
29. Pham V, Larsen E, Kure Ø, Engelstad P (2009) Routing of internal MANET traffic over external networks. *Mob Inf Syst* 5(3):291–311. <https://doi.org/10.3233/MIS-2009-0085>
30. AL-Khdour T, Baroudi U (2009) A generalized energy-efficient time-based communication protocol for wireless sensor networks. *Int J Internet Protocol Technol* 4(2):134–146. <https://doi.org/10.1504/IJIPT.2009.027338>

Automated Air Pollution Monitoring System



**G. Poornima, S. Lakshmi, D. Muthukumaran, T. Dinesh Kumar,
K. Umapathy, N. C. A. Boovarahan, M. A. Archana,
and Ahmed Hussein Alkhayyat**

Abstract This paper presents a system to monitor quality levels as a baseline for assessment by detecting the mixture of gases, vapours and particles in an indoor environment and display them with respect to Air Quality Index (AQI). The present situation in air pollution is mainly due to various physical and chemical activities in industries and vehicles which made the air quality to questionable point. Pollution levels are rising exponentially as a result of factors such as industry, urbanization, population growth and usage of automobiles which all lead to problems harmful to human health. If the air quality drops below a particular threshold, the indicator shows there are harmful gases in the air such as—carbon dioxide, smoke, alcohol, benzene, ammonia and nitrogen oxide. This vital measurement will give appropriate awareness among the public towards healthier life.

Keywords Air · Quality · Microcontroller · Sensor · Management

1 Introduction

The pollution in air is due to emission of toxic gases by companies, automobiles and certain particles present in the atmosphere. This is a real challenge to health of human beings. This situation creates a necessity for evaluating the quality of

G. Poornima · T. Dinesh Kumar · K. Umapathy (✉) · N. C. A. Boovarahan · M. A. Archana
SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India
e-mail: umapathykannan@gmail.com

S. Lakshmi
Thirumalai Engineering College, Kanchipuram, Tamil Nadu, India

D. Muthukumaran
Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

A. H. Alkhayyat
Scientific Research Centre of the Islamic University, The Islamic University, Najaf, Iraq

air in the environment so that appropriate steps can be taken in time. The objective is to construct an air quality monitoring system specifically designed for urban environments. Urban areas face significant air pollution challenges due to increased industrialization, traffic congestion and population density. Monitoring air quality is crucial for public health, environmental sustainability and policy-making. This proposed system will address the need for accurate and real-time data collection, analysis and visualization of air pollutants in urban areas. By addressing the need for a comprehensive air quality monitoring system in urban environments, this can contribute to create healthier and more sustainable cities, protecting public health and enabling evidence-based decision-making for air pollution mitigation strategies. The basic theory behind an air quality monitoring system involves the measurement and analysis of various air pollutants present in the atmosphere. The system typically consists of sensors, data collection mechanisms, data analysis algorithms and visualization tools. The quality of air will be monitored and analysed every day using a web server. An alarm will be triggered whenever the quality of air crosses a threshold level thereby indicating the content of toxic gases present in the environment.

2 Literature Review

Duangsuwan et al. reported a IoT-based architecture to monitor the content of pollution in the nearby environment [1, 2]. But calibration was not done appropriately. Tzortzakis et al. enunciated a monitoring arrangement for air pollution using the concept of LoRa [3]. The used sensors depend on mobile gateway for the cloud connection. Again LoRa was implemented for air monitoring using raspberry Pi controller [4, 5]. But the system is not suitable for huge scale implementation. The air pollution was carried out by employing PM sensors with LoRa arrangement [6, 7]. A LoRa dependent sensor node was developed to track parameters such as carbon dioxide, humidity and temperature [8, 9]. The discussion was done on usage of statistical models for monitoring and forecasting. An economical air quality tracking system was also demonstrated for monitoring various gases [10, 11]. This system can manipulate both quality of air index and air status in real-time mode. To measure quality of air in case of indoor environment, kalman filtering was implemented to correlate gas transport with consumption of energy [12, 13]. Li et al. implemented filter dependent methodologies in order to enhance the system accuracy and reliability for forecasting of energy [14, 15].

3 Materials and Methods

Air quality monitoring systems utilize advanced sensor technologies to detect and measure pollutant concentrations accurately. The monitoring system collects data from sensors placed in different locations within a region. These measure pollutant

levels at specific intervals and transmit the data to a central database or server. The collected data is analysed to determine the air quality index and identify trends, patterns and potential sources of pollution. Data analysis techniques such as statistical analysis, machine learning algorithms and geospatial mapping are employed to process and interpret the information effectively. Air quality monitoring systems generate reports and alerts based on the analysed data. These reports provide information on pollutant levels, trends and potential health risks. The design of an air quality monitoring system involves both hardware and software components. The sensors chosen for air pollutants are carbon dioxide (CO₂), ozone (O₃) and volatile organic compounds (VOCs). The complete system is shown in Fig. 1.

It provides accurate measurements of air pollutants to ensure reliable assessment of air quality. The sensors used have a high level of accuracy and calibration which are capable of measuring pollutants with a high degree of precision to detect even minor changes in air quality. These sensors detect low concentrations of pollutants to capture even slight variations in air quality and capable of operating continuously thereby providing accurate measurements over extended periods. They include selecting high quality sensors and components, thereby implementing redundancy measures and

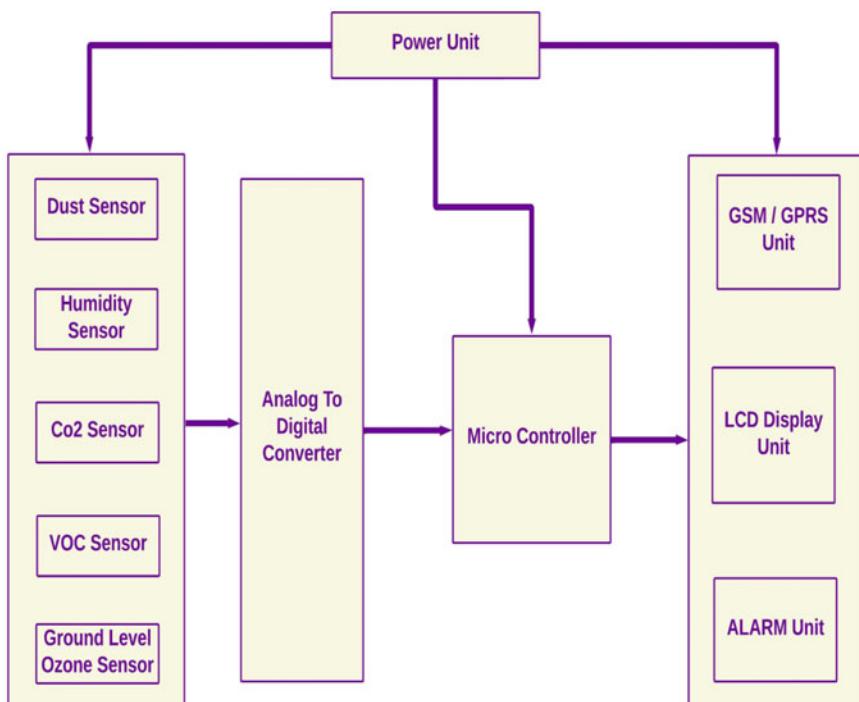


Fig. 1 Arrangement for proposed system

ensuring proper maintenance cum calibration. This proposed system provides real-time measurements to enable prompt identification and response to changes in air quality. The response time is minimized to provide timely alerts and warnings.

Hardware Requirements

Arduino controller is employed for implementation of system. A light scattering laser sensor is used for measurements as shown in Fig. 2. A beam of light is focused into particles in all directions and a detector is used to detect the light scattered from particles. The sensor will be able to manipulate the particle concentration within that particular region. The range of those particles lies between 0.2 and 2.4 μm .

DHT22 sensor is employed for measurement of humidity and temperature. It involves a thermistor arrangement and provides a signal at respective data pin. The new data can be obtained for every 2 s. The non-dispersive infrared (NDIR) CO₂ sensor is very easy to use, fantastic, inexpensive and accurate sensor as shown in Fig. 3. This is used for detection of the carbon dioxide in the air with good selectivity.

MP503 module provides high level of sensitivity along with features such as consumption of low power and longevity. This module has been cleaned and calibrated. It detects VOCs such as ammonia, hydrogen gas, carbon monoxide. The voltage organic compound (VOC) sensor is illustrated in Fig. 4.

MQ131 sensor has high level of sensitivity to gases such as ozone, carbon dioxide and NO₂. It responds to disturbances of organic gases in the other way. The ground

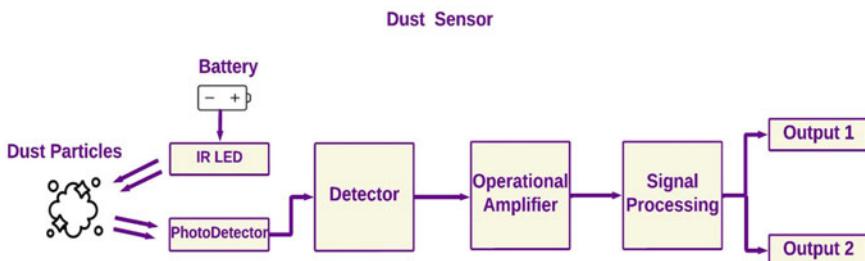


Fig. 2 Dust sensor

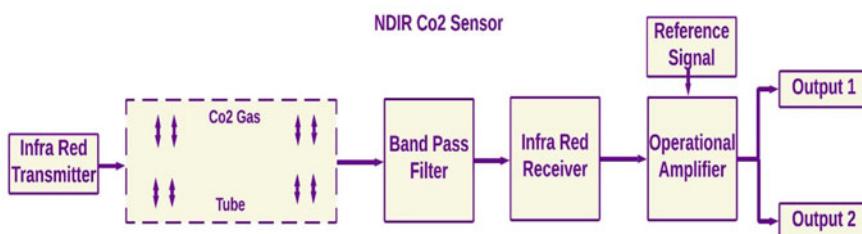


Fig. 3 Carbon-Di-Oxide (CO₂) sensor

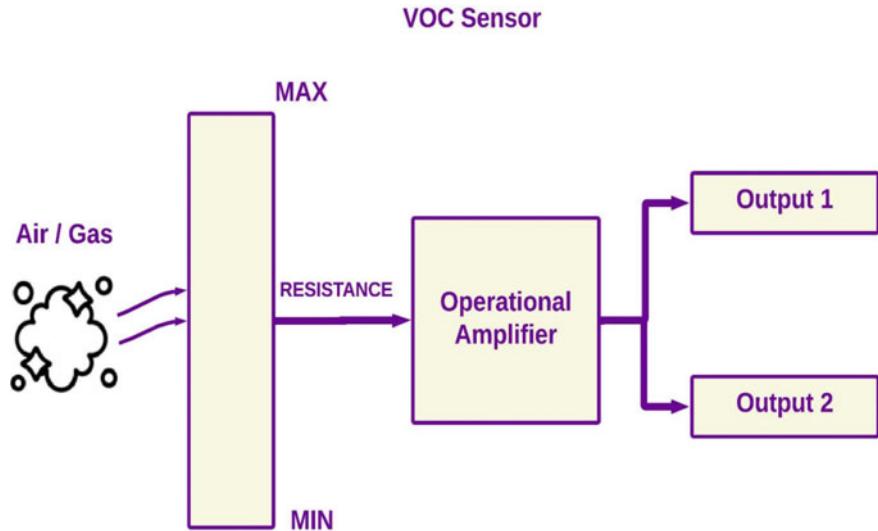


Fig. 4 Volatile organic compound (VOC) sensor

level ozone sensor is shown in Fig. 5. Flowchart of proposed system is illustrated in Fig. 6.

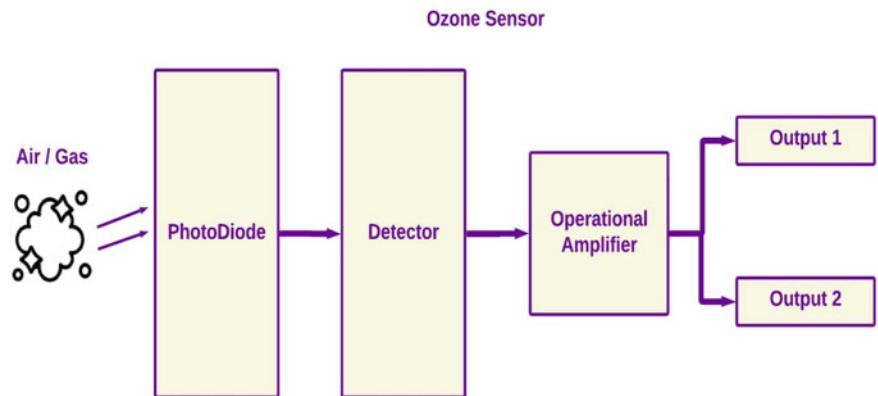
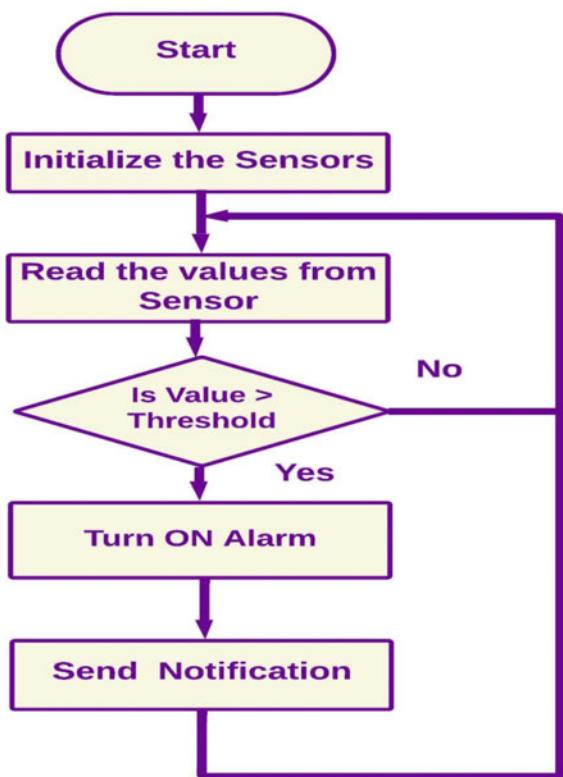


Fig. 5 Ground level ozone sensor

Fig. 6 Flowchart of proposed system



4 Results and Discussion

Figure 7 shows hardware connection of proposed system and Fig. 8 shows the total arrangement. Figure 9 shows the connection of CO₂ sensor and ground level ozone sensor connected laterally to the model. Figure 10 shows the levels of ozone, carbon monoxide, lead and ammonia concentration during third week of June 2023 using the proposed system.

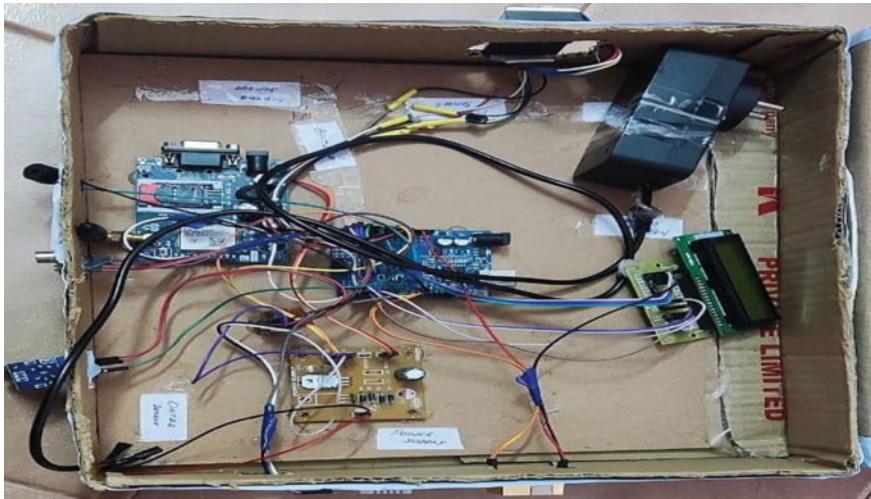


Fig. 7 Hardware connection of proposed system

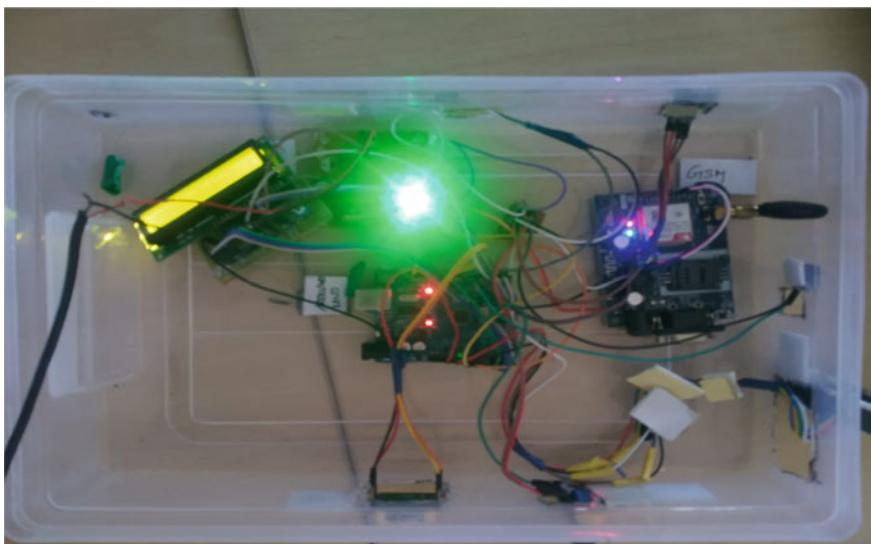


Fig. 8 Hardware setup of proposed system



Fig. 9 Co₂ sensor and ozone sensor connected to the model

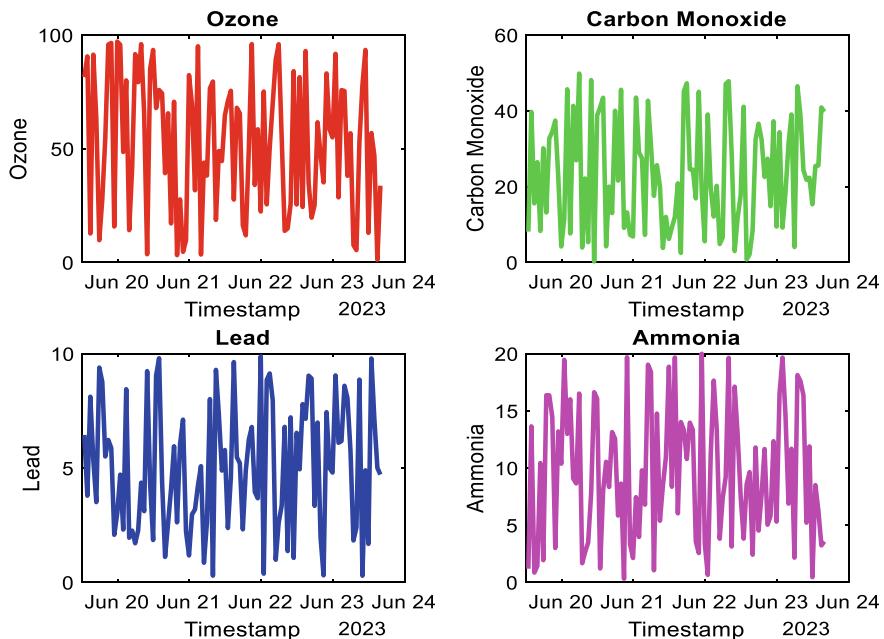


Fig. 10 Levels of ozone, carbon monoxide, lead and ammonia concentration during third week of June 2023

5 Conclusion

An air quality monitoring system is designed and constructed to measure and analyse the level of pollutants present in the air. This helps to assess the overall air quality in and around nearby environment. This valuable data is essentially required for environmental monitoring, public health and decision-making processes. The system aims to project the air quality and its impact on human health and the environment. They are often made available to the public through online platforms, mobile applications so that local authorities can raise awareness and facilitate informed decision-making. The future scope for air quality monitoring systems may include the use of nanotechnology, advanced optical sensors and miniaturized sensor arrays for detecting a wider range of pollutants with higher precision.

References

1. Duangsuwan S, Takarn A, Jamjareegulgarn P (2018) A development on air pollution detection sensors based on NB-IoT network for smart cities. In: 18th International symposium on communications and information technologies (ISCIT), pp 313–317
2. Duangsuwan S, Tákarn A, Nujankaew R, Jamjareegulgarn P (2018) A study of air pollution smart sensors LPWAN via NB-IoT for Thailand smart cities 4.0. In: 10th International conference on knowledge and smart technology (KST), pp 206–209
3. Tzortzakis K, Papafotis K, Sotiriadis PP (2017) Wireless self powered environmental monitoring system for smart cities based on LoRa. In: Panhellenic conference on electronics and telecommunications (PACET), pp1–4
4. Rosmiati M, Rizal MF, Susanti F, Alfisyahrin GF (2019) Air pollution monitoring system using LoRa modul as transceiver system. *Telkomnika* 17(2):586–592
5. Rossi M, Tosato P (2017) Energy neutral design of an IoT system for pollution monitoring. In: IEEE workshop on environmental, energy, and structural monitoring systems (EESMS), pp1–6
6. Sujuan L, Chuyu X, Zhenzhen Z (2016) A low-power real-time air quality monitoring system using LPWAN based on LoRa. In: 13th IEEE international conference on solid-state and integrated circuit technology (ICSICT), pp 379–381
7. Hsieh CL, Ye ZW, Huang CK, Lee YC, Sun CH, Wen TH, Juang JY, Jiang JA (2017) A vehicle monitoring system based on the LoRa technique. *World Academy of Science, Engineering and Technology*. *Int J Mech Aerosp Ind Mechatron Manuf Eng* 11(5):1093–1099
8. Thu MY, Htun W, Aung YL, Shwe PEE, Tun NM (2018) Smart air quality monitoring system with LoRaWAN. In: IEEE international conference on internet of things and intelligence system (IOTAIIS), pp 10–15
9. Rahmat M, Maulina W, Isnaeni, Miftah DYN, Sukmawati N, Rustami E, Azis M, Seminar KB, Yuwono AS, Cho YH, Alatas H (2014) Development of a novel ozone gas sensor based on sol–gel fabricated photonic crystal. *Sens Actuators A Phys* 220:53–61
10. Carminati M, Ferrari G, Grassetti R, Sampietro M (2012) Real-time data fusion and mems sensors fault detection in an aircraft emergency attitude unit based on kalman filtering. *IEEE Sens J* 12(10):2984–2992
11. Li X, Wen J (2016) System identification and data fusion for on-line adaptive energy forecasting in virtual and real commercial buildings. *Energy and Build* 129:227–237
12. Dinesh Kumar T, Archana MA, Umapathy K, Gayathri G, Bharathvaja V, Anandhi B (2023) RFID based smart electronic locking system for electric cycles. In: 4th International conference on electronics and sustainable communication systems (ICESC). Coimbatore, India, pp 76–81

13. Dinesh Kumar T, Archana MA (2022) Fundamentals of internet of things and its applications. Alpha International Publication
14. Umapathy K, Mangayarkarasi T, Subitha D, Sivagami A (2021) Android application and SMS alert based garbage monitoring and navigation system. *J Phys Conf Ser* 1964(6):062064-1 to 062064-7
15. Umapathy K, Omkumar S, Muthukumaran D, Chandramohan S, Sivakumar M (2023) Things-peak based garbage monitoring and collecting system. In: Lecture notes in networks and systems, vol 617. Springer Nature, pp 235–242

Simulation and Implementation of Solar Charge Controller by MPPT Algorithm



D. Vanitha, V. Malathi, and K. Umapathy

Abstract Power is one the most fundamental requirements for people in the present. Change of sun-based energy into power further develops age of power as well as decreases contamination because of petroleum derivatives. The result force of sunlight-based charger relies upon sun powered irradiance and heap impedance. Since heap impedance relies upon application, a DC converter is utilized for working on exhibition of sun powered charger. Sun-based irradiance and temperature are dynamic. Thus Internet-based calculation that progressively registers working mark of sunlight powered charger is required. The effective change of sunlight-based energy is conceivable with greatest power point following (MPPT) calculation. The different calculations in MPPT and their geography are talked about in this paper.

Keywords Power · Tracking · Energy · Converter · Voltage · Conductance

1 Introduction

The essential part of PV system is to monitor power point of photovoltaic array at maximum. Various strategies are examined in light of their execution in order to identify different neighborhood applications [1]. Sun powered parameters such as temperature and irradiance are not static. Thus Internet-based calculation that progressively figures the working place of the sun powered charger is required. The proficient change of sunlight-based power is conceivable with maximum power point tracking (MPPT) calculation. Various types of calculations are available with MPPT [2]. Sun power can be involved these days as generally dependable and ecological

D. Vanitha · V. Malathi

Department of EEE, Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Kanchipuram, Tamil Nadu, India

K. Umapathy (✉)

Department of ECE, Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Kanchipuram, Tamil Nadu, India

e-mail: umapathykannan@gmail.com

well-disposed energy source. To achieve better efficiency, greatest energy can be extricated from PV board by utilizing MPPT strategies. Innovation can be obtained by alluring choice in light of the fact that the highlights different benefits like as low upkeep necessity and ecological amicability [3]. The non-linear qualities correspond to conditions of weather with respect to sun oriented isolation. They are employed to support greatest power and speedy response from respective clusters of PV [4]. Integrated sun powered cells should give improved proficiency instead of their single intersection partners. Neuro Fuzzy-based calculation is planned using Brain conduct and fluffy rationale. The terminating point's ideal worth is determined and taken care of to converter [5, 6]. An unique Versatile Neuro-Fuzzy Deduction Framework for most extreme Pinnacle Power Move strategy for integrated sunlight-based cells is presented. Outputs are contrasted when compared to gradual strategy of conductance and performance of MPPT seem to be good compared to its different partners as far as transient state and the extent of voltage acquired [7, 8].

2 Proposed Method

The representation of proposed concept is illustrated in Fig. 1. The current sensor and voltage sensor senses respective values and interfaced with PIC (PERIPHERAL INTERFACE CONTROLLER) to generate the gate pulse to the buck/boost converter. Using these values, power is calculated by using formula $P = V * I$. The data is processed through PIC. The data of current, voltage, and power are displayed on the LCD for user interaction and given to load [9, 10].

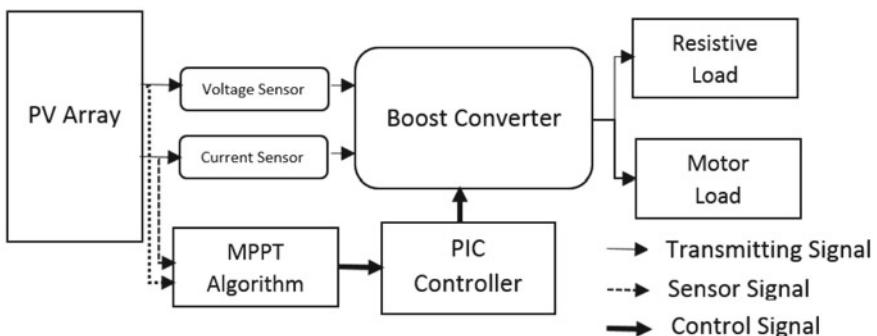


Fig. 1 Block diagram of proposed method

3 Boost Converters

In Fig. 2, a lift converter is planned in Simulink. In the circuit it comprises of DC voltage source, inductor, exchanging gadget (MOSFET) with heartbeat generator, diode, capacitor, obstruction, and a voltage estimation block is utilized to gauge the voltage. First DC voltage supply is given to the inductor which fixed inductance esteem (say as $L = 1 \text{ mH}$) then a MOSFET is associated with heartbeat generator to set off the pulse diode is utilized to make that the voltage streams in forward bearing then to decrease the waves id DC voltage waveform capacitor is utilized as channels. Yield diagrams are displayed in Fig. 3.

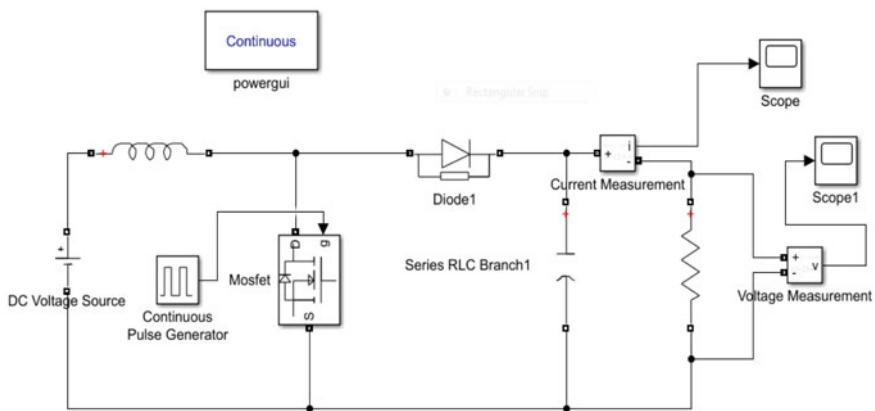


Fig. 2 Conventional boost converter circuit

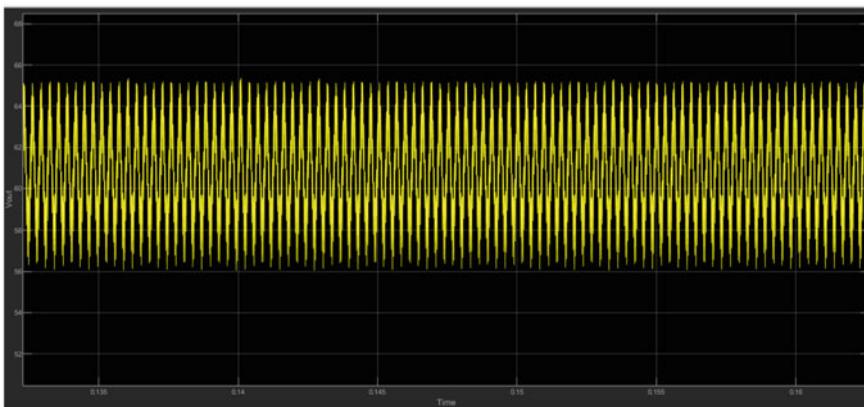


Fig. 3 Voltage waveform of boost converter

4 Pertrub and Observe Method

The calculation looks at power and voltages of time (K) with example at once and predicts an opportunity to way to deal with MPP. A little irritation in voltage modifies force of sunlight-based charger in the event that power modification going in a similar track. Yet, assuming delta power is negative, it demonstrates that MPP is far away and annoyance is diminished to arrive at the MPP. Figure 4 shows synopsis of P&O calculation [11, 12]. Accordingly, entire PV bend is actually looked at by little bothers to find MPP that expands reaction season of calculation. Alternately, when size of perturbation is amplified, it produces consistent oscillations with respect to MPP.

4.1 Simulink Diagram

Here in the Fig. 4 a set of 4 solar cells are placed in series with voltage value of 3 V with irradiance of 1000 w/m^2 . For measurements, ammeter and voltmeter are connected across the solar cells then it is connected to a boost converter which consists of switching element (MOSFET). Here the current and voltage values are given to the MPPT algorithm which compares the both the values according to that duty cycle is change this generated pulse is given to switching element as pulse trigger and output graph are shown in Fig. 5.

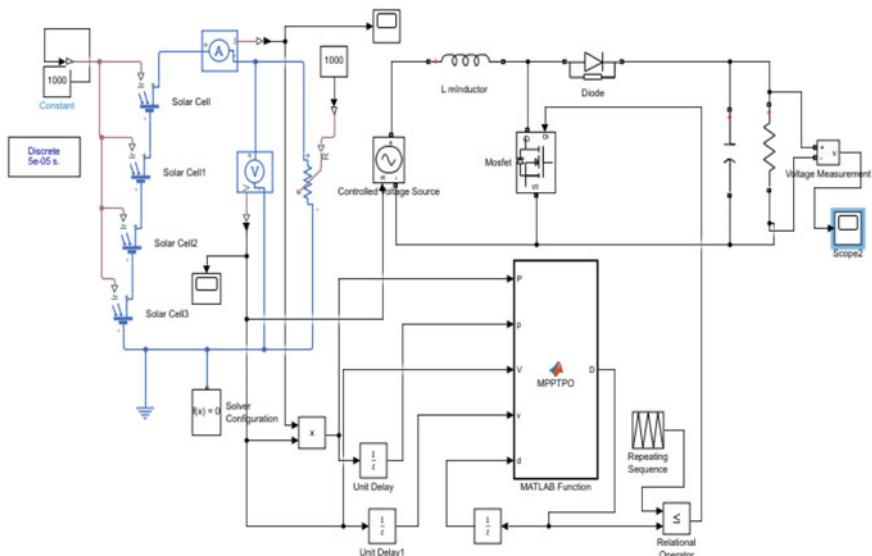


Fig. 4 P&O simulation diagram

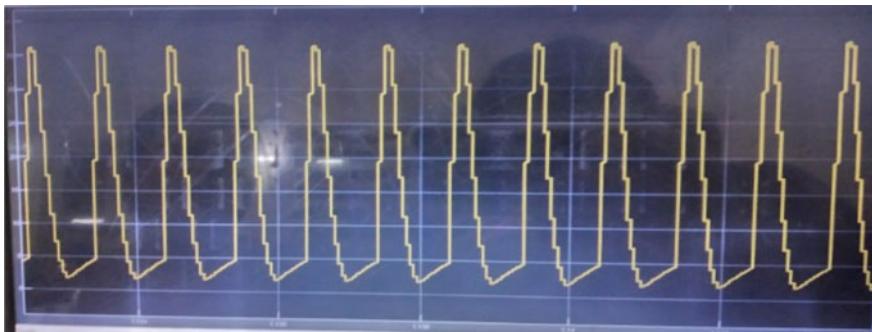


Fig. 5 Output voltage waveform of P&O method

4.2 *Output Waveforms*

See Fig. 5.

5 Incremental Conductance Method

MPPT is the main approach to expanding the proficiency of the sun oriented cell by removing the greatest power from the sunlight powered charger and conveying steady voltage regardless of variety in sun-based radiation. For example, sun-based power is conveyed straightforwardly to heap voltage from fall to nothing with DC technique. Consequently a framework with MPPT provides breakdown of voltage by maintaining working point close to most extreme point of power. An extensive variety of MPPT calculations are accessible [13, 14]. Of the multitude of accessible calculations Gradual Conductance Algorithm as represented in equation loans itself well to the DSP and Microcontroller. A correlation between the Irritate and Notice (P and O) and the Steady Conductance Calculation (INC) uncovers that the effectiveness of P and O technique is 95% and INC Calculation is 98.2%. Thus among every one of the strategies examined above Steady Conductance Calculation is viewed as best method and effectively versatile to the changing ecological circumstances. Steady Conductance Calculation.

5.1 *Simulation Diagram*

Here in Fig. 6, a set of 3 sun powered cells are associated in series with a voltage worth of 3 V with irradiance of 1000 W/m^2 to measure the current and voltage values ammeter and voltmeter are associated across the sun-based cells then it is associated

with a lift converter which comprises of exchanging component (MOSFET). Here the current and voltage values are given to the MPPT calculation which thinks about the both the qualities as per that obligation cycle is change this created beat is given to exchanging component as heartbeat trigger result. Output is displayed in Fig. 7

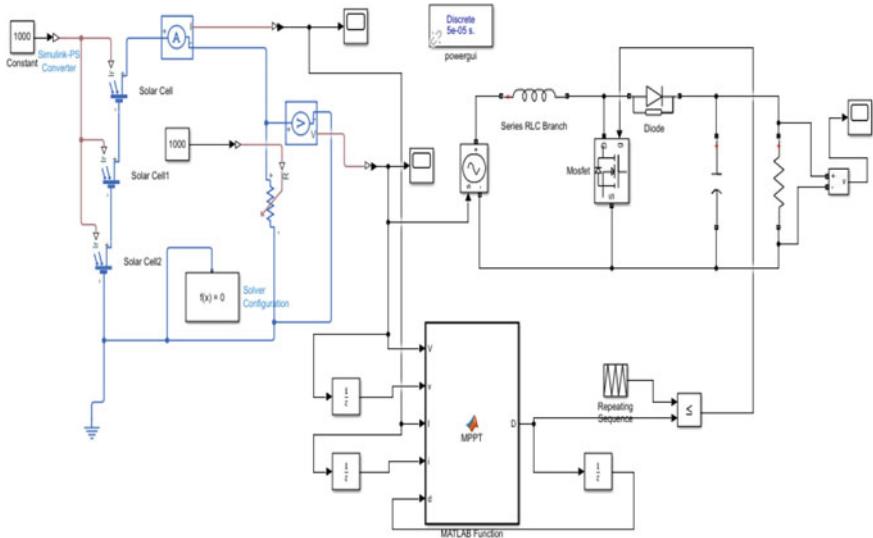


Fig. 6 INC simulation diagram

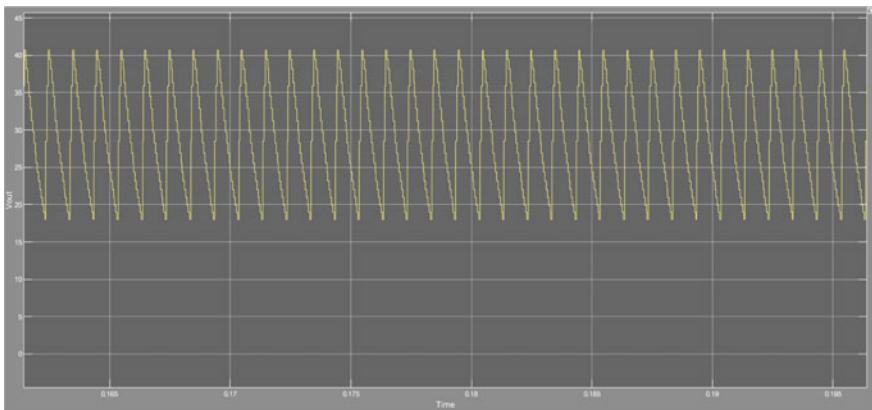


Fig. 7 Output voltage waveform of P&O method

6 Fuzzy Logic Method

Rule bases are created to control yield factors in a fluffy rationale framework. The principal point of fluffy frameworks is to make a hypothetical starting point for to make consistent suppositions and relationship among inaccurate terms of reference. In fluffy rationale mechanical frameworks, this is called rough thinking [15–18]. The primary subjects and action words of fluffy rationale are included fluffy sets and fluffy administrators, or something like that called “On the off chance that” rule articulations. These are applied while fostering the “On the off chance that” restriction proclamations that are the trademark and premise of fluffy rationale. In this way, for example, on the off chance that a fluffy decide declares that assuming that x is A, y should be B if A and B are values assigned by fluffy sets specifying a scope of X and Y, separately. In this standard, the “In the event that” segment is considered the forerunner or reason, while the piece is alluded to as the restrictive ensuing or end. Eventually, a fluffy rule is, at its center, an on off chance that standard that highlights the two circumstances and an end.

6.1 Simulation Diagram

Here a PV ARRAY with 52 individual cells with the temperature of 25 centigrade and irradiance of 1000 W/m^2 . from this array the voltage and power values are take and given to the boost converter and reference value and instantaneous of dv and dp here voltage and current is take as product. Memory block is used to store the previous values of power and voltage. “dv” and “dp” is given to multiplexer and to select a particular signal. A saturation block is used. The MUX input is given to fuzzy logic controller as shown in the Fig. 8.

Here in this block the set of fuzzy rules are inserted as coding this set of rules helps to maintain the higher efficiency with change of voltage and current values according to climatic changes. From the saturation block the signal from PWM signal generator is given to boost converter as pulse generator for the switching device (MOSFET). Figure 9 shows the output.

6.2 Output Waveform

See Fig. 9.

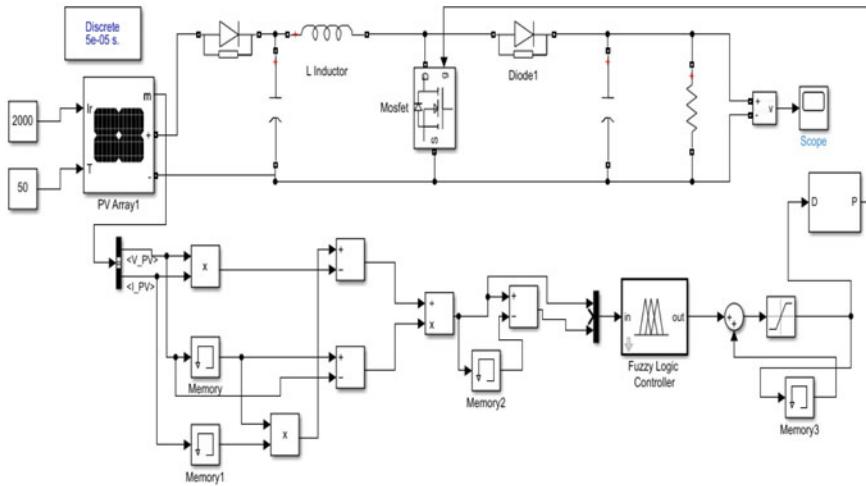


Fig. 8 Fuzzy logic control simulation diagram

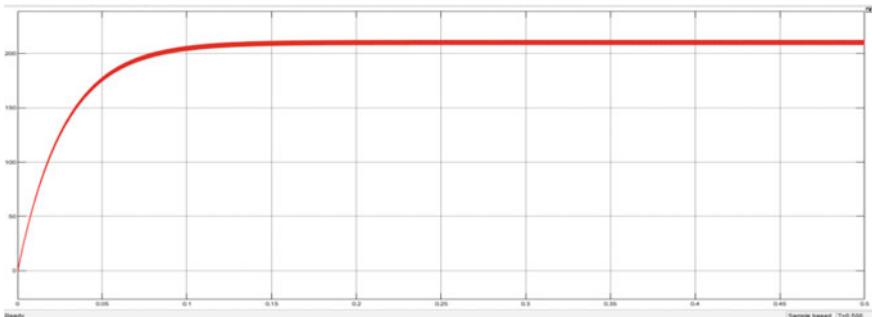


Fig. 9 Output voltage waveform of fuzzy logic control method

7 Hardware Implementation

Here in this equipment execution displayed in Fig. 10, Buck-Lift Converter Circuit comprises of DC power source, buck-support circuit, drive circuit (PIC), loads there are two DC voltage supplies are taken as input.

One of them is taken from sunlight-based charger or, in all likelihood from direct AC supply later changed over into DC voltage then voltage supply is given to help converter. Here in this circuit it comprises of exchanging component (MOSFET) one switch is buck and another is lift, inductor and capacitor. In order to create the beat to MOSFET (PIC) is utilized to produce the setting off heartbeat to MOSFET. Inductor raise the voltage with exchanging of MOSFET and capacitor is utilized to

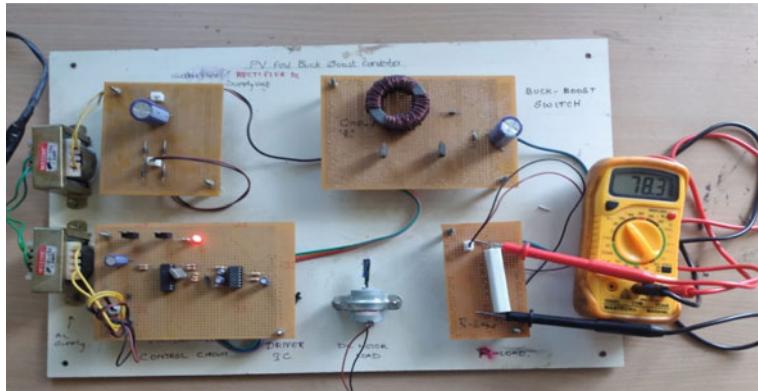


Fig. 10 Prototype hardware implementation

lessen the waves made in delivered DC voltage then voltage is given burden. Here in this circuit two kinds of burdens are there, they are DC engine and resistive burden.

8 Results and Discussion

The choice of MPPT method relies upon the particular prerequisites of the lift converter-based PV framework, like following exactness, execution intricacy, dynamic reaction, cost, and strength. While P&O is the most straightforward and most economical procedure to carry out, IC offers quicker unique reaction, and FLC gives the most noteworthy following exactness and power. A cautious examination of the advantages and downsides of every procedure can help decide the most fitting MPPT strategy for a specific lift converter-based PV framework application. FLC can possibly accomplish the most elevated following precision because of its capacity to adjust to changing ecological circumstances and streamline the PV framework's exhibition. IC is likewise equipped for accomplishing high following precision, especially under quickly changing natural circumstances, while P&O can accomplish a sensibly high following exactness under stable ecological circumstances. Eventually, the decision of MPPT procedure relies upon the particular necessities of the lift converter-based PV framework, like the ideal degree of following precision, the expense, and the intricacy of the execution.

The execution intricacy of the MPPT strategies in a lift converter-based PV framework follows the request for $P\&O < IC < FLC$. P&O is the least difficult and most generally accessible MPPT method, while FLC is the most mind boggling and requires particular information. The decision of MPPT procedure relies upon the particular prerequisites of the PV framework, like the ideal degree of following precision, the expense, and the intricacy of the execution.

Table 1 Comparison of MPPT methods

Types of MPPT techniques	Tracking accuracy	Implementation complexity	Dynamic response	Cost	Speed
Perturb and Observe	High	Simple	Slow	Low	Low
Incremental conductance	Low	More complex	Fast	High	Low
Fuzzy logic	Low	More complex	Fastest	High	High

The powerful reaction of the MPPT strategies in a lift converter-based PV framework follows the request for $P\&O < IC < FLC$. P&O has a somewhat sluggish unique reaction, while IC has a quicker reaction and is less inclined to motions. FLC has the quickest reaction and adjusts to changing PV conditions, bringing about a high following precision. The decision of MPPT procedure relies upon the particular necessities of the PV framework, for example, the ideal degree of following precision and the speed of reaction required for the application. P&O and IC strategies are easier and more affordable to carry out contrasted with FLC. Notwithstanding, the expense of execution can change contingent upon the particular necessities of the PV framework and the degree of precision and execution wanted. At times, the greater expense of executing FLC might be legitimate assuming that it gives better execution and exactness contrasted with P&O and IC procedures. Rundown of correlation are displayed in the Table 1.

P&O and IC procedures are basic and reasonable however may need vigor under quickly changing ecological circumstances. FLC offers higher exactness and heartiness yet requires more equipment assets and programming advancement. The decision of method relies upon the particular prerequisites of the PV framework, including the degree of precision and vigor required, the accessible equipment assets, and the expense and intricacy of execution.

9 Conclusion

P&O and IC methods are basic and economical yet may need heartiness under quickly changing ecological circumstances. FLC offers higher exactness and heartiness yet requires more equipment assets and programming advancement. The decision of method relies upon the particular prerequisites of the PV framework, including the degree of precision and vigor required, the accessible equipment assets, and the expense and intricacy of execution. From the above results the effectiveness of result voltage from sun powered charger is expanded with the assistance of maximum power point tracking is contrasted with ordinary board productivity. The model equipment model has been executed.

References

1. Esram T, Chapman PL (2007) Comparison of photovoltaic array maximum power point tracking techniques. *IEEE Trans Energy Convers* 22(2)
2. Hussein KH, Muta I, Hoshino T, Osakada M, Maximum photovoltaic power tracking: an algorithm for rapidly changing atmospheric conditions. *IEEE Proc Genre Transm Distrib*
3. Ananth DVN (2012) Performance evaluation of solar voltaic system using maximum power tracking algorithm with battery backup. *IEEE*
4. Sanjeeva Reddy BR, Narayana PB et al, PPT algorithm implementation for solar photovoltaic module using microcontroller
5. Ratsame C (2012) A new switching charger for photovoltaic power system by soft-switching. In: Proceedings on 12th international conference on control, automation and systems. ICC, Jeju Island, Korea, pp 17–21
6. Prasad SY et al (2010) Microcontroller based intelligent DC/DC converter to track maximum power point for solar photovoltaic module. *IEEE*
7. Sarvi M, Azadian A (2022) A comprehensive review and classified comparison of MPPT algorithms in PV systems. *Energy Syst* 13:281–320
8. Vanitha D, Rathinakumar M (2019) Modeling and investigation on PV based Luo converter and buck-boost converter with coupled inductor systems. *J Adv Res Dyn Control Syst* 11(01):741–751
9. Vanitha D, Rathinakumar M (2017) Fractional order PID controlled PV buck boost converter with coupled inductor. *Int J Power Electron Drive Syst (IJPEDS)* 8(3):1401–1407
10. Vanitha D, Jayanthi G (2016) Performance enhancement of PV system using Luo converter and FLC based P&O MPPT. *Middle-East J Sci Res* 24(S2):216–220
11. Boovarahan NCA, Umapathy K (2020) Power allocation based on channel state information in massive MIMO system. *IOP Conf Ser Mater Sci Eng* 12(2):1–9
12. Krishnamoorthy R, Umapathy K (2018) Design and implementation of micro-strip antenna for energy harvesting charging low power devices. In: Proceedings on 4th IEEE international conference on advances in electrical and electronics, information, communication and bio-informatics, Chennai
13. Sivakumar GP, Umapathy K (2017) Closed loop fuzzy logic controlled Class-F3 amplifier system with improved dynamic response. In: Proceedings on 2nd IEEE international conference on electrical, computer and communication technologies, Coimbatore
14. Balasubramaniam A, Umapathy K, Periyarselvam K, Malarvizhi C (2023) IoT and cloud server based industry power monitoring and management system. In: Proceedings on third IEEE international conference on power, energy, control and transmission systems, Chennai, pp 1–4
15. Umapathy K, Sridevi T, Navyasri M, Anuragh R (2020) Real time intruder surveillance system. *Int J Sci Technol Res (IJSTR)* 9(3):5833–5837
16. Vanitha D, Rathinakumar M (2018) Photovoltaic based proportional resonant controlled buck-boost converter with coupled inductor. In: IEEE Conference at Saveetha Engineering College, IEEE Digital Library
17. Malathi V, Sentamil Selvan S, Meikandasivam S (2022) Digital hysteresis control algorithm for switched inductor quasi Z source inverter with constant switching frequency. *Int J Electr Electron Res* 10(3):572–578
18. Dinesh Kumar T, Archana MA, Umapathy K, Gayathri G, Bharathvaja V, Anandhi B (2023) RFID based smart electronic locking system for electric cycles. In: 4th International conference on electronics and sustainable communication systems (ICESC). Coimbatore, India, pp 76–81

Nanoscale Multi-gate Graded Channel DG-MOSFET for Reduced Short Channel Effects



Ashutosh Pandey, Kousik Midya, Divya Sharma, and Seema Garg

Abstract Aggressive scaling of MOSFET results in degradation in their performance due to different short channel effects. The problem was addressed by use of high-k dielectric followed by different structural modifications. Among different modified devices, DG-MOSFET seems to be promising one due to scalability, simple structure and offering better control on channel. However, to obtain enhanced performance further structural modification of DG-MOSFET was introduced. This study aims at performance analysis of DG-MOSFET with graded channel and multi-gate. The results clearly indicate reduced DIBL effect for modified DG-MSFET structure.

Keywords DG-MOSFET · Drain induced barrier lowering · Simulation

1 Introduction

The major advantages of MOSFET are higher speed and lower power consumption. Although both the mentioned features are affected significantly in sub-100 nm regime due to increased short channel effects (SCEs).

Initially high-k materials were introduced to counter SCEs. But high-k materials fail to solve problems like drain induced barrier lowering (DIBL), hot carrier effects, (ION/IOFF) ratio [1]. Afterwards different structural modifications were introduced by the researchers. Devices like different types of dual gate MOSFETs, FinFETs, Tunnel FETs were proposed to reduce the short channel effects. DG-MOSFET became popular as it can provide simpler structure, better scalability and superior control on channel.

Different substrate materials with higher carrier mobility have been used to overcome the SCEs. However, use of material other than Si invites higher cost and fabrication complexities.

A. Pandey · K. Midya (✉) · D. Sharma · S. Garg

Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh 201009, India
e-mail: kmidya@gmail.com

Gate engineering and channel engineering are other approaches to obtain higher mobility and reduced hot carrier effects with the Si substrate. Different gate materials introduce different work function difference with the channel. Therefore, suitable gate material can produce step in band/surface potential distribution. Such distribution of surface potential can increase the carrier mobility and decrease in hot carrier effect. R. Ramesh reported the distribution of surface potential for single metal DG-MOSFET (SMDG), dual metal DG-MOSFET (DMDG) and tri metal DG-MOSFET (TMDG) [2]. His work shows, lowest value of surface potential was obtained for TGMOSFET and consequently most prominent subthreshold swing. TMDG has also been reported there with lowest DIBL value.

Hence, increased carrier mobility can be obtained without sacrifice in the device lifetime. Gate electrode with higher work function is used in the source end to achieve this in n-channel MOSFET. Such gate geometry produces abrupt increase in electric field and carrier mobility. Furthermore, it reduces electric field in drain end which generates reduced hot carrier effect. It is reported that such gate material engineering results in decrease in leakage current, drain induced barrier lowering (DIBL) effect and increase in transconductance and output resistances.

Additionally, graded channel doping has also been explored by researchers for better performance. S. Panigrahi and P. K. Sahu reported that channel engineering technique results in higher drive current compared to conventional DG-MOSFET, but at the same time it also produces lower leakage current [3]. M. A. Abdi et al. derived mathematical expression of surface potential for the graded junction [4]. Their work shows, for constant graded channel, decrease in the value of surface potential occurs with decrease in gate dielectric width and increase in dielectric constant. Bending along the channel due to non-uniform doping causes improved carrier mobility. Additionally, graded channel decreases the roll off effect in shorter channel devices [4].

This study aims at comparative analysis of short channel effects between single gate conventional DG-MOSFET and multi-gate graded channel DG-MOSFET. Simulation of the devices has been performed in Visual-TCAD version 1.8.0-1.

2 Device Specifications

Device 1: Relatively low doping concentration ($10^{16}/\text{cm}^3$) in the channel has been used in device 1 to avoid decrease in carrier mobility and variation in the threshold voltage. 30 nm of channel length has been used for all the devices. SiO_2 layer with thickness of 10 nm has been used as dielectric layer for all cases. Au has been used as gate electrode for the device 1.

Device 2: Doping variation has been used in the device 2. Channel has been divided in three equal regions with each length of 10 nm. Lower doping level ($10^{15}/\text{cm}^3$) in source end, medium doping level $5 \times 10^{15}/\text{cm}^3$ in the middle and highest doping level in the drain end ($10^{16}/\text{cm}^3$) have been used.

Device 3: Device 3 has structure similar to device 1 but with Al as gate electrode. 30 nm channel length with dielectric (SiO_2) thickness of 10 nm has been used in this device. Same substrate doping, as used in device 1, with the value $10^{16}/\text{cm}^3$ has been used in this device. 4.2 eV of work function for Al has been considered in simulation.

Device 4: In device 4 multi-gate has been introduced along with multi-channel structure. Gate materials have been chosen with decreasing work function (WF) as we move from source to drain end. Au, Ag and Al with WF 5.10 eV, 4.5 eV and 4.2 eV have been chosen for this device. Channel doping profile has been used in similar way of device 2, i.e. $10^{15}/\text{cm}^3$, $5 \times 10^{15}/\text{cm}^3$ and $10^{16}/\text{cm}^3$ doping concentration has been used for each 10 nm of channel length in increasing manner from source to drain end (Fig. 1).

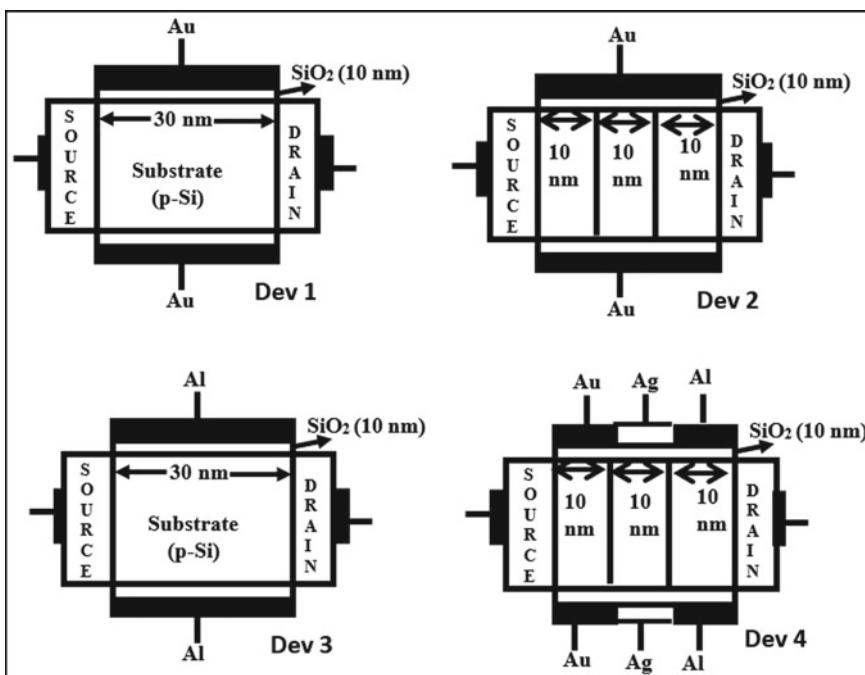
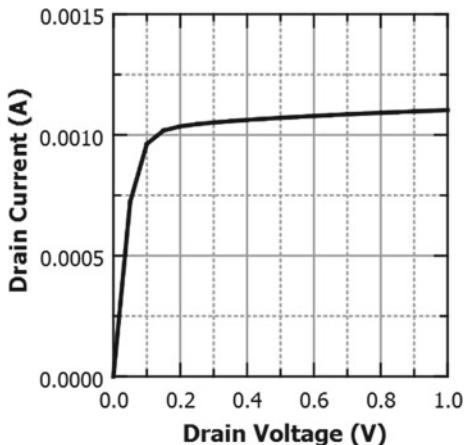


Fig. 1 Schematic diagram explaining structure of the devices under study

Fig. 2 I_D - V_{DS} characteristics for the device 1



3 Results and Discussion

3.1 Effect of Graded Channel

Figure 2 shows I_D - V_{DS} characteristics for device 1. Drain voltage is varied from 0 to 1 V with constant gate voltage at the value of 1.5 V.

I_D - V_{GS} characteristics for device 1 and device 2 are shown in Fig. 2. Gate voltage was varied from 0 to 1 V. Drain to source voltage was kept constant at the value of 1.5 V.

I_D - V_{GS} characteristics for device 1 and device 2 are shown in Fig. 3a. Gate voltage was varied from 0 to 1 V. Drain to source voltage was kept constant at the value of 1.5 V. The figure shows higher ON current (I_{ON}) for Dev. 2. As mentioned above, variation in doping concentration has been used along the channel in device 2. Channel has been divided in three regions of length 10 nm each and doping concentration has been increased from source to drain end. The approximate band diagram for device 1 and device 2, in equilibrium, is shown in Fig. 3b. Lower doping concentration in the source end results in lower conduction band offset. On the other hand, higher doping concentration (i.e. higher band offset) in drain end controls the roll off, a severe challenge in short channel devices.

3.2 Effect of Gate Electrode

Figure 4 shows the I_D - V_{GS} characteristics for Au and Al gate electrode.

The expression of threshold voltage is expressed as

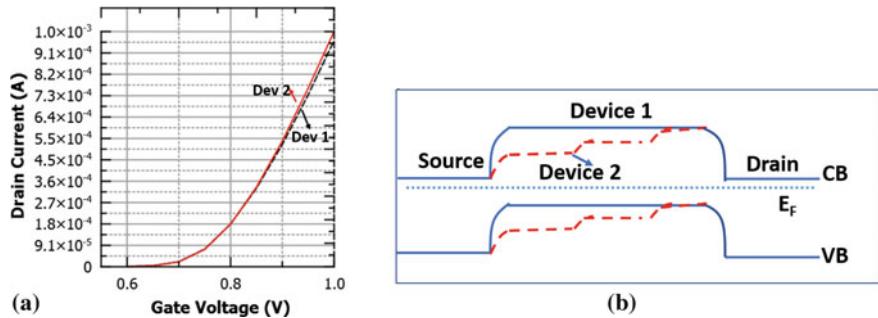
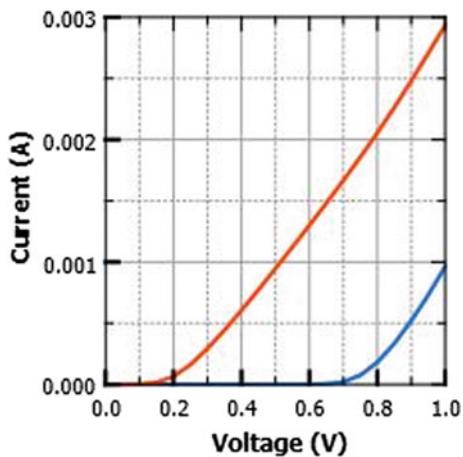


Fig. 3 **a** Input characteristics for device 1 and device 2; **b** approximate band diagram (in equilibrium) along channel length for device 1 (blue) and device 2 (red)

Fig. 4 I_D-V_{GS} characteristics of MOSFET with Al and Au gate



$$V_T = V_{FB} + 2\emptyset_B + \frac{\sqrt{4\epsilon_s q N_a \emptyset_B}}{C_{OX}} \quad (1)$$

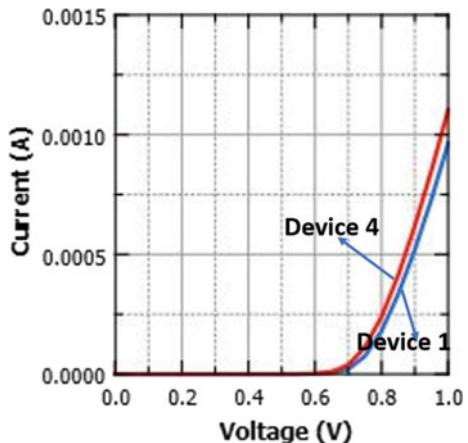
where

- V_T threshold voltage
- V_{FB} flat band voltage
- \emptyset_B bulk potential
- N_a substrate doping concentration
- C_{OX} oxide capacitance

Considering the expression of V_T (Eq. 1), it is evident that gate electrode with lower work function produces lower V_T . Al and Au has work function of 4.2 and 5.1 eV, respectively.

Although metal gate with lower work function produces lower V_T , it increases DIBL effect for smaller devices. Dual gate and triple gate MOSFET can provide

Fig. 5 Input characteristics of device 4 and device 1



better solution to reduce DIBL affect. In this work triple gate MOSFET with graded channel has been studied to obtain improved performance for device with lower dimensions.

3.3 Effect of Multi-gate

Figure 5 shows the input characteristics of device 1 and device 4. Device 4 shows lower V_T compared to device 1.

It should be noted here that, device 4 is with multi-gate structure (Au–Ag–Al) and V_T is with a value closer to device 1 and far away from the value of V_T of device 3 (Fig. 5). As carrier is injected from source end, the value of V_T is predominantly determined by gate electrode at source end.

3.4 Analysis of DIBL

To measure DIBL, I_D - V_{GS} characteristics is measured at two different constant drain voltages. DIBL is measured following the formula [5].

$$\text{DIBL} = \frac{\Delta V_T}{\Delta V_{DS}} \quad (2)$$

In this experiment, input characteristics for device 1 and device 4 has been measured for drain voltage 1.5 and 3 V.

Figure 6 shows the input characteristics for device 1 and device 3 for gate voltage sweep from 0 to 1 V.

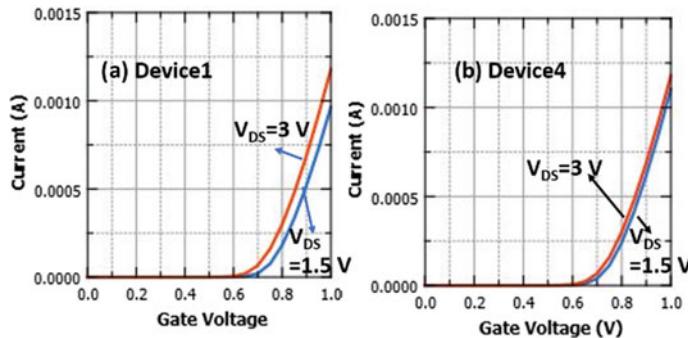


Fig. 6 Input characteristics at $V_{DS} = 1.5$ V and 3 V for **a** device 1 and **b** device 4

Figure 6 clearly indicates that change in threshold voltage, i.e. DIBL effect is more prominent for single gate device, compared to triple gate device. Value of DIBL for device 1 and device 4 is obtained as 0.46667 and 0.1333, respectively.

4 Conclusion

A comparative performance analysis between conventional DG-MOSFET and graded-channel-multi-gate DG-MOSFET has been presented in this study. Results indicate a minor change in V_T but a significant improvement in DIBL can be obtained with the use of graded channel and triple gate in the device structure.

References

1. Mohapatra SK, Pradhan KP, Sahu PK (2014) Impact of dual material gate and lateral asymmetric channel in GS-DG-MOSFET. *Int J Nano Biomater* 5(4):196–205
2. Ramesh R (2017) Influence of gate and channel engineering on multigate MOSFETs-a review. *Microelectron J* 66:136–154
3. Panigrahy S, Sahu PK (2013) Performance enhancement and reduction of short channel effects of nano-MOSFET by using graded channel engineering. In: 2013 international conference on circuits, power and computing technologies (ICCPCT). IEEE, pp 787–792
4. Abdi MA, Djeffal F, Meguellati M, Arar D (2009) Two-dimensional analytical threshold voltage model for nanoscale Graded Channel Gate Stack DG MOSFETs. In: 2009 16th IEEE international conference on electronics, circuits and systems-(ICECS 2009). IEEE, pp 892–895
5. Mendiratta N, Tripathi SL (2020) A review on performance comparison of advanced MOSFET structures below 45 nm technology node. *J Semicond* 41(6):061401

Performance Enhancement and Scheduling in Communication Networks—A Review into Various Approaches



Priya Kumari and Nitin Jain

Abstract Optimizing the communication network's performance under diverse service quality constraints to match the briskly expanding claims of wireless/mobile applications is the vital goal of imminent wireless networks. A conspicuous way to improve the network performance is through embodying several scheduling mechanisms. Though various scheduling schemes exist, improved schemes are still needed for performance breakthroughs. Therefore, this article provides intense research on scheduling and performance optimization of communication systems. It outlines the prime scope of scheduling resources and enhancing diverse performance measures for strengthening and facilitating wireless network performance. It investigates the existing studies on resource allotment, scheduling and performance enhancement of communication networks. This review work illuminates some vital performance metrics involved in performance upgradation. The paper finally presents the paramount research challenges explicitly involved in the performance betterment of communication networks for introducing and implementing optimal schemes and encouraging vast research in this direction.

Keywords Communication networks · Performance enhancement · Resource scheduling · Scheduling methods

1 Introduction

The swift advancement in communication platforms and their technologies have provoked the nascence of neoteric communication network structures like data station networks, millimeter-wave networks, ultra-dense diversified networks and cognitive

P. Kumari (✉) · N. Jain
Babu Banarsi Das University, Uttar Pradesh, Lucknow, India
e-mail: priya.sinha942@gmail.com

N. Jain
e-mail: hod.ec@bbdu.ac.in

networks. Every network's performance requisites and features may vary dynamically [1–7]. Furthermore, the augmenting network capacity allows an assortment of new applications and services like autonomous driving [8], augmented reality [9], edge computing [10] and online gaming [11], entailing highly stringent requisites on the communicating network. These applications communicate via various networks. Recent evolution in mobile computation and wireless systems strongly reveals that dynamic computers or machines and their communication links are integral to looming inter-networks [12]. Commonly, communication through wireless links is chiefly characterized by extreme bit-error rates, temporary disconnections, high latencies and limited bandwidth, which should be handled using network applications and protocols. Many existing communication networks often encounter hurdles in reaching the anticipated performance goals due to delay, throughput, latency, limited resources, unavailability of suitable channels, limited bandwidth, traffic congestion, unclear transmission/reception paths, etc. The congestion control, delay in resource allotment link disconnectivity or poor coverage directly or indirectly affect the network's bandwidth, throughput, latency and efficiency.

Moreover, these constraints hinder the operation and even degrade the entire network performance by lowering network throughput efficiency and augmenting latency packet transmission/reception delay. However, network performance optimization can be achieved by maximizing efficiency throughput and minimizing latency and delay. This is possible through appropriate congestion control, link connectivity monitoring and timely allocation of desired network resources, which can reduce the critical issues encountered in communication networks to a broad extent.

Scheduling techniques play an imperative role in the timely allocation of resources for enabling communication networks to perform their operations without delay, thereby aiding in avoiding traffic congestion and improving communication networks' operation. Many works adopt diverse scheduling strategies for effectively prioritizing and allocating resources to achieve expected network operation and strengthen performance. Scheduling deals with allotting constrained resources, particularly competing operations over time. It aids in satisfying real-time requisites of communication systems by properly assigning tasks to end networks and notifications to desired communication links without violating system resource constraints and reliance relations. In many instances, specifically in highly complex and vast networks, resource scheduling becomes tedious owing to the involvement of diverse components. Optimized scheduling strategies in such cases greatly help resolve resource scheduling impediments and enhance performance. Some studies also exploit congestion control techniques for facilitating smooth communication and task execution [13]. However, these techniques sometimes do not fit well in current or future complex and extremely dynamic networks because of innumerable and diversifying factors influencing the comprehensive network performance. Recently, machine learning (ML), deep reinforcement learning (DRL) and various other approaches have been exploited for realizing and predicting diverse network parameters for further enriching the performance [14–16]. ML helps to learn without being explicitly programmed to perform the desired action. It tries to construct algorithms

and models that can learn to make decisions directly from data without following pre-defined rules. DRL is typically a sub-field of ML that combines deep learning and reinforcement learning schemes for optimal decision-making. Thus, instead of utilizing the scheduling schemes alone, combining them with advanced technologies like ML, DRL, etc., can lead to optimized decision-making concerning resource scheduling and network performance enhancement. Therefore, this review aims to investigate the performances of ML, DRL-based scheduling schemes and general scheduling techniques employed by researchers for the performance amelioration of communication networks.

1.1 Contributions of the Survey

The paramount contributions of this study include:

- To explore distinct performance enhancement techniques in communication networks.
- To investigate several scheduling methods in communication networks.
- To analyze different performance measures employed in existing works.
- To determine the prevailing challenges in performance enhancement and scheduling.

1.2 Paper Organization

The rest of this paper is structured as follows: Section 2 explores scheduling methods for performance augmentation of communication networks. Section 3 gives a clear comparative study of reviewed works. Section 4 highlights the prevailing research hurdles, and Sect. 4 concludes the study.

2 Literature Survey

2.1 General Scheduling Methods

For upgrading the communication network's performance, the work [17] considered massive machine-like devices (MMDs) with resource scheduling and information aggregation. The employed aggregators aggregated data and scheduled resources between MMDs for addressing limited-resource problems. The resource scheduling (RS) was executed through channel-aware RS and random RS techniques. The metrics like the MMD success probability, successful channel usage probability

and average successful MMDs were examined to determine overall MMD communication performance. The channel-aware RS method exhibited superior performance over the random RS method. In [18], the issue of periodic parallel operation scheduling was addressed. Initially, a synchronous operation framework was considered wherein every operation comprised segments with an arbitrary amount of matching threads. Then, an operation decomposition technique was adopted, which decomposed every parallel operation into a genre of sequential operations. Composed tasks were scheduled using partitioned deadline monotonic scheduling (PDMS) and global EDF. It showed how these transformations could be applied, and identical resource expansion bounds could be maintained. In [19], a disseminated link scheduling scheme depending on subsequent interference cancellation was presented for nullifying the interference and coordinating the wireless link transmission in MIMO networks. Results clarified that this scheduling scheme significantly augmented the throughput and strengthened the communication network's performance.

The performance of two scheduling schemes, proportional fair (PF) and round-robin, was evaluated in [20] considering fairness index, delay and throughput metrics. Simulation outputs demonstrated the superiority of the round-robin scheme over the PF scheduling method. The round-robin scheme exhibited improved performance through achieving 18.29% delay and 3.65% throughput. The PF achieved a 0.994 average fairness index, while the round-robin achieved a 0.995 fairness index. In [21], three distinct scheduling strategies were compared: round-robin, PF scheduling and random scheduling. Results revealed that PF scheduling outperformed the random and round-robin methods in federated learning convergence rates. In [22], an optimum fair scheduling (OFS) approach was provided for throughput optimization. The OFS scheme's performance was evaluated regarding fairness and total throughput. The approach displayed better performance in terms of throughput optimization. In [23], a combined link adjustment and scheduling technique was presented for 5G minimal-latency communications. It addressed diverse challenges like inter-cell interference, resource scheduling and link adaptation. The presented technique enhanced the 5G system performance and substantially lowered the latency. In [24], partial priority and complete priority scheduling schemes were combined to lower the total makespan and satisfy the anticipated deadlines of top-priority applications. The proposed mixed scheduling technique achieved valuable optimization with timing constraints and better performance.

Though these methods performed satisfactorily, they failed to achieve effective decision-making concerning scheduling resources and performance enhancement. Hence, to address this shortcoming, researchers have integrated advanced techniques like ML/RL/DRL with scheduling methods for optimizing the scheduling performances.

2.2 *ML/DRL-Based Scheduling*

An ML-based scheduling scheme was exploited in [25] for achieving adaptable transmission time interval (TTI) scheduling. The duration of TTI was selected as per transmission channel and base station conditions. This ML-directed scheduling scheme performed resource scheduling and allotment according to distinct TTIs. Simulation outputs clarified that the presented ML-directed scheduler effectively decreased the packet loss rate delay and improved the network performance. In [26], an ML-directed scheduling solution was proposed for ameliorating live video streaming quality. It aimed at enhancing users' quality of experience and QoS provisioning. This solution specifically targeted extremely challenging scenarios involving high-bitrate live video streaming. Performance evaluation revealed that the presented ML scheduler outperformed diverse competing schedulers regarding throughput, packet loss and delay. In [27], a DRL-based scheduler was proposed for allotting resources in edge computing systems. This scheduler adopted deep Q-learning technology for developing an adaptive strategy for assigning computational resources. Evaluation of scheduling performance indicated that the presented scheduler performed better than random and equal scheduling benchmarks.

In [28], a DRL-directed scheduling technique capable of dynamically adapting to traffic changes and several reward functions was described for optimally scheduling network traffic. It offered a valuable footprint toward designing self-driving, autonomous networks capable of managing traffic using past data. In [29], RL was integrated with TCP to enhance the communication system's performance. Unlike pre-set rule-directed TCP, the proposed RL-TCP scheme exploited reinforcement signals for learning congestion control conditions from experience without requiring former knowledge about network dynamics. The learning agent applied a generalization-dependent Kanerva coding technique for search size and training period reduction. The proposed RL-TCP scheme exhibited higher throughput (about 59.5%) and minimal transmission latency.

ML/RL/DRL-directed scheduling schemes are found to perform better compared to general scheduling techniques. Therefore, these schemes are highly recommended for networks' performance enhancement.

3 Performance Metrics

The prominent performance measures affecting the communication network performance include jitter, throughput, fairness, latency, convergence speed, bandwidth usage, packet loss, retransmission and responsiveness (Fig. 1).

- **Jitter:** It indicates fluctuation in the delay of packets over a network connection. It occurs because of route changes, interference and network congestion.

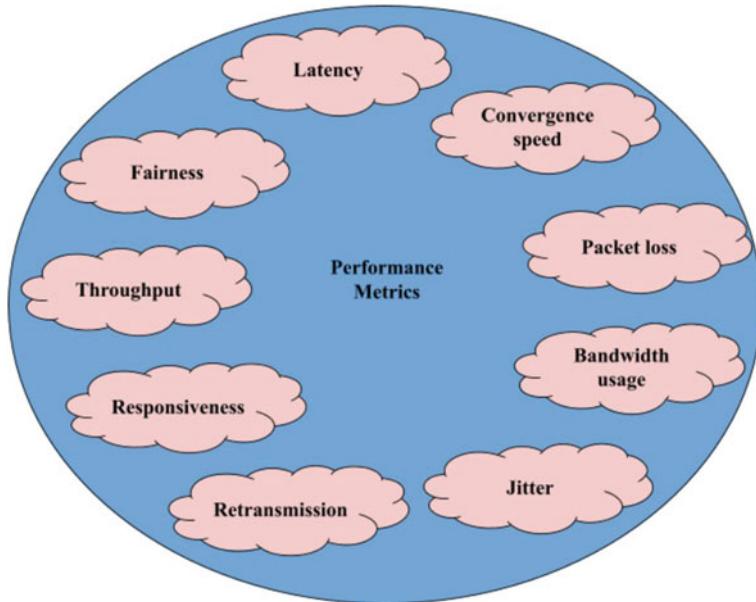


Fig. 1 Distinct performance metrics

- **Throughput:** It indicates the network's original information transmission rate. It is typically the rate of information delivered successfully over a channel. While network bandwidth estimates the theoretical extremity of information transfer, throughput determines how much information is normally sent. Particularly, throughput estimates the successfully transmitted percentage of information packets.
- **Fairness:** The fairness metric in the network determines whether applications or users are getting a fair/impartial share of network/system resources. The unfair resource assignment may cause resource wastage, starvation and redundant allocation. Therefore, the fairness metric is equally vital in smooth network operation.
- **Latency:** Latency signifies the time taken for certain information to reach the desired destination in the network. When estimating latency, odd spikes or consistent delays are prime indications of a serious performance issue occurring due to numerous reasons. From a user's viewpoint, most delays are generally undetectable but can create a large impact using unified communication networks like Skype, Zoom, Microsoft Teams, etc.
- **Convergence speed:** Generally, flow start/end signifies the two predominant events of the network. Slow response affects the network resource usage and user flow realization time. Convergence estimates how swift competing flows congregate to their impartial distribution of bandwidth.

- **Bandwidth usage:** Basically, bandwidth signifies the maximum possible information transmission rate. For reaching maximum network operations, a huge bandwidth is desired. However, most of the time, the bandwidth available is insufficient. It can degrade the network/system performance.
- **Packet loss** describes data packets traveling through a network failing to reach their destination. It occurs because of the errors in information transmission.
- **Retransmission:** Retransmission rate is useful in determining network congestion. By analyzing the retransmission delay or time utilized for the retransmission of dropped data packets, the time used for network recovery from packet loss can be identified.
- **Responsiveness:** It signifies the speed required to reach equilibrium. In common words, responsiveness indicates how swiftly a connection responds to augmented congestion by shrinking its window dimension.

4 Comparative Analysis

The different scheduling methods, including random RS, PF scheduling, round-robin, priority-based, ML/RL/DRL-directed scheduling and others reviewed in Table 1 are illustrated in Fig. 2.

From Fig. 2, it can be noticed that ML/RL/DRL-directed scheduling schemes are extensively adopted in most of the existing works compared to the rest of the scheduling schemes.

5 Open Issues

Despite the noteworthy contributions of existing performance enhancement and scheduling schemes, certain challenges are unresolved due to the dynamicity of extant resources and network conditions. As network resources and requisites vary extensively with location, time and task requirement, designing a generalized resource scheduling approach is tough. Simple conditions signifying the availability, unavailability, limited and sufficient status of resources are no longer supportive for resource characterization. Moreover, constant estimation of resource and network status for allocating dynamic or limited resources between competing entities and upgrading network performance is tedious because of network dynamicity, time and deployment region constraints. Many approaches fail to reach the anticipated network performance due to specific network requisites and unadaptable system conditions. Several times, variations of network conditions like congestion control, queue length and traffic model affect the scheduling process and network operation, causing performance degradation. All these constraints appear as barriers to the performance optimization of communication networks. Therefore, optimized decision-making

Table 1 Comparison of performance enhancement and scheduling techniques employed in the literature

Ref.	Scheduling techniques	Task performed	Merits	Demerits	Future scope
[17]	Channel-aware RS, random RS	RS	Improved overall MMD communication performance	Additional resources/ channels could deteriorate overall MMD communication performance at the information aggregation phase	Optimized scheduling methods must be adopted
[18]	PDMS, global EDF	RS	Achieved resource expansion bounds of 4 and 5	Resource contention, pre-emption penalties and cache effects were not inspected	Analyzation of system performance without conversion to synchronous framework
[19]	Disseminated link scheduling	Interference cancellation	Augmented the throughput	Only the throughput metric was considered for evaluation	Delay, latency and other metrics must be considered
[20]	Round-robin, PF scheduling	Performance enhancement	Employed schedulers achieved a reasonable fairness index	Latency and convergence speed parameters were not considered	Inspection of additional parameters affecting network performance is needed
[23]	Combined link adjustment scheduling	RS and latency minimization	Lowered the latency	Concentrated only on latency minimization	Communication network's performance assessment considering additional metrics is needed
[24]	Partial priority scheduling, complete priority scheduling	Application scheduling	Reduced the total system's makespan and optimized the scheduling performance	Only the timing constraint was considered	Other system constraints influencing scheduling performance must be included

(continued)

Table 1 (continued)

Ref.	Scheduling techniques	Task performed	Merits	Demerits	Future scope
[28]	DRL-directed scheduling	Scheduling of network traffic	Outperformed the heuristic schedulers in traffic scheduling	-	Further performance enhancement is needed, including additional parameters influencing traffic performance
[29]	RL with TCP	Congestion control	Achieved large throughput and minimal latency	Limited performance evaluation	Latency and other network factors must be examined
[30]	Feedforward neural network (FFNN)	Throughput optimization	Reduced latency and enhanced throughput	Failed to attain single cycle repetition interval for tiny FFNN topologies	Additional parameters, like delay, congestion control, etc., should be evaluated for performance enhancement
[31]	DRL	Throughput optimization and prioritized resource allotment	Enhanced whole system performance	Energy conversion efficacy, signal power, concurrent transmission, jamming attack and channel conflicts were not considered	Performance optimization considering additional measures is needed
[32]	DRL	RS	Outperformed greedy and random scheduling schemes	Throughput, traffic congestion and delay metrics were not considered	Other metrics affecting the RS performance should be inspected
[33]	DRL	Congestion control	Achieved large throughput without compromising fairness	Huge state space enhanced the complexity	Techniques for complexity abatement must be discovered
[34]	TCP-RL	Performance enhancement of TCP	Achieved large throughput	The performance of TCP-RL was untested in extremely dynamic network scenarios	The performance of TCP-RL should be tested in complex and extremely dynamic networks

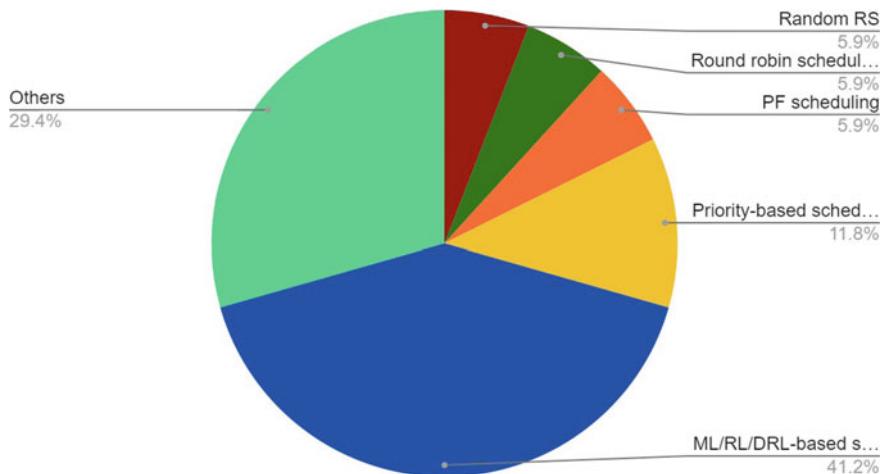


Fig. 2 Reviewed scheduling techniques

schemes integrated with scheduling techniques must be developed to achieve superior performance.

6 Conclusion

This survey article presented a descriptive study on performance optimization and scheduling in communication systems. The prime significance of scheduling in communication networks was explained. The pivotal performance metrics influencing communication network performance were elaborated. The formerly exploited performance optimization schemes and distinct scheduling mechanisms were reviewed. The tasks performed, merits, shortcomings and future directions of these reviewed techniques were tabulated. The prevalent and leading technical hindrances in performance upgradation and scheduling were presented to enlighten the researchers regarding the necessity for discovering and implementing better scheduling and performance modernization schemes for looming communication networks.

References

1. Xiao M, Mumtaz S, Huang Y, Dai L, Li Y, Matthaiou M, Karagiannidis GK, Björnson E, Yang K, Lin-I C, Ghosh A (2017) Millimeter wave communications for future mobile networks. IEEE J Sel Areas Commun 35(9):1909–1935. <https://doi.org/10.1109/JSAC.2017.2719924>

2. Shokri-Ghadikolaei H, Fischione C, Fodor G, Popovski P, Zorzi M (2015) Millimeter wave cellular networks: a MAC layer perspective. *IEEE Trans Commun* 63(10):3437–3458. <https://doi.org/10.1109/TCOMM.2015.2456093>
3. Rani P, Rohit S (2022) An experimental study of IEEE 802.11 n devices for vehicular networks with various propagation loss models. In: International conference on signal processing and integrated networks. Springer Nature Singapore, Singapore
4. Qiao J, Shen XS, Mark JW, Shen Q, He Y, Lei L (2015) Enabling device-to-device communications in millimeter-wave 5G cellular networks. *IEEE Commun Mag* 53(1):209–215. <https://doi.org/10.1109/MCOM.2015.7010536>
5. Fortuna C, Mohorcic M (2009) Trends in the development of communication networks: cognitive networks. *Comput Netw* 53(9):1354–1376. <https://doi.org/10.1016/j.comnet.2009.01.002>
6. Jeon SW, Devroye N, Vu M, Chung SY, Tarokh V (2011) Cognitive networks achieve throughput scaling of a homogeneous network. *IEEE Transact Inform Theory* 57(8):5103–5115. <https://doi.org/10.1109/WIOP.T.2009.5291610>
7. Kumar N, Rani P, Kumar V, Verma PK, Koundal D (2023) TEEECH:three-tier extended energy efficient clustering hierarchy protocol for heterogeneous wireless sensor network. *Expert Syst Appl* 216:119448
8. Levinson J, Askeland J, Becker J, Dolson J, Held D, Kammel S, Kolter JZ, Langer D, Pink O, Pratt V, Sokolsky M, Stanek G, Stavens D, Teichman A, Werling M, Thrun S (2011) Towards fully autonomous driving: Systems and algorithms. In: 2011 IEEE intelligent vehicles symposium (IV), IEEE, pp 163–168. <https://doi.org/10.1109/IVS.2011.5940562>
9. Carmignani J, Furth B, Anisetti M, Ceravolo P, Damiani E, Ivkovic M (2011) Augmented reality technologies, systems and applications. *Multimedia Tools Appl* 51(1):341–377. <https://doi.org/10.1007/s11042-010-0660-6>
10. Satyanarayanan M (2017) The emergence of edge computing. *Computer* 50(1):30–39. <https://doi.org/10.1109/MC.2017.9>
11. Cai W, Leung VC, Chen M (2013) Next generation mobile cloud gaming. In: 2013 IEEE seventh international symposium on service-oriented system engineering, IEEE, pp 551–560. <https://doi.org/10.1109/SOSE.2013.30>
12. Hussain N, Rani P, Kumar N, Chaudhary MG (2022) A deep comprehensive research architecture, characteristics, challenges, issues, and benefits of routing protocol for vehicular ad-hoc networks. *Int J Distrib Syst Technol (IJDST)* 13(8):1–23
13. Zhang T, Mao S (2020) Machine learning for end-to-end congestion control. *IEEE Commun Mag* 58(6):52–57. <https://doi.org/10.1109/MCOM.001.1900509>
14. Alqerm A (2018) Novel machine learning-based techniques for efficient resource allocation in next generation wireless networks (Doctoral dissertation)
15. Zhang C, Patras P, Haddadi H (2019) Deep learning in mobile and wireless networking: a survey. *IEEE Commun Surv Tut* 21(3):2224–2287. <https://doi.org/10.1109/COMST.2019.2904897>
16. Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. *Comput Electric Eng* 105:108543
17. Kumar N, Rani P, Kumar V, Athawale SV, Koundal D (2022) THWSN: Enhanced energy-efficient clustering approach for three-tier heterogeneous wireless sensor networks. *IEEE Sens J* 22(20):20053–20062
18. Saifullah J, Li K, Agrawal C, Lu C (2013) Gill, “Multi-core real-time scheduling for generalized parallel task models.” *Real-Time Syst* 49(4):404–435. <https://doi.org/10.1109/RTSS.2011.27>
19. Wu J, Lin D, Li G, Liu Y, Yin Y (2019) Distributed link scheduling algorithm based on successive interference cancellation in MIMO wireless networks. *Wireless Commun Mobile Comput* 2019. <https://doi.org/10.1155/2019/9083282>
20. Sanyoto N, Perdana D, Bissono YG (2019) Performance evaluation of round robin and proportional fair scheduling algorithms on 5G millimeter wave for node density scenarios. *Int J Simul-Syst Sci Technol* 20(2):17.1–17.2. <https://doi.org/10.5013/IJSSST.a.20.02.17>

21. Yang HH, Liu Z, Quek TQ, Poor HV (2019) Scheduling policies for federated learning in wireless networks. *IEEE Transact Commun* 68(1):317–333. <https://arxiv.org/abs/1908.06287>
22. Ojo MO, Giordano S, Adami D, Pagano M (2018) Throughput maximizing and fair scheduling algorithms in industrial internet of things networks. *IEEE Transact Industr Inform* 15(6):3400–3410. <https://doi.org/10.1109/TII.2018.2873974>
23. Pocovi G, Pedersen KI, Mogensen P (2018) Joint link adaptation and scheduling for 5G ultra-reliable low-latency communications. *IEEE Access* 6:28912–28922. <https://doi.org/10.1109/ACCESS.2018.2838585>
24. Xie G, Zeng G, Liu L, Li R, Li K (2016) Mixed real-time scheduling of multiple dags- based applications on heterogeneous multi-core processors. *Microprocess Microsyst* 47:93–103. <https://doi.org/10.1016/j.micpro.2016.04.007>
25. Zhang J, Xu X, Zhang K, Zhang B, Tao X, Zhang P (2019) Machine learning based flexible transmission time interval scheduling for eMBB and uRLLC coexistence scenario. *IEEE Access* 7:65811–65820
26. Comşa S, Muntean GM, Trestian R (2020) An innovative machine-learning-based scheduling solution for improving live UHD video streaming quality in highly dynamic network environments. *IEEE Transact Broadcast* 67(1):212–224
27. Yang T, Hu Y, Gursoy MC, Schmeink A, Mathar R (2018) Deep reinforcement learning based resource allocation in low latency edge computing networks. In: 2018 15th international symposium on wireless communication systems (ISWCS), IEEE, pp 1–5
28. Chinchali S, Hu P, Chu T, Sharma M, Bansal M, Misra R, Pavone M, Katti S (2018) Cellular network traffic scheduling with deep reinforcement learning. In: Thirty- second AAAI conference on artificial intelligence
29. Li W, Zhou F, Chowdhury KR, Meleis W (2018) QTCP: Adaptive congestion control with reinforcement learning. *IEEE Transact Netw Sci Eng* 6(3):445–458. <https://doi.org/10.1109/TNSE.2018.2835758>
30. Novickis R, Justs DJ, Ozols K, Greitāns M (2020) An approach of feedforward neural network throughput-optimized implementation in FPGA. *Electronics* 9(12):2193. <https://doi.org/10.3390/electronics9122193>
31. Yang Z, Feng L, Chang Z, Lu J, Liu R, Kadoch M, Cheriet M (2020) Prioritized uplink resource allocation in smart grid backscatter communication networks via deep reinforcement learning. *Electronics* 9(4):622
32. Atallah R, Assi C, Khabbaz M (2017) Deep reinforcement learning-based scheduling for road-side communication networks. In: 2017 15th international symposium on modeling and optimization in mobile, Ad Hoc, and wireless networks (WiOpt), IEEE, pp 1–8. <https://doi.org/10.23919/WIOPT.2017.7959912>
33. Xu Z, Tang J, Yin C, Wang Y, Xue G (2019) Experience-driven congestion control: When multi-path TCP meets deep reinforcement learning. *IEEE J Select Areas Commun* 37(6):1325–1336. <https://doi.org/10.1109/JSAC.2019.2904358>
34. Nie X, Zhao Y, Li Z, Chen G, Sui K, Zhang J, Ye Z, Pei D (2019) Dynamic TCP initial windows and congestion control schemes through reinforcement learning. *IEEE J Select Areas Commun* 37(6):1231–1247. <https://doi.org/10.1109/JSAC.2019.2904350>

A Survey of Network Protocols for Performance Enhancement in Wireless Sensor Networks



Abhishek Gupta, Devendra Kumar Sharma, and D. N. Sahai

Abstract With the help of present state-of-the-art microelectronic and electrical technologies, smart wireless sensor networks (WSNs) can be realized that are capable of sensing and monitoring diverse type of environmental parameters like air temperature, humidity, pollution, gas level, etc. Measured data provides a feedback that is used to operate different support systems like air conditioners and ventilators for environmental control. High energy efficiency along with high network throughput rates and low latency are critical design issues for a wireless sensor network. Different types of network protocols like node geographic location-based, multihop, data aggregation, and hierarchical clustering protocols have been proposed for performance enhancement of WSNs. This paper presents a detailed review of different routing protocols for WSN along with their merits and demerits. Node geographic location-based routing like Geographic Distance Routing (GEDIR) and hierarchical clustering (HC)-based routing protocol such as Distributed Energy-Efficient Clustering (DEEC) provides and increased network lifetime and network throughput performance along with reduced network latency performance of a WSN.

Keywords Wireless sensor network · Zigbee · Wi-Fi · Bluetooth · Routing protocol · Sink relocation protocol · Duty cycle scheduling · Data aggregation

A. Gupta

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India
e-mail: abhishek_gupta@bsnl.co.in

D. N. Sahai

INMARSAT, ALTTC, Ghaziabad, India

D. K. Sharma (✉)

Department of Electronics and Communication Engineering, ABES Engineering College, Ghaziabad, Uttar Pradesh, India
e-mail: d_k_s1970@yahoo.co.in

1 Introduction

Low cost implementation of tiny sensor nodes has been made possible by the availability of affordable small-sized microcontroller modules, ADC, RAM-ROM memory modules, Li-ion batteries, and a variety of sensors and transducers [1]. Sensor nodes can be constructed to monitor a wide range of environmental characteristics, such as moisture, air temperature, toxin level, noise, stress, or movement tracking [2] using the appropriate sensor modules. Nodes use the built-in radio-frequency transceivers such as Zigbee or Bluetooth to transmit the measured data packets to a distant reception node [3]. For the purpose of monitoring important parameters, such as moisture or air temperature, many wireless sensor nodes are installed in the target surveillance area. Thus, such nodes establish a wireless sensor network and exchanges data with a central base station sink node. In the present scenario, wireless sensor networks are being used in a broad range of applications like precision agriculture [4], e-Health [5], and home automation [6–8], etc.

Figure 1 shows the use of wireless sensor network for Precision Agriculture application. Several sensor nodes are installed in an agricultural farm. These nodes monitor the moisture level of the soil over there. Node data regarding the moisture content of the soil is transmitted to a central base station. Based on the moisture data, a computer controls the irrigation system that irrigates the farm whenever the required moisture level goes below a critical level.

A little battery resource is capable of supporting a wireless sensor node [9]. The two primary factors that contribute to rapid node battery depletion are data transmission and fast sampling data sensing. Committed data transceiver units consume the highest percentage of node power in a node circuit [10]. Nearby nodes often measure comparable information about frequent occurrences. Significant node power can be saved if nodes are able to identify such data redundancy and choose not to transmit it. Network performance is unaffected by redundancy reduction [11]. High data sampling rates also shorten the lifespan of sensor nodes [2]. Uneven node battery

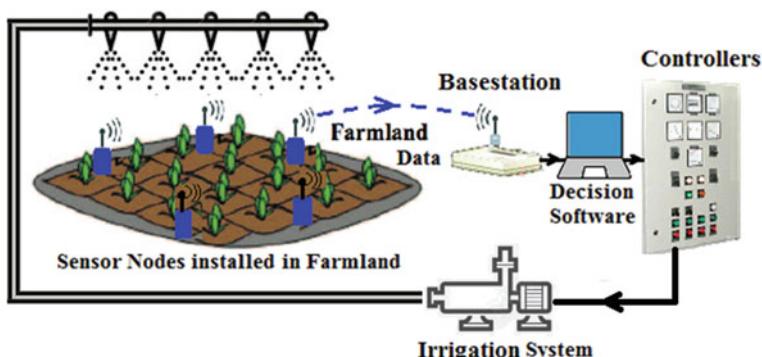
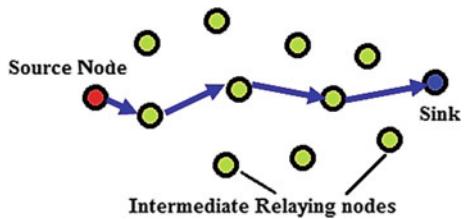


Fig. 1 Wireless sensor network for precision agriculture

Fig. 2 Multihop communication in wireless sensor network



energy consumption across the network is also a concern for multihop networks. As they must transmit substantial data traffic from peripheral nodes to the sink in multihop networks, nodes near the sink rapidly drain their batteries compared to nodes at the source [9]. Figure 2 shows a multihop wireless sensor network. Because there are so many dead nodes, their batteries are typically not replaceable. Additionally, it is a time-consuming and expensive process [4]. Dead nodes shorten the lifespan of the network and raise data loss rates.

Network protocols are the guidelines that network nodes and sink nodes adhere to when performing network operations including sensing, transmitting data to base stations, and processing that data at the stations. Variety of WSN protocols have been proposed by the researchers like dynamic routing protocol [12], hierarchical clustering based routing [13–15], data aggregation [16], sink relocation [11], local data processing [17], etc. Next section defines majorly reported algorithmic level solutions for performance enhancement of WSNs.

2 Different Types of Network Protocols

Different algorithmic level methods for performance enhancement of WSNs are discussed as follows:

A. Cooperative effort by sensor nodes

As per First order radio model [13], a wireless sensor node expands its energy in transmitting a data packet of N bits as follows:

$$E_{Tx} = E_{\text{Transmitter}} \times N + E_{\text{Amplifier}} \times N \times d^2 \quad (1)$$

where $E_{\text{Transmitter}}$ indicates the amount of energy that is required for running node transmitter circuit for transmitting one bit data. $E_{\text{Amplifier}}$ indicates the amount of energy that is required to run node transmitter amplifier for transmitting one bit data. The variable d indicates the transmission distance.

Hence, according to First order radio model, node energy consumption is a positive function of number of bits to be transmitted. In cooperative effort by sensor nodes or local data processing protocols [11, 17], sensor nodes are programmed to execute the sensed data locally. By local processing of the data, nodes remove the redundant

data and transmit the data only when it has changed from previously sensed data or it is beyond normal limits. These techniques efficiently reduce the amount of bits that must be transferred and eliminate the transmission of unnecessary data. Node transmission energy is conserved in this way.

B. Multihop data communication

In the case of multihop communication, instead of sending the data directly to sink, source nodes transmit it to nearby intermediate node [9]. Intermediate node relays the received data to another intermediate node in the direction of sink. Thus being relayed by multiple intermediate nodes, a data packet reaches ultimately to sink node. Figure 2 shows a multihop wireless sensor network.

C. Dynamic routing protocol

Routing protocols decides the optimal multihop data routing path in between leaf and sink node. Dynamic routing protocol [12, 18] chooses a new path for each new transmission round. In this manner data relaying load is distributed evenly among the network nodes and node energy consumption become even. It improves the network lifetime. Figure 3 shows a wireless sensor network with a dynamic routing scheme.

D. Data aggregation

In data aggregation-based routing protocols, intermediary relaying nodes that receive data packets of more than one node, carry out data aggregation on the received data packets [16]. Intermediate relaying nodes thus create an aggregated datum packet from received packets. Data aggregation includes the removal of similar data values, averaging of data values, finding the maxima or minima of data values, etc. This datum packet is sent to the base station sink node. This results in the significant reduction of data traffic and energy consumption in the network. Clustering protocols like LEACH [13], SEP [14], and DEEC [15] use data aggregation. Figure 4 demonstrates the data aggregation process.

E. Sink relocation

In a multihop wireless sensor network, source nodes relay their data to intermediary relaying nodes. Data is further sent to other intermediary nodes via relay nodes. In the end, the data from the source nodes finally reaches the sink node. Relaying nodes that are situated in close proximity to the sink carryout relaying operation at a very fast rate. It quickly depletes their batteries [11]. If the sink node is moved to the new locations in a scheduled manner, this relaying load is divided evenly among network

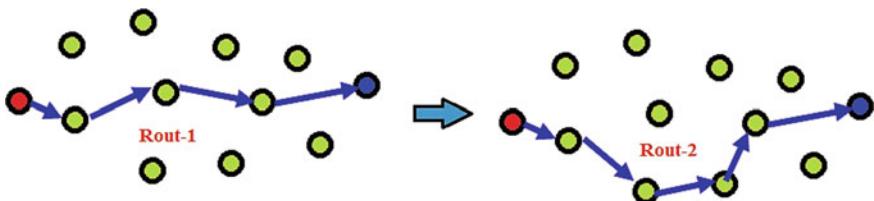
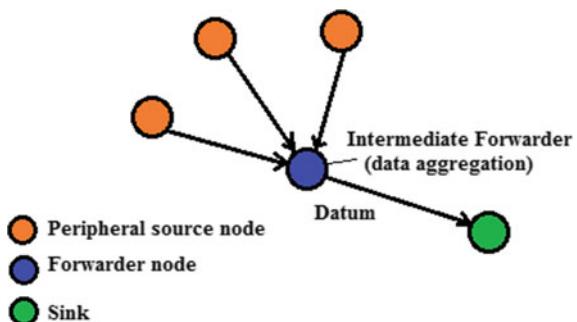


Fig.3 Dynamic routing algorithm

Fig. 4 Data aggregation scheme



nodes, resulting in an equitable distribution of energy consumption. The lifespan of the network and service quality are both positively impacted by sink relocation. Figure 5 shows the concept of sink relocation.

F. Duty cycle scheduling

If data cycle scheduling strategies are used, some of the network's underperforming nodes will be compelled to occasionally enter sleep mode while maintaining network consistency. These techniques increase network longevity and energy efficiency [11, 19]. Table 1 shows the general WSN design problems and their protocol level solutions.

In the areas of low power routing protocol designs and performance models, there is an extensive collection of research-based literature. The literature on routing protocol design and analysis will be reviewed in the following section.

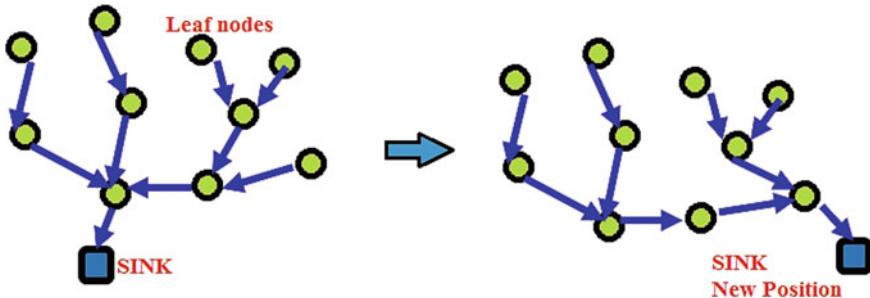


Fig. 5 Sink relocation in WSN [B12]

Table 1 Design issues for WSN

Design issues for WSN	Programming algorithm level solution
Low network lifetime	<ul style="list-style-type: none"> • Local data processing • Data aggregation • Low power routing protocol • Sink relocation protocol • Duty cycle rescheduling
High packet loss rates & high network latencies	<ul style="list-style-type: none"> • Guaranteed delivery routing protocol • Data aggregation • Tuning sampling frequencies fulfilling network lifetime, reliability, and end to end delay constraints

3 Review of Routing Protocols in Wireless Sensor Networks

Routing protocols select the optimum data transmission path from peripheral source nodes to sink node. This section provides a thorough analysis of several routing protocols found in the literature. Based on presented review, the routing protocols can be categorized in to the following categories—(i) Node geographic location-based multihop routing protocols (ii) Data aggregation-based cluster routing protocols. They are surveyed and reviewed as follow:

(i) Node geographic location-based multihop routing protocols

Multihop data routing methods based on node geographic location select data-paths that are based on geographic coordinates of sensor nodes [20]. Each node broadcasts a beacon message with its position coordinates to all other nodes in the network. This process is termed as localization [21, 22].

In case of geographic location-based multihop routing, a source node determines its next relay node based on its own position, the positions of neighbor relay nodes, the distance between nodes, and the progress supplied by the next relaying node in the direction of the base station node. Such routing methods consume less electricity since they need little computing and do not require network addresses. A review of node geographic location-based multihop routing techniques is provided as follows:

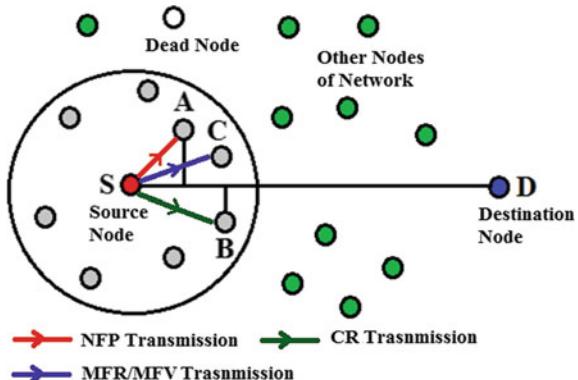
A. Most forward with fixed radius (MFR)

In most forward with fixed radius (MFR) [20] routing protocol, a node transmit its data packet to such relaying node so that its data packet gets maximum nearer to sink. It is called as greedy advancing scheme. The disadvantage of such a routing method is that the transmitted power of a transmitting node remains constant regardless of whether the recipient is distant or close.

B. Most forward with variable radius (MVR)

In case of most forward with variable radius (MVR) [23] routing protocol, nodes can vary their transmission power according to receiver node distance from them. Thus MVR scheme becomes more energy efficient as compared to MFR scheme.

Fig. 6 MFR, MVR, NFP, and CR location-based routing



C. Nearest with forward progress (NFP)

The nearest with forward progress (NFP) protocol allows a node to transmit a data packet to its nearest relaying intermediate node in the sink's direction. Transmission power is adjusted based on transmission distance. In densely populated networks, NFP outperforms MFR and MVR in terms of throughput efficiency [20].

D. Compass or directional routing (CR)

Compass or directional routing (CR) [24] is another form of node geographic location-based routing scheme where a node chooses a relay with the lowest slope to the line connecting the source and sink nodes. Figure 6 demonstrates the MFR, MVR, NFP, and CR schemes.

A dead-end relaying node cannot find the next relaying node in its broadcast range. As a result, a data packets get lost if it is relayed to a dead-end node. It has been demonstrated that the MFR, NFP, MVR, and CR routing protocols are unable to avoid dead-end relaying nodes. They do not give 100% guarantee for successful data transmission.

E. Geographic distance routing (GEDIR)

Geographic Distance Routing (GEDIR) [20, 25] has the ability to leave a dead-end relaying node. In case of GEDIR, if a node transmits a packet to a dead-end node, then the dead-end node sends it back to the previous transmitting node. The previous node then transmits it to the next relaying node. Thus GEDIR can avoid relaying dead ends. Table 2 shows a comparative analysis of MFR, NFP, MVR, CR, and GEDIR routing protocols.

MFR, MVR, NFP, and CR schemes offer sufficient network throughput rate in dense networks. However, in case of sparse network throughput get low due to dead-end relaying nodes. It has been established that the MFR, NFP, MVR, and CR routing protocols are unable to avoid dead-end relaying nodes. They do not give 100%guarantee for successful data transmission. GEDIR can avoid relaying dead ends and provide 95% throughput rate in dense networks. GEDIR is energy-efficient in a sparse network, but as the number of hops increases, their delivery rates decrease.

Table 2 Comparative analysis of different localized geographic routing schemes

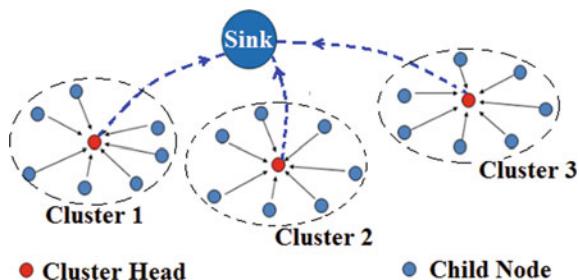
Routing schemes	Dead-end node avoidance guarantee	Network throughput rate		Power consumption	
		Sparse N/W (%)	Dense N/W (%)	Sparse N/W	Dense N/W
MVR	No	5	45	Low	Medium
MFR	No	5	45	Low	High
NFP	No	32	45	Low	Medium
CR	No	50	95	Low	Medium
GEDIR	Yes	50	95	Low	High

Applications requiring the analysis of high resolution data, such as chemical danger detection, forest fire monitoring, and gas leak detection in structures, need for node geographic position-based multihop routing protocols. Each sensor node's data is essential in these applications and must be relayed independently to the sink. Node geographic location-based multihop routing protocols do not suit the applications where a summary of certain physical parameter is required from the location being monitored. In such cases, data aggregation and clustering-based routing protocols proved to be more power efficient.

(ii) Data aggregation-based clustering routing protocols

Clustering-based routing protocols divide the nodes into different clusters. A high energy node from a cluster is allocated as the cluster head (CH) for the cluster. CH node collects packets of other nodes of its cluster and perform data aggregation on received packets. Through data aggregation, multiple data packets are combined into a single datum. CH transmits the datum to the sink node (the base station). Data aggregation may include discarding similar data values, average of data values, maxima or minima of data values, etc. Figure 7 shows the concept of clustering based data routing protocols. LEACH [13], SEP [14], and DEEC [15], and HEED [26] are clustering protocols which use data aggregation. Below is a review of such routing protocols:

A. Leach

Fig. 7 Hierarchically clustered sensor network

Low Energy Adaptive Clustering Hierarchy (LEACH) [13] is a well-known cluster routing scheme. This protocol performs dynamic node clustering and elects new cluster head (CH) for each node-cluster for each new round of transmission cycle. LEACH predicts the requisite ratio of cluster head nodes for each new data transmission round. Suppose n is the count of cluster head nodes that are required for current data transmission round whereas N represents total node count of the network. Then the requisite ratio of cluster head nodes for current data transmission round is $k = n/N$. Further, the LEACH estimates a probabilistic weight function $P_i(t)$ for each cluster node (node-i; $i \in 1, 2, 1 \dots N$) using following formula:

$$P_i(t) = \begin{cases} \frac{k}{N-k*(r \bmod \frac{N}{k})}; C_i(t) = 1 \\ 0; C_i(t) = 0 \end{cases} \quad (2)$$

Here r represents number of transmission round. $C_i(t)$ function is equal to 1 if node-i acted as CH in last $1/k$ number of rounds otherwise $C_i(t) = 0$. Nodes with $P_i(t) > 0$ becomes cluster heads. Then selected cluster heads broadcast an announcement message. After receiving announcement message from selected cluster heads, a non CH node becomes join a CH nearest to it and becomes subordinate node it. In this manner selected cluster heads create clusters. LEACH is a distributed protocol and performs data routing at efficiently.

LEACH does not consider node energy levels for cluster head selection. In this manner nodes having lower energy levels also get selected as CH. Hence, the randomly selected CH nodes lead to shortened lifetime of low energy nodes (if they are selected as cluster head). Hence, LEACH cannot offer even distribution of energy dissipation load across the network nodes. Although LEACH performance can be taken as acceptable for homogenous network with equal node energy levels of all network nodes. But for heterogeneous network (Different energy levels of different network nodes) LEACH performance becomes degraded.

B. SEP

Stable Election Protocol (SEP) [14] is designed to be implemented for heterogeneous WSNs in which the residual energies of the sink and source nodes are dissimilar. The election of cluster head nodes is based on probabilistic weight functions (similar to LEACH) but selection criteria includes the knowledge of node residual energy levels also. In this manner nodes with high energy levels are only chosen for cluster head job. Hence, SEP can offer even distribution of energy dissipation load across the network nodes. SEP provides high data packet throughput as well. SEP protocol requires the knowledge of node residual energy levels in each transmission round. SEP is designed for stationary network nodes only.

C. DEEC

Similar to SEP, the Distributed Energy-Efficient Clustering (DEEC) [15] is also designed to be implemented for heterogeneous WSNs in which the residual energies of the normal and advance nodes are dissimilar. For each node, DEEC assigns a cost value that is equal to the node energy divided by the overall network energy. The

Table 3 Comparative analysis of different clustering-based routing schemes

	LEACH	SEP	DEEC	HEED
Network architecture	Homogeneous	Heterogeneous	Heterogeneous	Heterogeneous
Cluster head selection criteria	Probabilistic weight function based	Probabilistic weight function based	Cost function based	Cost function based
Energy efficiency	Moderate	High	Moderate	Moderate
Cluster head distribution Uniformity among network node	Random	Moderate	Moderate	High
Network lifetime	Low	Moderate	High	Moderate

cluster head (CH) is chosen as the node in a cluster with the greatest cost value. The likelihood of advance nodes being chosen as the CH increases as a result, and they are commonly chosen. It shortens the lifespan of the advance node.

D. HEED

Hybrid Energy-Efficient Distributed clustering (HEED) protocol [26] targets the networks incorporating multi-layer architecture. In HEED, normal nodes create primary level clusters. HEED calculates a cost value for each node of each cluster using a cost function of node energies, node proximity to the base station, and number of immediate neighbors a node has. Highest residual energy node in a cluster which with least distance from sink and highest numbers of immediate neighbors is ideal candidate for cluster head (CH). Such node performs routing with least required energy. HEED selects such node as CH. CH selection is even and results in improved energy dissipation load balancing among network node. Clustering with several layers increases overall energy usage. Table 3 contains a comparative analysis of different clustering.

4 Conclusion

This paper presented a study of a wide range of data routing protocols from the existing literature that have been proposed as solutions to various design difficulties of wireless sensor networks, such as low network lifetime, high packet loss rates, and high network latency. The surveyed protocols include node geographic location-based multihop data routing protocols and data aggregation-based cluster routing protocols. MFR, MVR, NFP, CR, and GEDIR are the examples of node geographic location-based routing protocols. MFR, MVR, NFP, and CR protocols offer sufficient network throughput rate in dense networks. However, in case of sparse networks, their throughput get low due to dead-end relaying nodes. MFR, NFP, MVR, and

CR routing protocols are unable to avoid dead-end relaying nodes. Hence, they do not give 100% guarantee for successful data transmission. GEDIR can avoid relaying dead ends and provide 95% throughput rate in dense networks. GEDIR is energy-efficient in a sparse network, but as the number of hops increases, their delivery rates decrease. LEACH, SEP, DEEC, and HEED are the examples of data aggregation and hierarchical clustering-based routing protocols that are surveyed in details along with their merits and demerits. LEACH protocol does not consider node energy levels for cluster head selection. In this manner, nodes having lower energy levels also get selected as CH. Hence, the randomly selected CH nodes shorten the operating period of low energy nodes resulting in a reduced network lifetime. SEP, DEEC, and HEED consider node residual energy levels for cluster head selection. In this manner nodes with high energy levels are only chosen for cluster head job. DEEC and HEED protocols are designed for heterogeneous sensor networks and offer prolonged network lifetime in comparison of LEACH protocol.

References

1. Imam SA, Choudhary A, Sachan VK (2015) Design issues for wireless sensor networks and smart humidity sensors for precision agriculture: a review. In: 2015 International conference on soft computing techniques and implementations (ICSTI), Faridabad, pp 181–187
2. Silvius CF et al (2015) A low-power wireless sensor for online ambient monitoring. *IEEE Sens J* 15(2):742–749
3. Challoo R et al (2012) An overview and assessment of wireless technologies and co-existence of zigbee, bluetooth and wi-fi devices. *Elsevier Proc Comput Sci* 12:386–391
4. Cheick TK et al (2015) Performance management of IEEE 802.15.4 wireless sensor network for precision agriculture. *IEEE Sens J* 15(10)
5. Wu T, Wu F, Redouté JM, Yuce MR (2017) An autonomous wireless body area network implementation towards IoT connected healthcare applications. *IEEE Access* 5:11413–11422
6. Imam SA, Choudhary A, Zaidi AM, Singh MK, Sachan VK (2017) Cooperative effort based wireless sensor network clustering algorithm for smart home application. In: 2017 2nd IEEE international conference on integrated circuits and microsystems (ICICM), Nanjing, pp 304–308
7. Chakravarthi V et al (2013) Technology for smart home. In: Proceeding of international conference on VLSI, communication, advanced devices, signals & systems and networking (VCASAN-2013), Springer India, pp 7–12
8. Pavithra D et al (2015) IoT based monitoring and control system for home automation. In: Proceeding IEEE global conference on communication technologies (GCCT), pp 169–173, 23–24
9. Akyildiz IF et al (2002) Wireless sensor networks: a survey. *Comput Netw* 38(4):393–422
10. Vibhav Kumar S, Syed Akhtar I, Beg MT (2012) Energy-efficient communication methods in wireless sensor networks: a critical review. *Int J Comput Appl* 39(17):975–8887
11. Wang C-F, Shih J-D, Pan B-H, Wu T-Y (2014) A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks. *IEEE Sens J* 14(6):1932–1943
12. Nikolaos AP et al (2013) Energy-efficient routing protocols in wireless sensor networks: a survey. *IEEE Commun Surv Tut* 15(2):551–591
13. Wendi Rabiner H, Anantha C, Hari B Energy-efficient communication protocol for wireless microsensor networks. In: Proceeding of the 33rd Hawaii international conference on system sciences-(HICSS '00), IEEE Computer Society, Washington, DC, USA, vol 8

14. Georgios S, Ibrahim M, Azer B (2004) SEP: a stable election protocol for clustered heterogeneous wireless sensor networks. In: Proceeding second international workshop on sensor and actor network protocols and applications (SANPA 2004), Boston, MA
15. Qing L, Zhu Q, Wang M (2006) Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Comput Commun* 29(12):2230–2237
16. Frank YSL et al (2006) Multi-sink data aggregation routing and scheduling with dynamic radii in WSNs. *IEEE Commun Lett* 10(10):692–694
17. Kshitij S et al (2011) Intelligent humidity sensor for - wireless sensor network agricultural application. *Int J Wireless Mob Netw* 3(1):118–128
18. Datta S, Stojmenovic I, Wu J (2002) Internal node and shortcut based routing with guaranteed delivery in wireless networks. 2002 Kluwer Academic Publishers. Manufactured in The Netherlands pp 169–178
19. Chiara P et al (2014) ALBA-R: load-balancing geographic routing around connectivity holes in wireless sensor networks. *IEEE Transact Parallel Distribut Syst* 25(3):529–539
20. Hou T-C, Li V (1986) Transmission range control in multihop packet radio networks. *IEEE Trans. on Communications* 34(1):38–44
21. He T, Stankovic JA, Lu C, Abdelzaher T (2003) Speed: a stateless protocol for real-time communication in sensor networks. In: International conference on distributed computing systems (ICDCS), p 46
22. Ivan S, Xu L (2001) Power-aware localized routing in wireless networks. *IEEE Trans Parallel Distribut Syst* 12(11):1122–1133
23. Stojmenovic I, Lin X (2001) Loop-free hybrid single-path/flooding routing algorithms with guaranteed delivery for wireless networks. *IEEE Trans Parallel Distribut Syst* 12(10):1023–1032
24. Kranakis E, Singh H, Urrutia J (1999) Compass routing on geometric networks. In: Proceeding 11th Canadian conference computational geometry, pp 51–54
25. Stojmenovic I, Lin X (1998) A power aware distributed routing in ad hoc wireless networks. Technical Report TR-98-11, Computer Science, SITE, Univ. of Ottawa
26. Younis O, Fahmy S (2004) HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans Mob Comput* 3(4):366–379

Airbnb Price Prediction Using Advanced Regression Techniques and Deployment Using Streamlit



Ayan Sar, Tanupriya Choudhury, Tridha Bajaj, Ketan Kotecha, and Mirtha Silvana Garat de Marin

Abstract This article seeks to anticipate AirBnB prices using advanced regression approaches. Extensive data analysis was done on different databases spanning diverse variables such as location, property type, facility, and user level. The database is constructed utilizing robust approaches such as feature augmentation, outlier reduction, and value loss. A number of complex regression models, such as linear regression, decision tree, random forest, gradient regression, are generated on the pre-developed database. The model is improved through hyperparameter adjustment to increase prediction accuracy. A cross-validation approach was employed to examine the performance and resilience of the model. In addition, a feature significance study was undertaken to discover the most significant elements impacting Airbnb prices. The experimental findings suggest that the improved regression approach delivers greater prediction accuracy than the standard model. The results of this study add to

A. Sar · T. Bajaj

Informatics Cluster, School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand 248007, India

T. Choudhury (✉)

CSE Department, Symbiosis Institute of Technology, Symbiosis International University, Lavale Campus, Pune, Maharashtra 412115, India

e-mail: tanupriya.choudhury@sitpune.edu.in; tanupriya@ddn.upes.ac.in

SoCS, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand 248007, India

K. Kotecha

Symbiosis Center for Applied Artificial Intelligence, Symbiosis Institute of Technology, Symbiosis International University, Pune 411045, India

M. S. G. de Marin

Engineering Research and Innovation Group, Universidad Europea del Atlántico, C/Isabel Torres 21, 39011 Santander, Spain

e-mail: silvana.marin@unic.co.ao

Department of Project Management, Universidade Internacional do Cuanza, Estrada Nacional 250, Bairro Kaluapanda, Cuito-Bié, Angola

K. Kotecha

e-mail: Director@sitpune.edu.in

Airbnb's pricing system and can promote improved decision-making for hosts and visitors searching for competitive pricing.

Keywords Advanced regression · AirBnB · Hyperparameter tuning · Streamlit · Forecasting

1 Introduction

The increasing drift in the rise of Internet platforms like Airbnb has recently changed the way customers and individuals book housing, hotels, and homestays. This peer-to-peer accommodation in the housing had allowed homeowners to rent out their leftover property, starting from the spare rooms to fully furnished homes to the tourists who went on searching for unique and pleasant housing properties in minimum pricing range. With the increase in popularity of Airbnb, and to be more accurate, anticipating rental pricing has become a huge difficulty, especially for owners and visitors. Accurately pricing forecasting and predicting is very important for landlords because it can help them establish competitive prices that can extensively attract guests with more potential while optimizing their income. By properly utilizing the power of machine learning and statistical analysis, a trustworthy model is constructed that can be reliably used to predict the rental price based on different factors connected to the property, location, and different characteristics of the house owner or the landlord. A vast database of Airbnb listings is studied, including information such as house type, number of beds, amenities, area features, and host-related aspects. The regression methods implemented in this study can go beyond typical linear regression models and further into more complex algorithmic solutions such as ensemble methods, support vector regression, random forest regression, and gradient boosting algorithms. The comparison of different strategies and the assessment of their effectiveness results in the discovery of the most efficient strategies for estimating the accurate value of Airbnb. On the other hand, the customers and guests may utilize these estimations to build more accurate and educated judgments and ensure that the housing or homestay can meet their expectations on every basis, like budget and interests.

2 Problem Identification

This research article addresses the difficulty of reliably estimating AirBnB listing pricing using sophisticated regression algorithms. The major difficulty may be separated into the following key aspects:

Price Fluctuations: AirBnB properties vary in price owing to different factors such as location, home type, amenities, and seasonal demand.

Multidimensional Variables: Airbnb listings contain several variables that might affect pricing, including home size, number of bedrooms, amenities, proximity to famous destinations, and more.

Non-linearity and Complex Relationships: Relationships between qualities and enumeration values can be non-linear or straightforward. This challenge involves understanding and employing advanced regression algorithms that can represent such complexity.

Data Quality and Processing: AirBnB data may have missing numbers, outliers, and inconsistencies that may influence the accuracy of the prediction model. Proper data processing techniques should be employed to overcome these challenges and assure the quality and dependability of the database.

Generalization and Scalability: The constructed regression model should not only produce accurate predictions on the basis of training but should also be able to generalize to unseen data.

Evaluation Criteria: In order to correctly analyze the performance of the price forecasting model, it is required to pick suitable evaluation criteria.

3 Literature Review

Recent research has been conducted in this area, and multiple studies have been done on these price prediction models for application in different sectors [1] mainly focuses using Principal Component Analysis (PCA) to enhance the model's accuracy on the home pricing and rent. They found that ridge regression is best for continuous price prediction, and the random forest classification is best for predicting price ranges individually. Analyzes the various machine learning algorithms to predict the rental pricing for shared warehouses using linear regression and random forest algorithms [2]. Mainly examines the factors responsible for the impact on home service prices on the website of Airbnb, utilizing the ordinary least squares and quantile regression analysis [3]. Explored the hotel pricing strategies in Kyiv using regression and time-series analysis [4]. Gives an overall idea of gradient boosting in the field of machine learning [5]. Proposes a dynamic pricing strategy for hotel pricing, considering important factors like demand and sensitivity [6]. Looks upon the factors influencing hotel room pricing in Beirut, Lebanon through the hedonic pricing models and Ordinary Least Squares (OLS) regression analysis [7]. Puts efforts into accurate pricing for Airbnb listings and the importance of data preprocessing using machine learning models like linear regression, decision trees, and random forests [8]. Focuses on variable feature selection and various machine learning approaches like XGBoost and neural networks for price prediction [9]. Analyzes the Airbnb housing prices in the USA, putting into use linear regression, decision trees, and random forest algorithms [10]. Predicts the Airbnb rental pricing using classification techniques and feature selection methods [11]. Studies the Airbnb price prediction in London using various machine learning algorithms, with random forest providing the best results and potential future scope [12].

4 Methodology

Data Collection and Preprocessing: The dataset consisted of Airbnb listings, which included relevant features such as location, home type, amenities, proximity to famous destinations, and price. The handle of missing values cleaned the dataset along with the outliers and inconsistencies. Appropriate data imputation techniques were also applied, along with outlier detection methods. The normalization was also applied with scaling numerical features to ensure the equal contribution of variables to the regression model.

Feature Selection and Engineering: A thorough dataset analysis was conducted to determine the most significant features that can contribute to the influence of pricing. Statistical analysis was also used for the selection of relevant features along with feature engineering.

Model Selection: A detailed experiment was done with the various regression algorithms to find the best-suited one for the problem. Common regression models like Linear Regression, Polynomial Regression, Random Forests, Gradient Boosting, and CatBoost regression were considered for their simplicity and interpretability with robustness.

Model Training and Evaluation: The dataset was split between training and testing, and suitable evaluation criteria were also utilized, such as Mean Squared Error, Root Mean Squared Error, and Mean Absolute Error for the performance assessment.

Hyperparameter Tuning: The hyperparameter tuning was done to optimize its performance. Suitable techniques, such as cross-validation, were also utilized to avoid overfitting and improve the model's ability to generalize unseen data.

Model Deployment and Scalability: After achieving the satisfactory regression model, the model will be deployed for price prediction in a web-based app using Streamlit to portray and test its basic use. It was also ensured that the model was scalable and could handle many AirBnB listings from different cities and regions.

5 Datasets and Algorithms Discussed

This dataset was collected from the Kaggle dataset on AirBnB price prediction and the corresponding competition. The AirBnB dataset consisted of 29 variables and 74,111 observations related to various rental properties. The variables in the dataset included room type, property type, city, latitude, longitude, reviews, score rating, etc.

The selective choice of algorithms for the comparative analysis on price prediction depends on several factors, including dataset characteristics, main problem identified, and the solution proposed. The reasons include:

- (i) **Simplicity and interpretability:** Linear regression and polynomial regression are very simple and easy to implement, making them suitable as baselines. The

level of interpretability provided by them is also helpful in understanding the impact of features individually.

- (ii) **Ensemble Methods:** CatBoost, Gradient Boosting, XGBoost, and Random Forest are all ensemble methods that combine multiple weak learners to make a powerful predictive model.
- (iii) **Robustness:** The dataset used here contains complex relationships between the features and prices. These methods perform the task of non-linear pattern recognition.
- (iv) **Feature importance:** These methods can also be very useful in providing information about the feature importance, which is also valuable for understanding the features playing a significant role in the price prediction of Airbnb listings.

Linear Regression: Linear regression is a basic algorithm used to predict continuous numerical values. It presupposes a linear connection between input features and goal variables. The method determines the best fit that minimizes the difference between the predicted value and the actual value of the target variable.

$$J = \frac{1}{n} \sum_{i=1}^n (\mathbf{B}_0 + \mathbf{B}_1 \cdot \mathbf{x}_i - y_i)^2$$

$$\frac{\partial J}{\partial \mathbf{B}_0} = \frac{2}{n} \sum_{i=1}^n (\mathbf{B}_0 + \mathbf{B}_1 \cdot \mathbf{x}_i - y_i)$$

$$\frac{\partial J}{\partial \mathbf{B}_1} = \frac{2}{n} \sum_{i=1}^n (\mathbf{B}_0 + \mathbf{B}_1 \cdot \mathbf{x}_i - y_i) \cdot \mathbf{x}_i$$

$$J = \frac{1}{n} \sum_{i=1}^n (\mathbf{B}_0 + \mathbf{B}_1 \cdot \mathbf{x}_i - y_i)^2$$

$$\mathbf{B}_0 = \mathbf{B}_0 - \alpha \cdot \frac{2}{n} \sum_{i=1}^n (pred_i - y_i)$$

$$\mathbf{B}_1 = \mathbf{B}_1 - \alpha \cdot \frac{2}{n} \sum_{i=1}^n (pred_i - y_i) \cdot \mathbf{x}_i \quad (1)$$

Equational derivation for linear regression.

The line is determined by assessing the coefficients for each input feature. Linear regression is commonly used due to its simplicity and clarity.

Polynomial Regression: Polynomial regression extends linear regression by incorporating polynomial terms of the input features in the model. It captures non-linear correlations between characteristics and target variables.

$$\begin{aligned}
 \frac{\partial \mathbf{R}}{\partial \mathbf{a}} &= \sum_{i=0}^n \mathbf{a}x_i^4 + \mathbf{b}x_i^3 + \mathbf{c}x_i^2 - y_i x_i^2 = 0 \\
 \frac{\partial \mathbf{R}}{\partial \mathbf{b}} &= \sum_{i=0}^n \mathbf{a}x_i^3 + \mathbf{b}x_i^2 + \mathbf{c}x_i - y_i x_i = 0 \\
 \frac{\partial \mathbf{R}}{\partial \mathbf{c}} &= \sum_{i=0}^n \mathbf{a}x_i^2 + \mathbf{b}x_i + \mathbf{c} - y_i = 0
 \end{aligned} \tag{2}$$

Equation of polynomial regression of minima with partial derivatives set to 0.

By incorporating higher-order variables (e.g., square or cube) of the input characteristics, polynomial regression can better capture the skewed pattern in the data. This flexibility allows for more accurate predictions when the connection between characteristics and target variables is not linear.

CatBoost Regressor: CatBoost is a gradient boosting method that is more efficient at processing categorical features. Categorical characteristics are variables that may take on specified values, such as status or property type, in an Airbnb listing.

$$\text{Objective} = \frac{1}{n} \sum_{i=1}^n L(y_i, F(x_i)) + \frac{1}{2} \lambda \|\theta\|^2 \tag{3}$$

Objective function for CatBoost Regressor.

$$F(\mathbf{x}) = \sum_{k=1}^K f_k(\mathbf{x}) \tag{4}$$

Ensemble function used by decision trees for prediction.

Gradient Boosting Regressor: Gradient boosting is an ensemble learning approach that combines numerous weak prediction models to generate a strong prediction model. In terms of regression, weak models are frequently decision trees.

$$\begin{aligned}
 F_0(\mathbf{x}) &= \arg_{\gamma} \min \sum_{i=1}^n L(y_i, \gamma) \\
 L &= \frac{1}{n} \sum_{i=0}^n (y_i - \gamma_i)^2 \\
 \frac{dL}{d\gamma} &= \frac{2}{2} \left(\sum_{i=0}^n (y_i - \gamma_i) \right) = - \sum_{i=0}^n (y_i - \gamma_i)
 \end{aligned} \tag{5}$$

Equational derivation for proposed Gradient Boosting Regressor.

Ascending gradients generate models in a stepwise approach, where each new model corrects the mistakes of the prior model. The method iterates across the ensemble decision tree, focusing on circumstances where the prior model has underperformed. The final estimate is generated by adding the estimates from the various trees.

XGBoost Regressor: XGBoost is a gradient boosting method noted for its efficiency and performance. Like CatBoost and Gradient Boosting, XGBoost generates an ensemble of weak prediction models.

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t)$$

Equation 6 XGBoost objective function analysis.

For complicated models, it incorporates regularization approaches to avoid overfitting, such as adding a penalty to the loss function. XGBoost can handle missing values and categorical characteristics, making it a strong method for various datasets.

Random Forest Regressor: Random forest is an ensemble learning method that creates several decision trees and combines their predictions. This avoids duplication and gives trustworthy predictions by evaluating a subset of random characteristics in each distribution.

$$RFfi_i = \frac{\sum_j normfi_{ij}}{\sum_{j \in all\ features, k \in all\ trees} normfi_{jk}} \quad (7)$$

Equation for calculating feature importance inside random forest regression model.

6 Proposed System Design

The proposed system design is depicted in the Fig. 1:

7 UML Class Diagram

See Fig. 2.

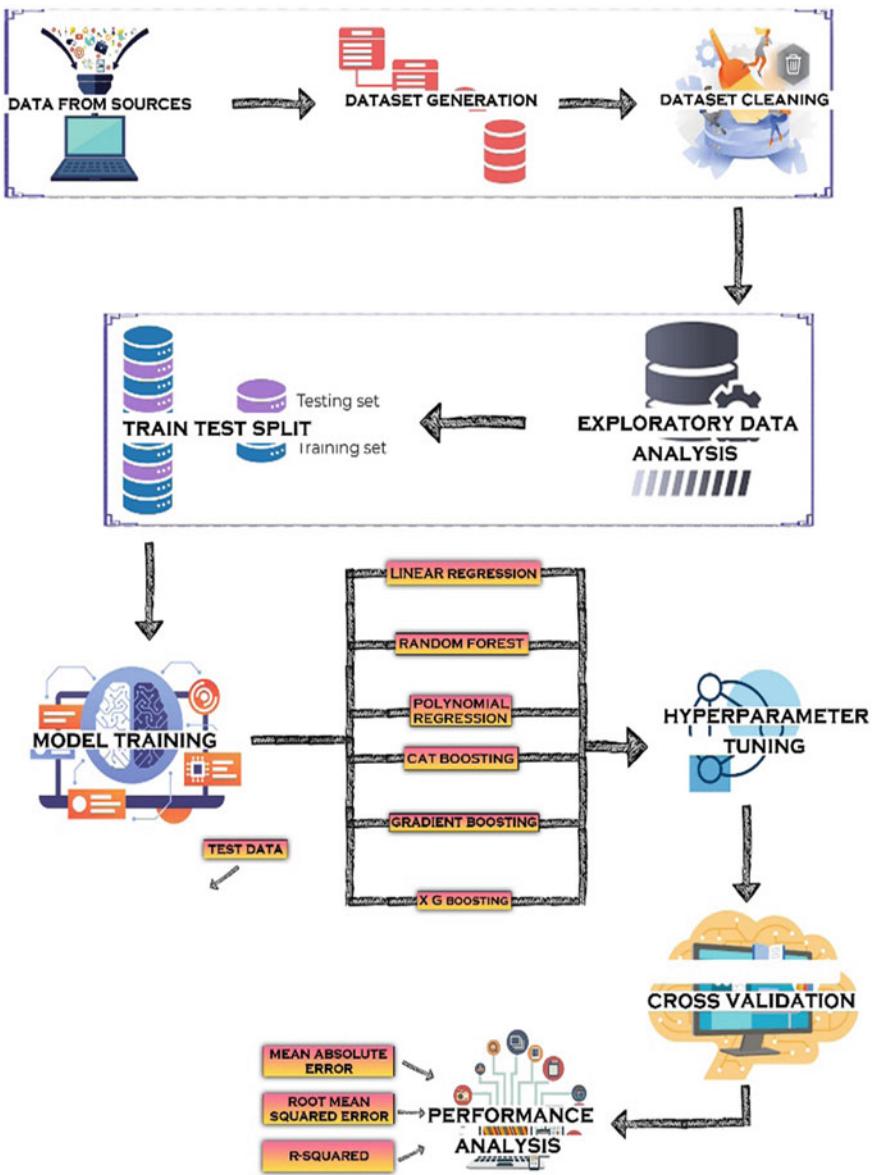


Fig. 1 Graphical abstract for the proposed system design

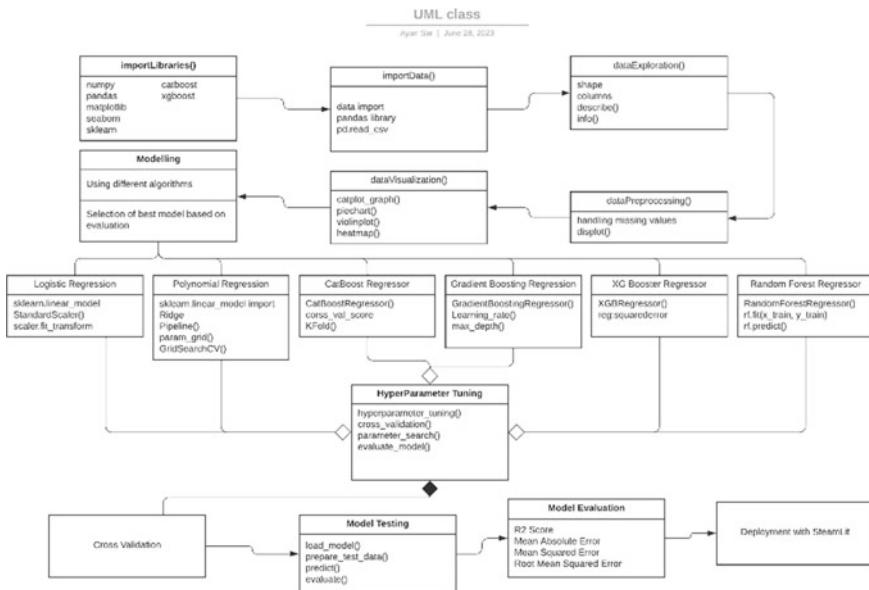


Fig. 2 UML class diagram of the proposed framework

8 Results and Discussion

The system on which the deep neural networks was implemented and the dependencies used for the successful training and implementation of the model:

Processor: 11th Gen Intel(R) Core (TM) i5-11400H @ 2.70GHz 2.69 GHz.

Memory: 24.0 GB (23.7 GB usable).

Disk: NVMe Micron_2450_SSD (512GB).

Network: MediaTek Wi-Fi 6 MT7921 Wireless Lan Card (100Mbps up and down).

GPU 1: Intel® UHD Graphics 6000.

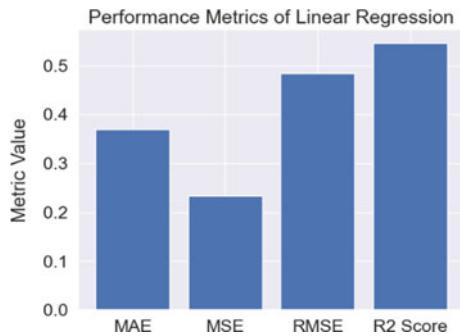
GPU 2: Nvidia GeForce RTX 3050 Ti Laptop GPU.

Dependencies and packages: TensorFlow, Keras, CUDA Toolkit, Matplotlib, cuDNN libraries should be configured as per the GPU available.

In this part, we give an evaluation of the effectiveness of several machine learning models for forecasting Airbnb listing pricing. Models tested include linear regression, polynomial regression, CatBoost regression, gradient boosting regression, XGBoost regression, and random forest regression. Evaluation based on many metrics including mean squared error (MSE), root mean squared error (RMSE), absolute error (MAE), and R-squared (R2) score.

Linear Regression: Linear regression is a simple and extensively used model for predicting continuous variables.

Fig. 3 Performance metrics of linear regression



These findings reveal as in Fig. 3 that the linear regression model captures roughly 54.7% of the variance in the target variable. Although these findings give a baseline for comparison, it is evident that a more complicated model can yield superior performance.

Polynomial Regression: Polynomial regression expands the possibilities of linear regression by incorporating polynomial features. The incorporation of multivariate features as illustrated in Fig. 4 allows the model to capture non-linear interactions between input characteristics and target variables, resulting in enhanced prediction performance. The R2 value of 0.606 implies that about 60.6% of the target variable is explained by the polynomial regression model.

CatBoost Regressor: CatBoost is a gradient boosting technique that automatically organizes category characteristics without the requirement for explicit coding.

These findings suggest that CatBoost successfully as shown in Fig. 5 catches the complex associations in the database and leads to a considerable improvement in prediction accuracy. A strong R2 value suggests that about 72% of the target variable is explained by the CatBoost Regressor.

Gradient Boosting Regressor: Gradient boosting is an ensemble learning strategy that combines numerous weak learners to generate a powerful prediction model. These findings as given in Fig. 6 reveal that gradient boosting is better

Fig. 4 Graphical comparison of the evaluation metrics of polynomial regression

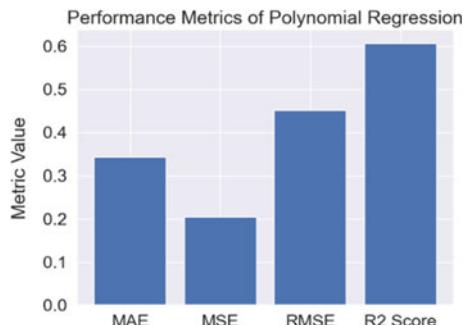


Fig. 5 Graphical comparison of the evaluation metrics of CatBoost regressor

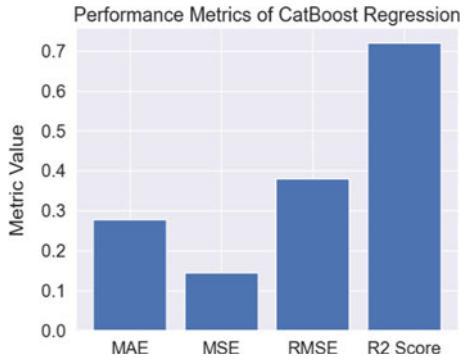
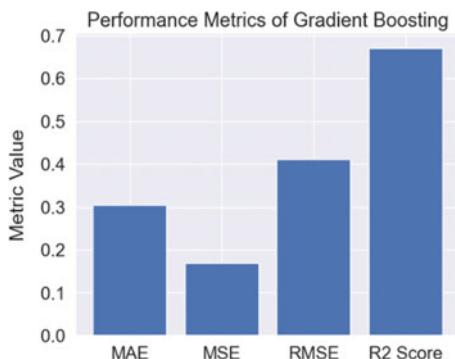


Fig. 6 Graphical comparison of the evaluation metrics of gradient boosting



than linear regression and polynomial regression, but somewhat behind CatBoost. However, the model captures around 67.1% of the variability of the target variable and indicates its capacity to produce relatively accurate predictions.

XGBoost Regressor: XGBoost is a strong gradient boosting technique noted for its speed and performance. The findings are comparable to CatBoost as shown in Fig. 7, proving that XGBoost can efficiently express complicated connections in databases. The R2 score greater than 0.712 suggests that around 71.2% of the target variable is explained by the XGBoost Regressor.

Random Forest Regressor: Random forest is an ensemble learning technique that creates several decision trees and combines their predictions. These findings are consistent as represented in Fig. 8 with the performance of gradient boosting and XGBoost, further confirming the efficiency of the ensemble technique. An R2 value of 0.700 shows that nearly 70% of the target variable is explained by the random forest regressor.

Fig. 7 Graphical comparison of the evaluation metrics of XGBoost regressor

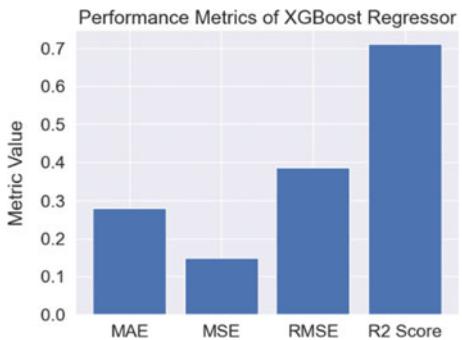


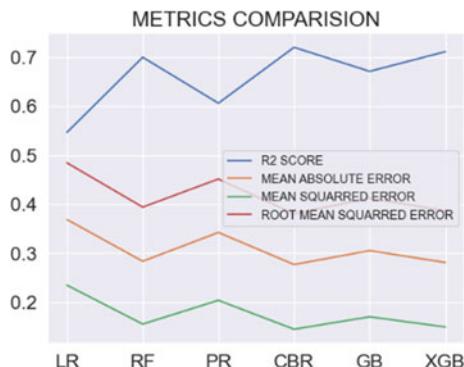
Fig. 8 Graphical comparison of the evaluation metrics of random forest



9 Comparative Study

In this comparative study, we examine and compare the performance of multiple machine learning models for forecasting Airbnb listing pricing. The models investigated in this study are linear regression, polynomial regression, CatBoost regression, gradient boosting regression, XGBoost regression, and random forest regression. We assessed their performance based on several metrics such as mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and R-squared (R2) score. Among the tested models, CatBoost Regressor emerged as the top model, followed by XGBoost Regressor and random forest regressor. XGBoost Regressor shows performance similar to CatBoost, demonstrating its ability to capture complex relationships in the database and provide accurate predictions. Also, random forest regressor gives strong performance with MAE 0.284, MSE 0.155, RMSE 0.394, and R2 0.700. These results are consistent with the performance of the ensemble model, demonstrating the effectiveness of the random forest algorithm in capturing complex relationships and achieving accurate predictions as per (Fig. 9).

Fig. 9 Metrics comparison of the different evaluation methods for the different algorithms. LR—Linear Regression; RF—Random Forest Regression; PR—Polynomial Regression; CBR—CatBoost Regressor; GB—Gradient Boosting; XGB—XGBoost Regression



10 Deployment

In this research study, we show the implementation of an advanced regression model to forecast the value of Airbnb using the Streamlit framework. The purpose is to give consumers an interactive online interface to enter property data and obtain a price estimate based on a trained model. Deploying Airbnb's price prediction model includes multiple phases, including loading the trained model, developing a user interface using Streamlit, and adding user input to make forecasts.

References

1. Bayoumi AE, Saleh M, Atiya AF, Aziz H (2012) Dynamic pricing for hotel revenue management using price multipliers. *J Rev Pric Manage* 12(3):271–285. <https://doi.org/10.1057/rpm.2012.44>
2. Chen CF, Rothschild R (2010) An application of hedonic pricing analysis to the case of hotel rooms in Taipei. *Tour Econ* 16(3):685–694. <https://doi.org/10.5367/000000010792278310>
3. Gite S, Khatavkar H, Kotecha K, Srivastava S, Maheshwari P, Pandey N (2021) Explainable stock prices prediction from financial news articles using sentiment analysis. *PeerJ* 7:e340. <https://doi.org/10.7717/peerj.cs.340>
4. Ma Y, Zhang Z, Ihler AT, Pan B (2018) Estimating warehouse rental price using machine learning techniques. *Int J Comput Commun Control* 13(2):235–250. <https://doi.org/10.15837/ijccc.2018.2.3034>
5. Masiero L, Nicolau JL, Law R (2015) A demand-driven analysis of tourist accommodation price: a quantile regression of room bookings. *Int J Hosp Manag* 50:1–8. <https://doi.org/10.1016/j.ijhm.2015.06.009>
6. Gupta A, Chaudhary DK, Choudhury T (2017) Stock prediction using functional link artificial neural network (FLANN). IEEE. <https://doi.org/10.1109/cine.2017.25>
7. Wang D, Nicolau JL (2017) Price determinants of sharing economy-based accommodation rental: a study of listings from 33 cities on Airbnb.com. *Int J Hospital Manage* 62:120–131. <https://doi.org/10.1016/j.ijhm.2016.12.007>
8. (n.d.-a). Anintuitive explanation of gradient boosting. Stanford University. https://www.cse.chalmers.se/~richajo/dit866/files/gb_explainer.pdf
9. Mehta P, Pandya S, Kotecha K (2021) Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. *PeerJ* 7:e476. <https://doi.org/10.7717/peerj.cs.476>

10. Kumara BA, Kodabagi MM, Choudhury T, Um J (2021) Improved email classification through enhanced data preprocessing approach. *Spatial Inform Res* 29(2):247–255. <https://doi.org/10.1007/s41324-020-00378-y>
11. (n.d.-b) Real Estate Price Prediction with Regression and Classification. Standford University. Autumn 2016(CS 229). http://cs229.stanford.edu/proj2016/report/WuYu_HousingPrice_report.pdf
12. Garlapati A, Garlapati K, Malisetty N, Krishna DR, Narayana G (2021) Price listing predictions and forthcoming analysis of airbnb.<https://doi.org/10.1109/iccenc51525.2021.9579773>
13. Yang S (2021) Learning-based airbnb price prediction model. In: 2021 2nd international conference on e-commerce and internet technology (ECIT). <https://doi.org/10.1109/ecit52743.2021.00068>
14. Patel J, Shah S, Thakkar P, Kotecha K (2015) Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Syst Appl* 42(1):259–268. <https://doi.org/10.1016/j.eswa.2014.07.040>
15. Rohini A, SudalaiMuthu T, Choudhury T (2021) Probabilistic Machine learning using social network analysis. In: Algorithms for intelligent systems, pp 1–12. https://doi.org/10.1007/978-981-33-4087-9_1
16. Dhillon J, Eluri NP, Kaur D, Chhipa A, Gadupudi A, Eravi RC, Pirouz M (2021). Anal Airbnb Prices us Mach Learn Tech. <https://doi.org/10.1109/ccwc51732.2021.9376144>
17. Lektorov A, Abdelfattah E, Joshi S (2023). Airbnb Rent Price Predict Mach Learn Models. <https://doi.org/10.1109/ccwc57344.2023.10099266>
18. Mahyoub M, Ataby AA, Upadhyay Y, Mustafina J (2023) AIRBNB price prediction using machine learning. In: 2023 15th international conference on developments in eSystems Engineering (DeSE), Baghdad and Anbar, Iraq, pp 166–171. <https://doi.org/10.1109/DeSE58274.2023.10099909>
19. Patel J, Shah S, Thakkar P, Kotecha K (2015) Predicting stock market index using fusion of machine learning techniques. *Expert Syst Appl* 42(4):2162–2172. <https://doi.org/10.1016/j.eswa.2014.10.031>
20. Singh S, Anand A, Mukherjee S, Choudhury T (2022) Machine learning applications in decision intelligence analytics. In: Springer eBooks, pp 163–178. https://doi.org/10.1007/978-3-030-82763-2_15

Tiger Community Analysis in the Sundarbans



Richa Choudhary, Tanupriya Choudhury, and Susheela Dahiya

Abstract It is common knowledge that the Sundarbans tiger reserve, which spans the border between India and Bangladesh, is one of the most important habitats in the world for the endangered Royal Bengal tiger (*Panthera tigris tigris*). The health of a tropical forest is strongly dependent on the variety of tiger species that live there, as well as their patterns of movement, the richness of their groups across time, and the ecosystem services that they provide. The purpose of this study is to investigate the relationship between tiger migration and the distribution of prey and topological features in this region. We are undertaking a community analysis of the tigers that live in the Sundarbans by looking at the distribution of their prey and mapping the vegetation in the area. We have been provided with datasets that comprise information on the distribution of prey, GPS locations of tigers, and vegetation mapping across the Sundarbans region by the Wildlife Institute of India (WII). The results of the research indicate that prey distribution, the kind of vegetation, and the time of day all play a significant role in determining the movements of tigers as a collective community

R. Choudhary (✉)

School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand 248007, India

e-mail: richachoudhary.86@gmail.com; r.choudhary@ddn.upes.ac.in

T. Choudhury

CSE Department, Symbiosis International (Deemed University), Symbiosis Institute of Technology, Pune, Maharashtra 412115, India

Ex-Professor, SoCS, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand 248007, India

S. Dahiya

School of Computer Science, Graphic Era Hill University, Dehradun, Uttarakhand 248002, India

Present Address:

T. Choudhury (✉)

Professor (Research), Graphic Era Deemed to be University, Dehradun, Uttarakhand 248002, India

e-mail: tanupriyachoudhury.cse@geu.ac.in; tanupriya.choudhury@sitpune.edu.in;
tanupriya@ddn.upes.ac.in; tanupriya1986@gmail.com

component, which can further assist researchers to develop measures and strategies for tiger conservation at Sundarbans region.

Keywords Royal Bengal tiger · Community analysis · Sundarbans

1 Introduction

Tigers, as vital apex predators, have a crucial role in preserving the delicate balance of their ecosystems. Regrettably, they have been confronted with numerous threats to their survival for many years, primarily due to human activities such as deforestation, poaching, and illegal trade in tiger parts [1]. Other factors contributing to the decline of tigers include conflicts between humans and tigers resulting in retaliatory killings due to attacks on livestock or humans by tigers; climate change impacting prey availability leading to starvation among tigers; and genetic issues caused by small fragmented subpopulations resulting in inbreeding depression [2]. By occupying an indispensable position as apex predators within food chains, tigers play a critical role in upholding ecological balance and conserving wildlife diversity [1]. They help regulate populations of prey species like deer, wild boar, and antelopes while also serving as indicators for habitat quality since they require vast areas with diverse vegetation types for thriving. Safeguarding tigers ensures the protection of other species that depend on similar habitats [3]. Situated at the junction of India and Bangladesh, the Sundarbans is a vast mangrove forest that holds not only UNESCO World Heritage status but also serves as a habitat for one of nature's most iconic and endangered creatures—the Bengal tiger (*Panthera tigris tigris*) [4]. Within this unique ecosystem, an intricate relationship exists between vegetation, prey availability, and water sources that significantly shape the behavior and survival of these majestic felines. The existence of flora inside the Sundarbans region is of utmost significance in assessing the appropriateness of habitats for tigers [5]. The dense mangrove forests offer vital cover for hunting activities as well as periods of rest. With its tall grasses, thickets, and interwoven roots acting as effective camouflage during stalking pursuits while providing sanctuary from potential dangers. By comprehending how different types of vegetation impact tiger movement patterns, conservationists can develop strategies to safeguard critical habitats while ensuring optimal conditions for successful hunts [5, 6]. Prey availability stands out as another pivotal factor influencing tiger populations in the Sundarbans [6]. The distribution and availability of prey species have a direct impact on the migrations and population dynamics of tigers [7]. To understand how tigers adapt their hunting behaviors, it is important to assess variations in prey densities across different regions within the Sundarbans. By identifying areas with higher concentrations or seasonal fluctuations in prey populations, conservation efforts can focus on protecting these crucial feeding grounds [8]. The presence of water sources is essential to the functioning of this one-of-a-kind ecosystem because these water sources both give tigers the

opportunity to drink and act as pathways that connect various sections of their territory [6]. Access to freshwater rivers, creeks, or tidal channels can have an effect on the migrations of tigers. Natural water resources are becoming scarce due to tidal oscillations & changes in the salinity levels [6]. The studies conducted to understand how tigers interact with their surroundings including vegetation, water sources, other species is known as community analysis. In order to perform community analysis, statistical analysis is conducted on the GPS data of tigers, topography of Sundarbans and prey distribution of the region. To understand the behavioral aspect of tiger in their ecological niche, it is important to analyze their movements & their interactions with the geographical elements [8]. This study is conducted to acquire better understanding of their social structure, daily activity cycles, individual behavior, and their interaction with other species. This can help in providing tailored conservation strategies that can help to address important issues impacting the overall population trends. These strategies can be implemented to conserve the tiger populations. While performing tiger's community analysis, it is important to understand their diets and how their diets can shift due to the scarcity of prey [9]. Another important factor to be considered are vegetation of the region and water supplies among the others [10]. Another crucial factor is examining their daily activity cycle within a community might provide valuable insights into their behavior [11]. The findings reveal that vegetation cover has a significant influence on tiger movement patterns, as they tend to frequent areas with dense vegetation rather than sparse or aquatic regions. Additionally, tigers show more activity during evening and nighttime hours. These results are further elaborated upon in the corresponding section of this report.

2 Materials and Methods

2.1 Study Area

Sundarbans is a world heritage UNESCO site as it is the world's largest Mangrove forest where tigers reside. It is a delta formed by the confluence of Brahmaputra, Ganges, and Meghna rivers in Bay of Bengal. Sundarbans spans across India and Bangladesh, covering an area of approximately 10,000 km² [12]. Sundarbans is an ecosystem comprising varied flora and fauna with geographical features including extensive mangrove forests, tidal streams, mudflats, sandy beaches, and number of small islands [6, 12]. The forest's network of rivers provides a habitat for species such as fish, crabs, and shrimp. The tides routinely replenish the mud layers on the forest floor. This ecosystem not only acts as a natural home for wildlife species like as tigers, but it also acts as a buffer against coastal erosion and storm surges because of the mangrove forests, which helps protect both the local residents that live in this region as well as their environment from the possible damage that can be caused by natural disasters [12]. In the course of time, sediment that was deposited by river currents generated a number of small islands or 'char', which contributed to the

specific characteristics of the area. The Sundarbans are distinguished by the extraordinary ecological diversity that they harbor, which is complimented by the fact that they play an important part in ensuring the survival of tigers. This part of the world is famous for being home to the biggest number of Bengal tigers anywhere in the world [1]. These magnificent animals are able to move around with ease within this deep mangrove forest because of the linked waterways that are great hunting grounds provided by this ecosystem's unique features. The Sundarbans Tiger Reserve is a protected region in the state of West Bengal in India that was formed in 1973 with the purpose of conserving the Bengal tiger, which is on the verge of extinction [1]. The reserve has a total area of around 2585 km^2 and is home to approximately one hundred Royal Bengal Tigers [1, 12]. These magnificent cats are excellent swimmers and seek water animals for their food. These tigers have adjusted to their environment, which includes a number of obstacles such as thick vegetation, tidal streams, and mudflats. However, they have been able to overcome these obstacles. Their unique ability to swim long distances and hunt aquatic prey like fish, crabs, and small crocodiles is one of their most notable features [3, 6]. Additionally, they are known to climb trees when water levels rise during high tide. Despite being protected by conservation measures, these tigers still face multiple threats within their habitat. The primary threat comes from poaching for body parts like bones or skin that are highly valued in traditional Chinese medicine [3]. Other dangers include habitat loss due to deforestation or land conversion for agriculture; human-wildlife conflict resulting from encroachment into tiger habitats; climate change-induced sea level rise, etc. The reserve also supports diverse flora and fauna but faces significant threats from human activities such as poaching and habitat destruction due to deforestation for timber or agriculture [3, 6, 12]. Sundarbans' unique landscape plays a crucial role in supporting rich biodiversity while also providing livelihoods for millions who depend on fishing, honey collection, and other activities. Nevertheless, the vulnerable environment of Sundarbans is confronted with substantial risks due to climate change-induced sea level rise, as well as human activities like as deforestation and pollution. Conservation measures implemented by government agencies include increasing patrols against poaching, promoting eco-tourism, establishing community-based conservation programs involving local communities in protecting natural resources while providing alternative livelihood options, and reforestation efforts. To combat these threats, government agencies along with NGOs focused on wildlife protection have implemented various conservation efforts over time [1] and the study area for this work—Sunderbans region has been shown in (Fig. 1).

Dataset Used

This work is using the data of 4 tigers and 11 prey for the Sundarbans region. The data is collected by the Wildlife Institute of India (WII), located in Dehradun, India. This data is collected using telemetry collars on tigers, it provides the locational information of the species [13]. The data contains the location of the tigers of approximately each day for around a year. Locations are the GPS coordinates of the sampled region for respective tigers, the dataset contains around 4494 cumulative GPS locations of tigers with no outliers. The dataset comprises tuples that consist of a collection of

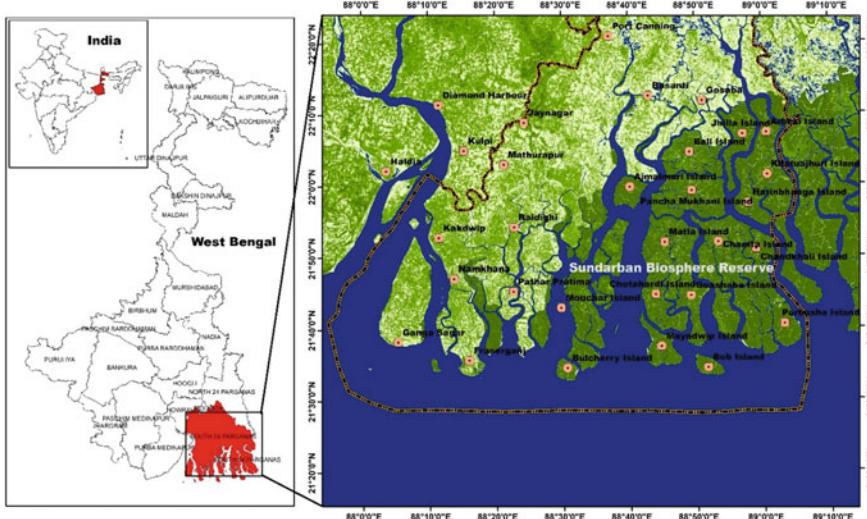


Fig. 1 Study area for this work—Sundarbans Region

attributes, as ID of the collared tiger, Date, Time, Latitude, and Longitude. To collect the data, tigers were captured and collared by the professionals using collars such as Telonix VHF MOD 400, GPS PLUS IRIDIUM. The radio collars were equipped with a customizable GPS schedule that had the ability to record location data at intervals ranging from 1 to 3 h. The acquired GPS locations of these 4 tigers are mapped and presented in Fig. 2. The data regarding prey distribution includes the geographical coordinates (latitude, longitude, or GPS coordinates) of 11 different prey species in the specified region. This information is collected annually by WII and obtained from their records. There were a total of eleven distinct prey species observed, including the Wild Pig, Rhesus Macaque, Chital, Estuarine Crocodile, Human, King Cobra, Lesser Adjutant Stork, Otter, Red Jungle Fowl, Water Monitor Lizard, and Egret. The data contains many GPS coordinates for the prey. Another dataset received from WII is of vegetation mapping for the region Sundarbans. It contains the NDVI value of the landscape.

2.2 Topography (Vegetation Mapping) of Sundarbans

The research utilized a Landsat TM satellite image, specifically focusing on the Indian sector of the Sundarbans region. The resolution of this image, which was acquired from the Earth Explorer website of the United States Geological Survey (USGS), is 30 m. The Earth Resource Data Analysis System (ERDAS) Imagine program was utilized to employ the vegetation delineation function in order to account for atmospheric variables. Furthermore, the haze reduction capabilities available in ERDAS

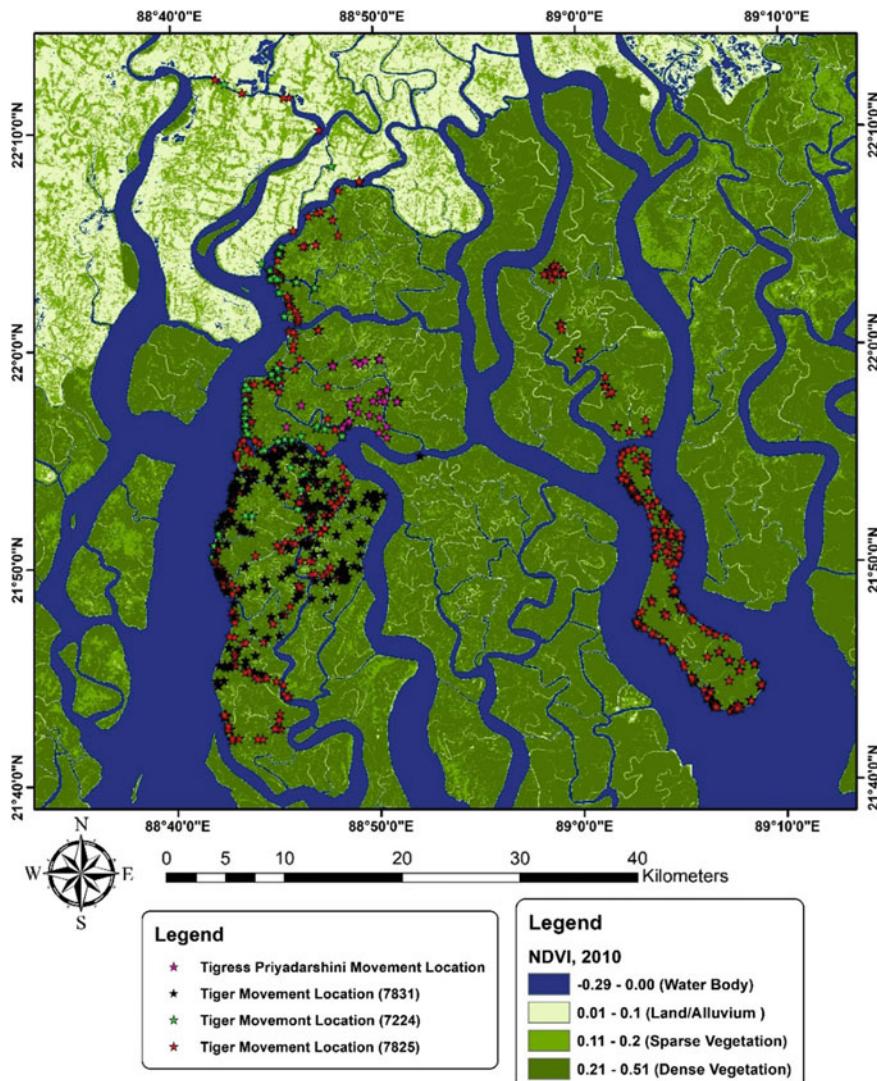


Fig. 2 Vegetation mapping of Sundarbans

software were employed to effectively reduce small regions affected by various forms of haze present in the photos. Near-infrared (NIR) and red bands are utilized to find the normalized difference vegetation index (NDVI) maps of the region. The reclassification of these maps was conducted by employing a range of NDVI values spanning from -1 to $+1$, which corresponds to vegetation levels determined through analysis of spectral reflections.

$$\text{NDVI} = (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$$

The above mentioned formula is used for the NDVI analysis, resulting in the identification of four separate categories: water body (-0.29 to 0.00), land/alluvium (0.01 to 0.1), sparse vegetation (0.11 to 0.2), and thick vegetation (0.21 to 0.51). In order to achieve the aims of the study, the researchers employed digital image processing, and ArcGIS software, for the purposes of picture alteration, categorization, analysis, and the production of the NDVI map. The utilization of ArcGIS played a crucial role in the generation of a fake color composite through the amalgamation of near-infrared (NIR), red, and green spectral bands extracted from Landsat TM 2010 satellite images. The utilization of this composite enabled the discernment of vegetation as a result of the higher reflectance of chlorophyll in the near-infrared spectrum in contrast to wavelengths within the visible range.

2.3 Prey Distribution in Sundarbans

The pursuit of prey is a significant factor contributing to the locomotion patterns observed in carnivorous animals. These creatures exhibit locomotion and engage in predation according to their capabilities [14]. Tigers inhabit regions where viable prey sources are present. Certain movements may have a shorter range, while others may have a greater range [15]. The velocity and overall condition of the tiger are determined by its movement patterns. The research study focused on observing the actions of four tigers as a representative sample, specifically in relation to their choosing of prey. The research incorporated the prey distribution dataset, which was obtained from the Wildlife Institute of India (WII). There were a total of eleven distinct prey species observed, specifically identified as ‘Wild Pig’, ‘Chital’, ‘Human’, ‘Lesser Adjutant Stork’, ‘Otter’, ‘Red Jungle Fowl’, ‘King Cobra’, ‘Rhesus Macaque’, ‘Water Monitor Lizard’, and ‘Estuarine Crocodile’. The information contains many GPS positions for the prey.

2.4 Data Analysis and Interpretation

This study utilizes three datasets obtained from WII to predict the movement patterns of tigers. Since movements are closely tied to behavioral traits [16], analyzing these patterns is crucial for understanding tiger behavior. To gain insight into their movement patterns, telemetry data was collected from the Sundarbans region along with prey distribution information. Furthermore, it is crucial to examine the landscape characteristics and vegetation mapping data of this particular area in order to have a comprehensive understanding of tiger ecology, as these elements significantly influence their patterns of migration [17]. The details of the datasets are already discussed under dataset section. After acquiring the data, data preprocessing is the next valid

step, the data used in this study was of high quality and collected by the WII, resulting in no outliers. Any missing entries or values were addressed by dropping empty rows at the beginning or end of the data set. For missing values within the data, a 5-period moving average was substituted for latitude and longitude columns while timestamp values were incremented. The wildlife institute's dataset consists of time series data, as it includes timestamped information on the movements of tigers in search of prey. To perform the community analysis of tigers in Sundarbans, it is divided into 3 parts, firstly the analysis of vegetation mapping is done w.r.t the tigers, then it is analyzed how prey distribution impact their movements, finally their temporal movement patterns are analyzed.

Community Analysis of Vegetation Cover and Water Resources

The movement of tigers is greatly affected by vegetation, playing a vital role in their behavior, hunting patterns, and overall survival [18]. Here are some important factors to consider: Vegetation provides crucial habitat for tigers as it offers cover for resting, hiding, and stalking prey. Dense vegetation such as tall grasses, thickets, or forests allows tigers to effectively blend in while patiently waiting to ambush their prey [19]. Tigers primarily hunt alone and rely on stealthy approaches to catch their prey. Vegetation helps them remain hidden from potential victims until they can get close enough for a successful attack. The presence of dense vegetation increases the success rate of tiger hunts by providing effective concealment during the stalk. Vegetation acts as natural pathways connecting different habitats within a landscape. Dense vegetation often supports higher populations of herbivores like deer or wild boar—which serve as primary prey for tigers [3, 20]. Suitable vegetative cover encourages these herbivores to thrive due to the availability of food sources and protection from predators. Vegetation plays an important part in the community formation of any species [21].

To perform this analysis the following features are utilized:

- (a) Four classes are created ‘Land’, ‘Sparse vegetation’, ‘Dense vegetation’, ‘Water Body’.
- (b) Time is divided into 4 classes ‘morning: {5:00, 12:00}’, ‘Afternoon: {12:00, 17:00}’, ‘Evening: {17:00, 20:00}’, ‘Night: {20:00, 5:00}’.
- (c) The data on the location of tigers, along with timestamps, is compared to determine how many times a tiger has been found in a particular vegetation type. This is done by increasing the corresponding vegetation counter based on pre-defined time classes.

Community Analysis of Prey Species

In order to conduct an analysis on the geographical locations of tigers and their proximity to the average value of prey, the latitude–longitude or GPS coordinates are transformed into a kilometer-based scale. The following formula is used to calculate the distance between data points in kilometers approximately [22].

$$a = \sin^2(\Delta\varphi/2) + \cos\varphi_1 \cdot \cos\varphi_2 \cdot \sin^2(\Delta\lambda/2) \quad (1)$$

$$c = 2 \cdot \text{atan2}\left(\sqrt{a}, \sqrt{(1-a)}\right) \quad (2)$$

$$d = R \cdot c \quad (3)$$

where

$\Delta\phi$ is representing difference between latitudes of tiger/prey I.E. Latitude1–Latitude2 for given GPS coordinates.

$\Delta\lambda$ is representing difference between longitudes of tiger/prey I.E. Longitude1–Longitude2 for given GPS coordinates.

R is radius of earth in Km = 6371.

D is distance between two GPS locations in Km.

To analyze the impact of prey distribution on tigers, we have taken the location of different preys and checked how many times it was within the range of 1 sigma, 2 sigma, 3 sigma deviation of tiger location.

Community Analysis of Daily Movements of Tigers

The locational data of sampled tigers is analyzed and descriptive statistics is performed to understand the basic patterns in the data. The data is analyzed to find the average movement in 1 h, total area explored by the tiger, percentage of Sundarbans area covered by the tiger. Table 1 shows data of these 4 tigers.

The average distance moved in a time span of 1 h by tiger 7831 is 0.3606 km. Tiger 7825 exhibited a mean displacement of 0.2985 km during the course of one hour. The tiger with identification number 7224 exhibited an average displacement of 0.9443 km, whilst Tigress_Priyadarshini traveled an average distance of 1.5721 km within a one-hour time period. After calculating the comparative movement of all the four subjects individually with respect to all of them together, in the areas of Sundarbans forest, it is concluded that the tiger 7831 moved 11.37%. Tiger 7825 moved 9.41% in comparison to the whole. Tiger 7224 moved 29.72% more while Tigress_Priyadarshini moved and covered 49.49% of the area. Upon calculating the male–female movement ratio for the given four subjects, the ratio came up to be 0.5344:1.5721. Tiger displacement was another parameter taken under consideration for the data wherein statistical analysis is performed to find out how much did the

Table 1 The average hourly displacement of radio collared tigers in the Sundarbans region

Tigers	Movement track (1 h Average) (in KMs)	Comparative Movement (in %)	Male–female movement ratio (male: female)	Area explored (in sq. km)	Area of tiger (x)/area of Sundarbans (y) (%)
7831	0.3606	11.37	0.5344:1.5721	11	0.18
7825	0.2985	9.41		70	1.17
7224	0.9443	29.72		70	1.17
Priyadarshini	1.5721	49.49		15	0.25

tiger displace in an hour's time interval. Tiger 7831 dispersed a total of 11 km², tiger 7825 and tiger 7224 displaced 70 km². in an hour, individually. Tigress_Priyadarshini covered an area of 15 km². (on an average) in an hour.

3 Results and Discussion

The main objective of this study is to enhance knowledge of tiger's community structure. To achieve this aim, the study has examined factors such as the movement data of tigers, prey distribution, and vegetation of the region. Statistical analysis is performed to understand the correlations that may exist between tiger behavior/prey availability/vegetation characteristics. Detailed assessments of vegetation preferences are conducted by analyzing movement data over the period of time. The analysis of movement data in context to prey is performed to provide a holistic understanding of the community structure of tigers. The findings from this research can then be used to inform conservation strategies and management plans aimed at ensuring the long-term survival and well-being of these magnificent creatures in their natural habitats.

3.1 Movements of Tigers in Sundarbans Landscape

The plots in the Fig. 3 explain the fact that movement is subject to a normal distribution and tigers were mostly confined within 3 sigma deviation from the mean. It can be further seen from the graphs that there can be huge difference in the movement dynamics of the tigers and 1 sigma deviation can be as little as around 5 kms for female tigers and go upto 40 kms for individual male tigers. This can extend upto 120 in case of 3 sigma deviations from the mean. It is also observed from the following charts that there is difference between the movement of male and female tigers. In the wild, there exists a notable distinction in the conduct displayed by males and females. Both genders are required to engage in hunting activities for their sustenance. Initially it is assumed that there is no noteworthy contrast in the movement patterns exhibited by male and female tigers during their hunt for food. So, the null hypothesis is constructed as

$$H_0 = \text{Male and female tigers show equal amount of displacement}$$

A statistical analysis using a T test was performed on the movement data of the four tigers, resulting in a p-value of 0.148120. Given that the p-value surpasses the threshold of 0.05, it is appropriate to infer that the observed outcome lacks statistical significance and the hypothesis is rejected. Figure 3 illustrates deviation plots from the average latitude and longitude coordinates. There are following observations from the plots as in Fig. 3:

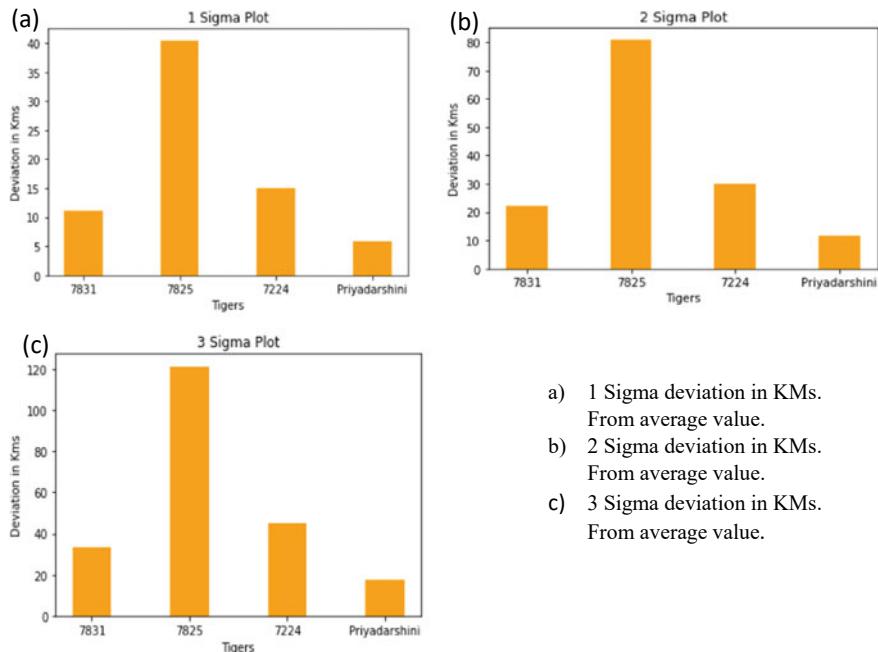


Fig. 3 Movement of tigers in Sundarbans

1. Female displacement is minimum as compared to male counterparts.
2. Tiger 7825 shows the maximum displacement as compared to other tigers, indicating that some tigers can have extreme displacements from their mean.

3.2 *The Influence of Vegetation on the Temporal Displacement of Tigers*

In this study to understand the impact of vegetation in Sundarbans for tigers, following hypothesis is assumed.

$$H_0 : \text{Time of day has no impact on tiger movement}$$

To validate the hypothesis, NDVI average of the grid in which the tiger was present was taken and count was generated. This was further segregated upon the time of day as shown in Fig. 4. Contrary to the notion that the time of day has no impact on tiger movement, it is worth noting that tigers are crepuscular creatures, meaning they are most active during dawn and dusk. Although tigers can be active at any hour, their hunting behavior and movement tendencies are more pronounced during these periods of low light. The time of day influences tiger movement for a variety of

reasons: Tigers have adapted to primarily hunt during twilight hours when visibility is diminished. Their exceptional night vision combined with the concealment provided by darkness enables them to approach prey without being detected [3, 22]. This grants them a strategic advantage in capturing their preferred prey species. Tigers inhabit regions with high temperatures, particularly in tropical climates. Tigers frequently exhibit a behavioral pattern of seeking shade and engaging in periods of rest during the most intense periods of daylight, as a means to conserve energy and mitigate the risk of overheating. Consequently, they may display less activity during midday hours when temperatures reach their peak. Many herbivores (such as deer), which constitute a significant portion of a tiger's diet [23], also exhibit heightened activity levels during early morning or late afternoon when temperatures are cooler and vegetation is more accessible for grazing purposes. Tigers adjust their movements accordingly to capitalize on these predictable patterns among their prey.

As evident from the movement plots versus time of day the following observations can be made:

1. Tigers have preference of moving during the night time,
2. Tigers prefer vegetation over land and water bodies,
3. Dense vegetation being the most favored among all these.

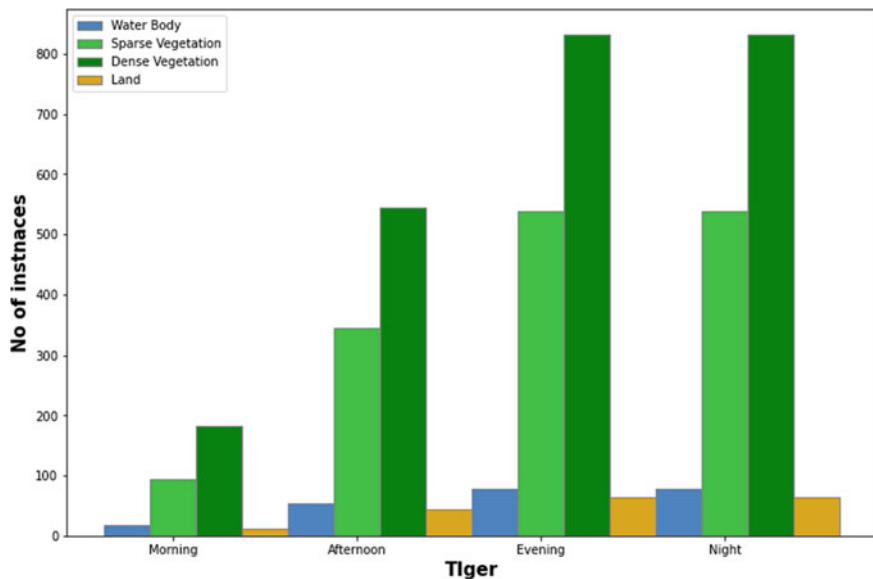


Fig. 4 Bar chart showing temporal presence of tigers in different types of vegetation

3.3 Impact of Prey Distribution in Region

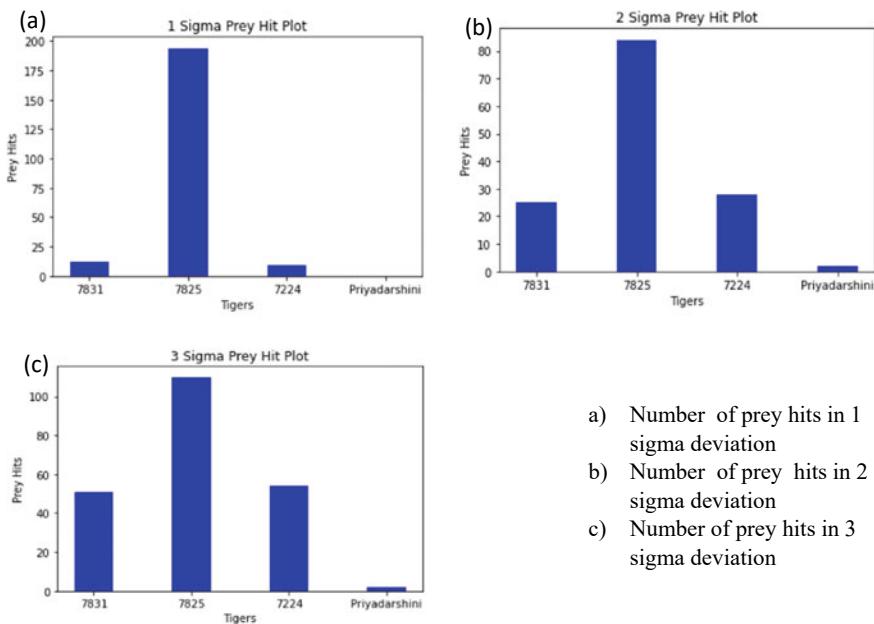
Tigers, being apex predators, are highly dependent on the availability and abundance of suitable prey species in order to ensure their survival. In order to fulfill their hunting needs, tigers require large territories that provide access to sufficient prey resources. The size of a tiger's home range depends on the density and distribution of its preferred prey species [24]. If there is an ample concentration of viable prey in an area, tigers may have smaller home ranges as they can find enough food within a more confined space. The availability of prey directly influences how tigers hunt and navigate through their habitat. When there is an abundance of prey in one specific area, tigers tend to concentrate their hunting efforts there, resulting in more frequent movements within that particular territory or zone. Certain regions experience seasonal variations in the availability of prey due to migration patterns or changes in vegetation growth cycles. The way prey is distributed in a particular area has a significant impact on how tigers move and behave. The following plots explain the hunting for tigers. It can be observed that there is difference between prey hunts of male and female tigers. The male tigers were able to hunt much more than the female tigers. Female tiger/tigress is able to successfully hunt in the 3 sigma deviation range. Also, number of successful hunts seems to be dependent upon the movement from the mean deviation. Tigers adapt their movements accordingly by following herds or concentrating around areas where certain types of prey become more abundant during different times of the year. Prey distribution also plays a role in determining which natural corridors tigers use as they travel between different habitats or territories while searching for food sources.

The following observations can be made from the Fig. 5:

1. Female tigers have least prey hunts,
2. Tiger 7825, shows the maximum success while hunting depicting that higher movement is positively correlated with movement of tigers,
3. All tigers are able to hunt in 2 sigma deviation.

4 Conclusion

There is a distinct population of royal Bengal tigers that can only be found in the Sundarbans delta, which is situated in the Bay of Bengal. This region's environment is one of a kind and extremely diversified. This population has differentiated themselves into their own distinct community as a result of their adaptation to the presence of mangrove trees inside the delta basin. The vegetation is a significant factor that plays a role in determining the patterns of movement of these tigers. During times when they are on the prowl for prey, tigers can make excellent use of the mangroves' thick undergrowth to conceal themselves. Tigers are known to favor regions that have a lot of flora because it helps them blend in with their environment and improves their chances of effectively ambushing their victim. In addition, research that was carried



- a) Number of prey hits in 1 sigma deviation
- b) Number of prey hits in 2 sigma deviation
- c) Number of prey hits in 3 sigma deviation

Fig. 5 Hunt patterns based on prey distribution of Sundarbans region

out on this group of tigers revealed that the most of the time, they are nocturnal animals. They have a tendency to be more active during the nocturnal hours. In terms of movement patterns, the analysis that was carried out as part of this research project revealed fascinating data linked to the distance that tigers travel from their typical site, also known as their home range. It was found that every individual was able to hunt within a mobility range that was no more than three standard deviations (sigmas) away from their typical location. This was one of the observations that was made. This points to a bell-shaped or Gaussian distribution pattern for the tiger's movements being the most likely explanation. These findings provide important new information about the ways in which royal Bengal tigers adapt and survive in the ecosystem of the Sundarbans delta. The fact that they like to spend the night in heavily wooded areas is evidence that they are able to effectively exploit available resources. Understanding these precise facts of tiger behavior can improve conservation efforts overall by assisting researchers in locating crucial habitats and developing management measures that are specifically customized for the one-of-a-kind population that can only be found in the Sundarbans delta region.

References

1. Qureshi Q, Jhala YV, Yadav SP, Mallick A (2023) Status of tigers, co-predators and prey in India, 2022. National tiger conservation authority, government of India, New Delhi, and Wildlife institute of India, Dehradun ISBN No: 81–85496–92–7
2. Bisht S, Banerjee S, Qureshi Q, Jhala Y (2019) Demography of a high-density tiger population and its implications for tiger recovery. *J Appl Ecol* 56(7):1725–1740. <https://doi.org/10.1111/1365-2664.13410>
3. Mukherjee S, Sen Sarkar N (2013) The range of prey size of the royal Bengal tiger of Sundarbans. *J Ecosyst* 2013:1–7. <https://doi.org/10.1155/2013/351756>
4. Singh SK, Mishra S, Aspi J, Kvist L, Nigam P, Pandey P, Sharma R, Goyal SP (2015) Tigers of Sundarbans in India: Is the population a separate conservation unit? *PLoS One* 10(4). <https://doi.org/10.1371/journal.pone.0118846>
5. Saklani A, Navneet B, Bhandari S (2019) A community analysis of woody species in tropical forest of Rajaji tiger reserve. *Environ Ecol Res* 37:48–55
6. Naha D, Jhala YV, Qureshi Q, Roy M, Sankar K, Gopal R (2016) Ranging, activity and habitat use by tigers in the mangrove forests of the sundarban. *PLoS One* 11(4). <https://doi.org/10.1371/journal.pone.0152119>
7. Nilsen EB, Christianson D, Gaillard JM, Halley D, Linnell JDC, Odden M, Panzacchi M, Toigo C, Zimmermann B (2012) Describing food habits and predation: field methods and statistical considerations. *Carnivore Ecol Conserv* 256–272. <https://doi.org/10.1093/acprof:oso/9780199558520.003.0011>
8. Li Z, Kang A, Gu J, Xue Y, Ren Y, Zhu Z, Liu P, Ma J, Jiang G (2017) Effects of human disturbance on vegetation, prey and Amur tigers in Hunchun Nature Reserve, China. *Ecol Model* 353:28–36. <https://doi.org/10.1016/j.ecolmodel.2016.08.014>
9. Choudhary R, Choudhury T, Dahiya S (2023) Exploring tiger movement pattern according to prey context: a case study in Sundarbans region of India. *Spat Inf Res*. <https://doi.org/10.1007/s41324-023-00525-1>
10. Koulgi PS, Clinton N, Karanth KK (2019) Extensive vegetation browning and drying in forests of India's tiger reserves. *Scientific Rep* 9(1). <https://doi.org/10.1038/s41598-019-51118-8>
11. Reddy HS, Srinivasulu C, Rao KT (2004) Prey selection by the Indian tiger (*Panthera tigris tigris*) in Nagarjunasagar Srisailam Tiger reserve India. *Mammalian Biol* 69(6):384–391. <https://doi.org/10.1078/1616-5047-00160>
12. Wikimedia Foundation. (2023) Sundarbans National park. Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Sundarbans_National_Park
13. Choudhary R, Dahiya S, Choudhury T (2023) A review of tiger conservation studies using nonlinear trajectory: a telemetry data approach. *Nonlinear Eng* 12(1):20220235. <https://doi.org/10.1515/nle-2022-0235>
14. Duangchtrasiri S, Jornburom P, Jinamoy S, Pattanvibool A, Hines JE, Arnold TW, Fieberg J, Smith JL (2019) Impact of prey occupancy and other ecological and anthropogenic factors on tiger distribution in Thailand's western forest complex. *Ecol Evol* 9(5):2449–2458. <https://doi.org/10.1002/ece3.4845>
15. Sarkar MS, Amonge DE, Pradhan N, Naing H, Huang Z, Lodhi MS (2021) A review of two decades of conservation efforts on tigers, co-predators and prey at the junction of three global biodiversity hotspots in the transboundary far-eastern Himalayan landscape. *Animals* 11(8):2365. <https://doi.org/10.3390/ani11082365>
16. Rew J, Park S, Cho Y, Jung S, Hwang E (2019) Animal movement prediction based on predictive recurrent neural network. *Sensors* 19(20):4411
17. Athreya V, Navya R, Punjabi GA, Linnell JD, Odden M, Khetarpal S, Karanth KU (2014) Movement and activity pattern of a collared tigress in a human-dominated landscape in Central India. *Tropic Conserv Sci* 7(1):75–86. <https://doi.org/10.1177/194008291400700111>
18. Bhardwaj G (2021) The spacing pattern of reintroduced tigers in human-dominated Sariska tiger reserve. *J Wildlife Biodiv* 5:1–14. <https://doi.org/10.22120/jwb.2020.124591.1129>

19. Wang Y, Cheng W, Guan Y, Qi J, Roberts NJ, Wen D, Cheng Z, Shan F, Zhao Y, Gu J (2023) The fine-scale movement pattern of Amur tiger (*Panthera tigris altaica*) responds to winter habitat permeability. *Wildlife Lett.* <https://doi.org/10.1002/wll2.12020>
20. Yadav PK, Brownlee MT, Kapoor M (2022) A systematic scoping review of tiger conservation in the terai arc landscape and Himalayas. *Oryx* 56(6):888–896. <https://doi.org/10.1017/s0030605322001156>
21. Sunarto S, Kelly MJ, Parakkasi K, Klenzendorf S, Septayuda E, Kurniawan H (2012) Tigers need cover: Multi-scale occupancy study of the big cat in Sumatran forest and plantation landscapes. *PLoS One* 7(1). <https://doi.org/10.1371/journal.pone.0030859>
22. Lior K (2017) Calculate distance between two latitude-longitude points? (Haversine formula). Stack Overflow. Retrieved from <https://stackoverflow.com/questions/27928/calculating-distance-between-two-latitude-longitude-points-haversine-formula>
23. Prajapati RK, Triptathi S, Mishra RM (2014) Habitat modeling for Tiger (*Panthera tigris*) using geo-spatial technology of Panna Tiger Reserve (M.P.) India. *Int J Sci Res Environ Sci* 2(8):269–288. <https://doi.org/10.12983/ijres-2014-p0269-0288>
24. Cushman SA, Krishnamurthy R (2023) Modelling landscape permeability for dispersal and colonization of tigers (*Panthera tigris*) in the greater panna landscape Central India. *Landscape Ecol* 38(3):797–819. <https://doi.org/10.1007/s10980-022-01590-x>

An Overview of the Use of Deep Learning Algorithms to Predict Bankruptcy



**Kamred Udham Singh, Ankit Kumar, Gaurav Kumar, Teekam Singh,
Tanupriya Choudhury, and Ketan Kotecha**

Abstract The financial forecasting of different firms in the area of financial status aims to determine whether the company will go bankrupt in the near future or not. This is a critical problem for these companies. Several companies have shown a strong interest in this area, particularly since they are concerned about the future of their companies from a financial perspective and want to determine whether or not they will go out of business. Therefore, in the work that we have done, we have presented

K. U. Singh

School of Computing, Graphic Era Hill University, Dehradun, India

A. Kumar

Department of Information Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, Chhattisgarh, India

e-mail: iiita.ankit@gmail.com

T. Singh

Department of Computer Science and Engineering, Graphic Era Deemed to Be University, Dehradun, India

e-mail: tsingh@ma.iitr.ac.in

T. Choudhury (✉)

CSE Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

e-mail: tanupriyachoudhury.cse@geu.ac.in; tanupriya.choudhury@sitpune.edu.in;
tanupriya1986@gmail.com

K. Kotecha

Symbiosis Centre for Applied Artificial Intelligence, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune 411045, India

Present Address:

T. Choudhury

Professor (Research), CSE Department, Graphic Era Deemed to be University, Dehradun, Uttarakhand 248002, India

G. Kumar

Department of Computer Engineering & Application's, GLA University, Mathura, India
e-mail: gaurav.kumar_phd.ca21@gla.ac.in

three well-known technologies of deep learning in conjunction with ensemble classifiers and boosting ensemble classifiers for the purpose of failure prediction. During our investigation, we used an uneven dataset consisting of businesses from Spain, Poland, and Taiwan. In addition to this, we applied approaches such as oversampling, hybrid balancing, and clustering-based balancing to get rid of the inconsistent data. When taking into account a real-life financial dataset with an appropriate amount of complexity, it was discovered that the MLP-6L model with the SOMTE-ENN balancing approach had the most remarkable performance when measured against the metrics.

Keywords MLP-6L · Deep learning · SVM · RF · SOMTE-ENN

1 Introduction

The issue of insolvency among companies resonates beyond their home countries, impacting a wide array of investors who hold financial stakes in these enterprises. The problem of bankruptcy among businesses extends beyond their national borders, affecting a broad range of investors who have financial interests in these corporations. As a consequence of this situation, the deep learning [1, 2] community has been more interested in bankruptcy prediction, which has led to the creation of a wide variety of approaches. A significant number of organizations are ready to collect thorough financial information and evaluate the state of enterprises that have been through the bankruptcy process. In the actual world, bankruptcies are not very common events, which results in the generation of unbalanced datasets in this field. Our particular research endeavors attempt to examine a variety of different methods for forecasting the possibility of bankruptcy for a certain organization [3]. We used financial information from Spanish enterprises in our comparison study of the random forest (RF), Naive Bayes, and J48 decision tree classifiers. The RF classifier was our primary focus in this study. Given the imbalanced composition of these data collections, we opted for a tactical method to address the issue of data irregularities. This involved implementing techniques such as random oversampling, undersampling, and combined methods. Considering the uneven nature of these data sets, we chose a strategic approach to tackle the problem of data inconsistencies. This included the use of methods like random oversampling, undersampling, and a mix of both techniques. This allowed us to combat the issue of data inconsistency. Combining the J48, KNN, and MLP classifiers with the Dynamic Laplace Relevance (DLR) status space improved our models' predicted accuracy even more. The performance of the classifiers was greatly enhanced as a consequence of this enhancement, and it was even able to outperform the results produced through traditional ensemble voting or from solo classifiers. For instance, the DLR method produced much better results.

In the course of this study, we investigated numerous sophisticated classification methods, namely MLP-6L, LSTM, and DBN, to determine the possibility of bankruptcy across a number of different organizations. We also investigated ensemble

techniques, including bagging-based approaches like RF, SVM, and KNN, in addition to boosting-based strategies like AdaBoost and XGBoost [4].

Our study was carried out with the assistance of datasets that needed to be revised, and they represented genuine businesses from Poland, Spain, and Taiwan. Notable differences existed across the datasets in terms of the complexity and variety of data types. The dataset from Spain included both financial and non-financial data, while the dataset from Taiwan claimed the highest number of characteristics and the biggest number of those attributes combined. On the other hand, the dataset pertaining to Poland was distinguished by having the greatest number of samples. We used a total of eight different sophisticated approaches to data balancing in order to overcome the problem of inconsistent data. Oversampling, under-sampling, hybrid oversampling, and approaches based on clustering were all included in these strategies. In order to guarantee the durability and dependability of our bankruptcy prediction models in the face of unbalanced datasets and various sorts of data, these procedures were essential.

2 Datasets Considered

In the field of deep learning [5, 6], the selection and use of our datasets constitute the central pillar of our inquiry. This is especially true in terms of the research that we are doing. Specifically, we have been concentrating our efforts on the datasets that originate from businesses in Spain, Poland, and Taiwan. We have relied on the statistics that Infotel has gathered. This resource is well-known for its extensive collection of data pertaining to corporations to get the information that comes from Spanish businesses. Infotel rigorously collects information on the beginning and development of these firms, as addition to the following financial and operational state of these businesses [7]. This thorough dataset includes both financial indicators and non-financial characteristics, offering a complete perspective of each company's path, including whether or not they have experienced financial solvency or bankruptcy [8]. The data was collected throughout the USA.

Our database contains genuine information obtained over a period of six years from 471 Spanish businesses. There are a total of 2859 examples that have been painstakingly recorded inside this dataset. 39 distinct categories and numerical characteristics characterize each of these instances. We have deliberately cut down our focus to evaluate just 16 relevant independent variables to simplify our study. Any features that were judged irrelevant have been eliminated from consideration. Table 1 provides an overview of the dataset's organizational structure [9, 10]. There is a substantial disparity in the data that is represented by the dataset that is comprised of Spanish enterprises. A stunning 98% of the data reflects financially stable companies, while just 2% of the data corresponds to organizations that have declared bankruptcy. This apparent imbalance presents a difficulty for machine learning algorithms, which often gravitate toward the easy route of predicting outcomes based on the class that constitutes most of the population [11].

Table 1 Variables by eliminating the irrelevant attributes the sample

Financial variables	Description	Type
Debt structure	Long-term liabilities/current liabilities	Real
Debt paying ability	Operating cash flow/total liabilities	Real
Operating income margin	Operating income/net sales	Real
Warranty	Financial warrant	Real
Stock turnover	Cost of sales/average inventory	Real
Debt ratio	Total assets	Real
Working capital	Working capital/total assets	Real
Debt cost	Interest cost/total liabilities	Real
Return on operating assets	Operating income/average operating assets	Real
Asset rotation	Asset allocation decisions	Real
Financial solvency	Current assets	Real
Asset turnover	Net sales	Real
Receivable turnover	Net sales	Real
Acid test	(Cash equivalent + marketable securities + net receivables)/current liabilities	Real

We have also included data from Taiwanese firms since we are aware of the need to use a varied collection of datasets to conduct an exhaustive analysis of our methods. The complexity of this dataset, which was gathered from the Taiwan Economic Journal, is far higher than that of its equivalent in Spanish. It covers a period of ten years and includes a total of 6819 records, out of which 6599 indicate enterprises that are in good financial standing and 220 indicate companies that have been forced into bankruptcy [12]. Our investigation is made more in-depth and complicated as a result of the fact that it consists of 95 different financial characteristics.

The Polish company dataset, which is the subject of our third and current investigation, stands out as the most complicated of the datasets that we've looked at. This specific dataset has a sizeable 10,000 samples, which provides a plethora of information for our investigation. The vast majority of the samples in this massive dataset, totaling 9797, represent profitable businesses [13]. Only 2.03% of the samples relate to organizations involved in a bankruptcy proceeding. Because of the difficulty of this dataset as well as the rarity of bankruptcy cases, it constitutes a severe obstacle for our study as well as an essential addition to our existing inventory.

3 Classification Algorithm Compared

An essential part of the investigation that we are doing is an in-depth analysis and analysis of the many categorization methods that are available. These algorithms are the foundation of our prediction models, and they enable us to determine the chance that a certain business would go bankrupt with their assistance. In the following, we will explain the most important classification methods that we have used as well as the reasoning for comparing them:

3.1 DBN

This model has garnered attention for its unique architecture and capabilities in capturing complex patterns in data. In Fig. 1, we illustrate the structure of this model, which comprises multiple layers of RBMs. In this configuration, the lower RBMs constitute the hidden layers, while the upper RBMs form the visible layer. Notably, there are undirected links connecting the two layers, creating a sophisticated network for information processing [14]. The training of this model follows a greedy approach, with each RBM being trained in an unsupervised manner independently. This means that the model progressively learns and refines its representations layer by layer, with each RBM capturing distinct data features.

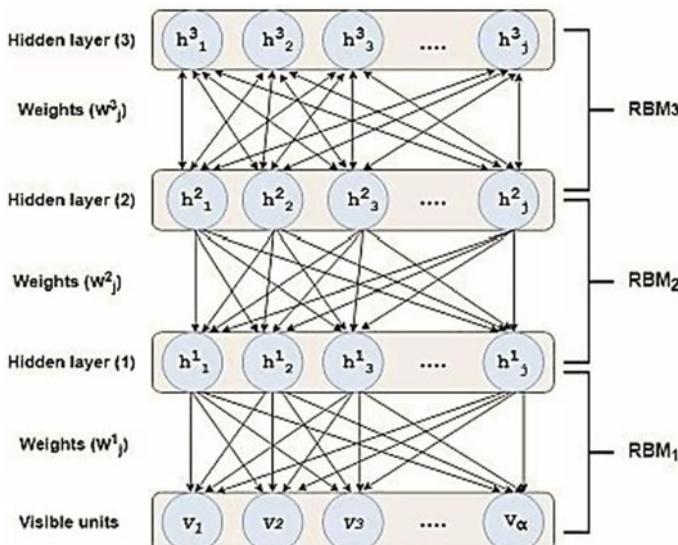


Fig. 1 Architecture of DBN

Within the hidden layers of this model, we find a binary randomized vector represented as ‘ hi ,’ with individual elements labeled as ‘ h_j^i .’ These elements correspond to the activations within the hidden layers and play a crucial role in encoding and extracting valuable information from the input data. Hinon et al.’s model represents an innovative and effective approach in the field of machine learning, particularly for tasks requiring hierarchical feature extraction and representation learning. Its stacked RBM architecture, combined with the unsupervised training strategy, offers promising prospects for various applications, including data analysis and pattern recognition.

$$(V, h^1, \dots, h^P) = (V|h^1) \underline{P}(h^1|h^2) \cdot P(h^P - 1|h^P) \quad (1)$$

$$(h^i|h^{i+1}) = \prod_{j=1}^{ni} (h_j^i|h^{i+1}) \quad (2)$$

With h_j^i as stochastic unit and equation first as activation, equation third will represent the probability of various hidden units.

$$(h^1 = 1|h^{1+1}) = \text{sigm}_j(b^1 + \sum_{k=1}^{n+1} W_{jk}^i h_k^{1+1}) \quad (3)$$

$$\text{sigm}(x) = \frac{1}{1 + \exp(-x)} \quad (4)$$

Equation fourth will represent the activation of a function.

3.2 LSTM

The cell that makes up this memory serves as the primary functional unit. It was designed to take the role of the RNN’s buried layer of neurons and is comprised of primarily three gated components (input, output, and forget gate), as shown in Fig. 2. Its design contributes to the process of determining whether to forget the most recent concealed state or to refresh it [15].

The following six equations explain the various stages of operation for this kind of memory:

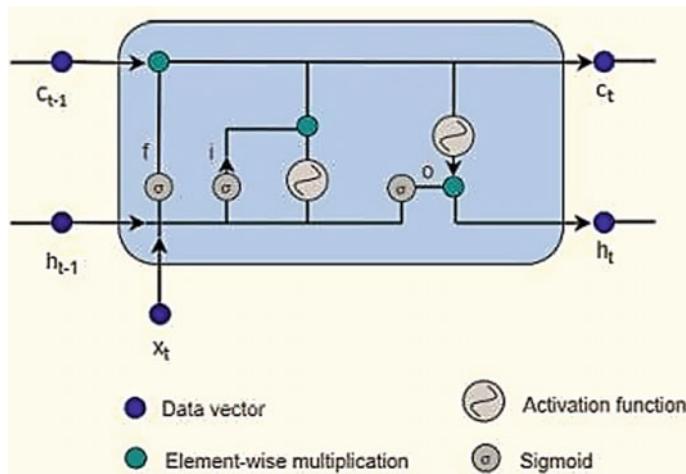


Fig. 2 Architecture of LSTM

$$\begin{aligned}
 f_t &= (W_f x_t + U_f h_{t-1} + b_f) \\
 i_t &= (W_i x_t + U_i h_{t-1} + b_i) \\
 \tilde{c}_t &= \tanh(W_c + U_c h_{t-1} + b_c) \\
 c_t &= i_t \odot \tilde{c}_t + f_t \odot c_{t-1} \\
 o_t &= (W_o x_t + U_o h_{t-1} + V_o c_t + b_o) \\
 h_t &= o_t \odot \tanh(c_t)
 \end{aligned} \tag{5}$$

3.3 MLP-6l

The majority of it is made up of three layers. In this feed-forward network, the first layer is the input layer, followed by the hidden layer, and finally the output layer. There are a total of four hidden layers in between the input and output layers [16]. The majority of the time, it is used for supervised learning. It includes the framework of a network that is completely linked to all other nodes, as shown in Fig. 3. In its whole, it constitutes all six layers.

3.4 Random Forest (RF)

Bermin is credited with the development of the classification algorithm known as the random forest algorithm. This method that we have developed consists of many

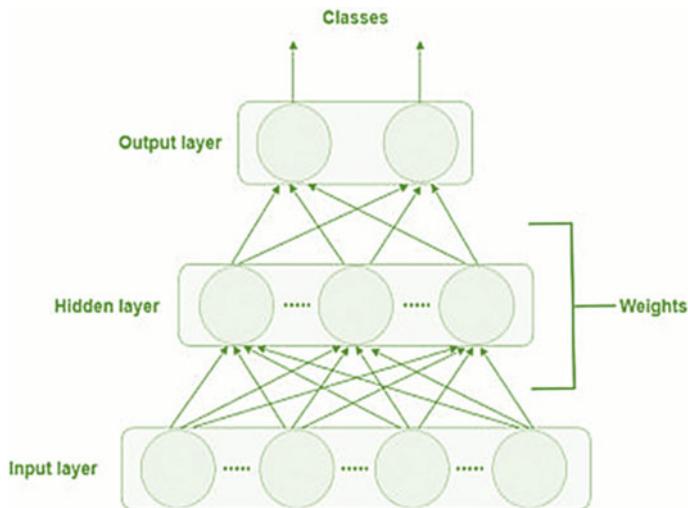


Fig. 3 Architecture of MLP-6L

sets of decision trees, and its purpose is to locate the optimal output for a variety of datasets [17]. The bootstrapping method is utilized to generate such datasets, and on the basis of those datasets, multiple decision trees are generated by the system. Figure 4 provides a more detailed description of the RF design.

4 Development Setup

Since the purpose of our study is to forecast the likelihood of bankruptcy filings by a variety of businesses, we are concentrating more on the evaluation metric than on the accuracy of our predictions in this study. In this study article, we made use of a variety of different datasets. While we were working on those datasets, we concluded that the data needs to be more balanced due to the fact that the bankrupted data makes up the vast bulk of the dataset. The inconsistency of the data was improved by using the many different enhancement approaches that were suggested after we had applied the various procedures [18]. In this particular instance, we made use of a number of different advanced strategies for resampling data. The primary purpose of using them is to improve the classification performance by raising the occurrences of the minority class while lowering the instances of the majority class.

In Table 2, the data is shown both before and after being subjected to several resampling techniques. As can be shown in Table 2, using a variety of resampling methods, balanced datasets were produced. These datasets include virtually an equal number of members from the majority and minority groups. When compared to other

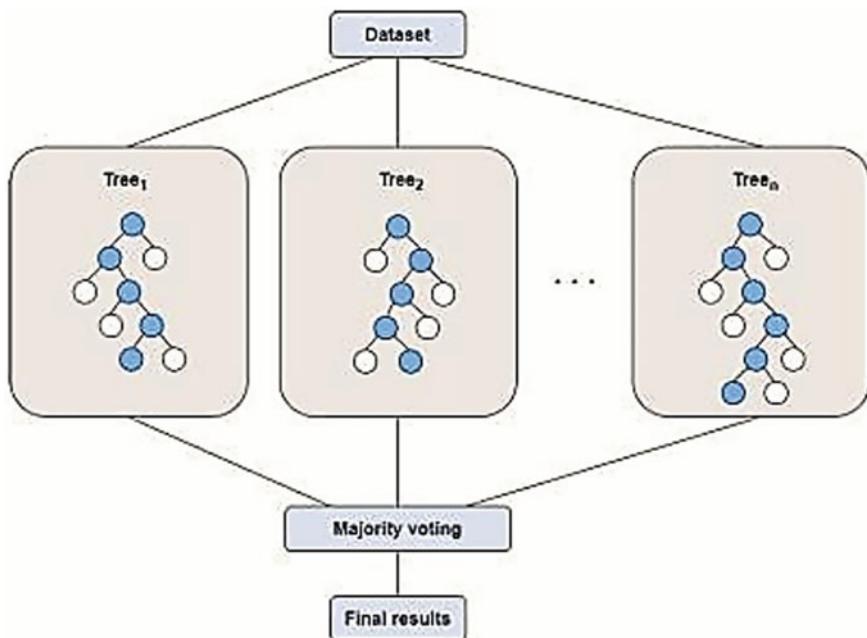


Fig. 4 Architecture of random forest

sampling approaches, ADASYN has successfully handled large minority-class populations. On the other hand, SVM-SMOTE produced data that was less balanced but was judged to be more realistic due to the fact that it created fewer fake data points. Using hybrid methods such as SMOTE-ENN and SMOTE-Tomek, the researchers first oversampled the data from the minority, and then they under-sampled the data from the majority to bring it up to the level of the minority data variation [19]. Tomek was previously responsible for removing connections from the dataset, but now ENN assists in the cleaning of deeper data than SMOTE-Tomek can do alone. Therefore, with sufficient oversampling, the impact will become apparent. In conclusion, we utilize SMOTE-NC to build the dedicated balancing datasets. These datasets comprise nominal and continuous characteristics that are supplied in the previously described datasets.

5 Result

The extensive analysis that we carried out on the datasets from Spain, Taiwan, and Poland resulted in the discovery of illuminating discoveries, which are outlined in Fig. 5. Following the application of a number of distinct algorithms to these datasets, we have been able to recognize useful patterns and distinguish between varying levels

Table 2 Before and after applying sampling technique

	Balancing techniques	Spanish companies' data		Taiwanese companies' data		Polish companies' data	
		Bankrupt	Solvent	Bankrupt	Solvent	Bankrupt	Solvent
	Original dataset	75	2945	312	7451	353	11,201
Oversampling	SVM-SMOTE	2144	2011	4217	5496	9894	9865
	ADASYN	2546	2352	6645	6541	6215	9541
	BL-SMOTE	1542	2145	2542	7145	3945	8976
	SMOTE	2415	2415	5214	5624	8768	9874
	SMOTE-NC	2366	2325	–	–	–	–
Clustering-based oversampling	K-means SMOTE	2354	1254	5694	6542	5746	8976
Oversampling under-sampling	SMOTE-Tomek	2564	1987	5946	6547	7954	7454
	SMOTE-ENN	1256	1652	5679	4565	6521	7541

of algorithmic performance. It is important to highlight that the deep belief network (DBN) has, on average and across all datasets, showed considerably lower levels of performance. This discovery highlights how important it is to investigate various algorithms for certain jobs that fall within the realm of financial status prediction. There have been a number of algorithms that have shown remarkable accuracy in forecasting the current financial situation of a variety of businesses. Notable among them are the multi-layer perceptron with six layers (MLP-6L), long short-term memory (LSTM), random forest (RF), and XGBoost, all of which have repeatedly shown excellent results and attained metric values that are more than 0.9899. These algorithms have shown their superiority in terms of their ability to process complicated data and derive relevant conclusions. Support vector machine (SVM), K-nearest neighbors (KNN), and AdaBoost are all examples of ensemble algorithms that have shown outstanding performance. These algorithms have regularly attained metric values higher than 0.9599, which demonstrates their dependability in performing tasks involving the prediction of financial position.

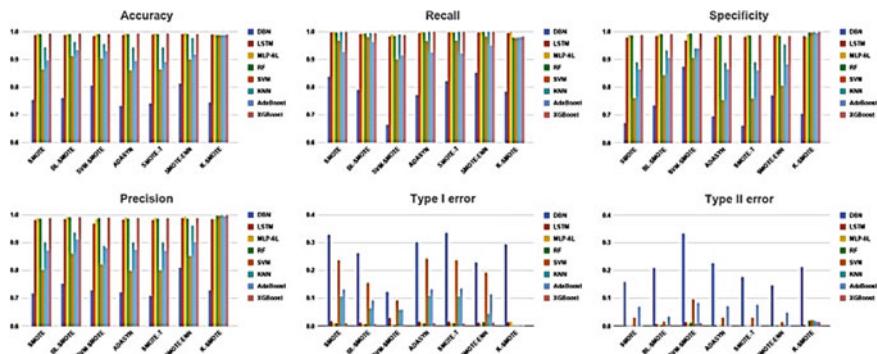


Fig. 5 Result of accuracy, recall, and specificity

6 Conclusion

Our research endeavors have resulted in a conclusive finding, which we arrived at after conducting an exhaustive analysis of a wide variety of algorithms and giving serious consideration to a number of different methods for resampling data. The multi-layer perceptron with six layers (MLP-6L), coupled with the SMOTE-ENN resampling algorithm, and is an outstanding performance because it demonstrates remarkable efficiency in terms of both accuracy and recall. The MLP-6L is the winner of this competition. This potent combination, when used together, reliably produces the highest metric values of all of the algorithms that we have investigated. The selection of MLP-6L, which was strengthened by the SMOTE-ENN resampling approach, reflects the excellent prediction powers it has in determining the current financial situation of businesses. The amazing accuracy and recall that this model was able to accomplish lend credence to its dependability and efficiency in managing un-balanced datasets. This issue is typical in real-world financial research. This conclusion functions as a helpful guideline for many stakeholders who are looking for optimum solutions in the context of financial stability analysis. The MLP-6L with SMOTE-ENN emerges as an appealing option, positioned to give strong results and aid decision-makers in making educated judgments about the financial well-being of firms. This combination is poised to deliver robust results and help decision-makers make informed judgments regarding the financial well-being of organizations.

Acknowledgements The generous backing for this research came in the form of an Research Support Fund (RSF) Grant from Symbiosis International (Deemed University), Pune, India.

References

1. Vohnout R et al (2023) Living lab long-term sustainability in hybrid access positive energy districts-a Prosumager smart fog computing perspective. In: IEEE internet of things journal. <https://doi.org/10.1109/JIOT.2023.3280594>
2. Yu X, Li W, Zhou X et al (2023) Deep learning personalized recommendation-based construction method of hybrid blockchain model. Sci Rep 13:17915. <https://doi.org/10.1038/s41598-023-39564-x>
3. Lin WY, Hu YH, Tsai CF (2012) Machine learning in financial crisis prediction: a survey. IEEE Trans Syst Man Cybern Part C Appl Rev 42(4):421–436
4. Du Y et al (2023) The research on prediction for financial distress in car company listed combining financial indicators and text data. In: Lecture notes in electrical engineering
5. Goyal B et al (2023) Detection of fake accounts on social media using multimodal data with deep learning. In: IEEE transactions on computational social systems. <https://doi.org/10.1109/TCSS.2023.3296837>
6. Praveen Malik S, Sharma R, Ghosh U, Alnumay WS (2023) Internet of things and long-range antenna's; challenges, solutions and comparison in next generation systems. Microprocess Microsyst 104934:ISSN 0141-9331.<https://doi.org/10.1016/j.micpro.2023.104934>
7. Da Z, Engelberg J, Gao P (2015) The sum of all FEARS investor sentiment and asset prices. Rev Financ Stud 28(1):1–32
8. Hernandez Tinoco M, Wilson N (2013) Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. Int Re v Financ Anal 30:394–419
9. Elhoseny M et al (2022) Deep learning-based model for financial distress prediction. Annal Operat Res
10. Gicic A, Donko D (2023) Proposal of a model for credit risk prediction based on deep learning methods and SMOTE techniques for imbalanced dataset. In: 2023 29th international conference on information, communication and automation technologies, ICAT 2023—proceedings
11. Alkhoshi E, Belkasim S (2018) Stable stock market prediction using NARX algorithm. In: ACM international conference proceeding series
12. Soui M et al (2020) Bankruptcy prediction using stacked auto-encoders. Appl Artif Intell 34(1):80–100
13. Jo H, Han I, Lee H (1997) Bankruptcy prediction using case-based reasoning, neural net- works, and discriminant analysis. Expert Syst Appl 13(2):97–108
14. Priyadarshini I, Kumar R, Alkhayyat A, Sharma R, Yadav K, Lulwah MA, Sachin K (2023) Survivability of industrial internet of things using machine learning and smart contracts. Comput Electric Eng 107:108617, ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108617>
15. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma RK (2023) SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. Trans Emerg Tel Tech e4758. <https://doi.org/10.1002/ett.4758>
16. Rohit S, Rajeev A (2023) Secured mobile IOT ecosystem using enhanced multi-level intelligent trust scheme. Comput Elect Eng 108:108715, ISSN 0045–7906. <https://doi.org/10.1016/j.compeleceng.2023.108715>
17. Haitao D, Jing H, Rohit S, Mingsen M, Yongjun R (2023) NVAS: a non-interactive verifiable federated learning aggregation scheme for COVID-19 based on game theory. Comput Commun ISSN 0140–3664. <https://doi.org/10.1016/j.comcom.2023.04.026>
18. Sharma A, Rani S, Shah SH, Sharma R, Yu F, Hassan MM (2023) An efficient hybrid deep learning model for denial of service detection in cyber physical systems. In: IEEE transactions on network science and engineering. <https://doi.org/10.1109/TNSE.2023.3273301>
19. Gupta U, Sharma R (2023) Analysis of criminal spatial events in India using exploratory data analysis and regression. Comput Electric Eng 109(Part A):108761, ISSN 0045-7906.<https://doi.org/10.1016/j.compeleceng.2023.108761>

An Analytical Study of Improved Machine Learning Approaches for Predicting Mode of Delivery



Vaishali Bhargava and Sharvan Kumar Garg

Abstract Machine learning approaches came about as a game-changer in modern healthcare, leading to more reliable medical predictions and enhanced patient care. Predicting the way of delivery during labor is critical to protecting mother and newborn health. This study offers a thorough comparison of machine learning (ML) approaches aiming at predicting an optimal mode of delivery. The efficacy of enhanced ML algorithms is evaluated in improving prediction accuracy using a dataset containing maternal health indicators. This study compares the effectiveness of five distinct machine learning approaches: J48, Logistic Model Trees, Random Forests, Random Tree, and Multilayer Perceptron. We analyze their prediction capabilities and applicability for the task at hand through a thorough experimental procedure. Our findings show that different advanced ML techniques have varying degrees' effectiveness in forecasting the way of delivery. The performance parameters under consideration are accuracy, precision, and recall. Among all the performance metrics considered, J48 exhibited the most favorable performance.

Keywords Machine learning · Cesarean section · Classification · J48 · Random Forest

1 Introduction

Health care has undergone a tremendous transition in recent years, owing to breakthroughs in machine learning (ML) along with data analytics. Prediction of medical outcomes is one crucial area that has benefited from technological advances, leading to better care for patients as well as better informed decision-making by medical professionals. Predicting the type of childbirth, a difficult medical decision that has a considerable impact on both mother and newborn health, is an essential aspect of this evolution.

V. Bhargava · S. K. Garg
Swami Vivekanand Subharti University, Meerut, UP 250005, India
e-mail: bhargavaavaishali23@gmail.com

The type of delivery, whether vaginal or cesarean, is crucial in determining the health of both the expecting mother and infant. Effective prediction of the best way to deliver is critical for optimizing healthcare resources, minimizing problems, and assuring a safe delivery process. While traditional clinical methods depend on clinical expertise and prior information, incorporating advanced machine learning technologies offers the potential to improve forecast accuracy as well as offer more personalized suggestions.

The purpose of the research is to inspect and evaluate the impact of enhanced ML methods in forecasting the delivery mode. We hope to overcome the restrictions of existing models for prediction and contribute to the improvement of clinical decision-making procedures by exploiting varied datasets and employing algorithms that are innovative.

2 Literature Review

An ensemble learning approach is adopted by researchers [1], implementing five distinct machine learning algorithms—J48, Random Forests (RFs), AdaBoosting of decision trees (ADA-B), Gradient Boosting, and Decorate. These algorithms are implemented within the Knime analytics platform to accomplish the task of classifying patients into two delivery categories: vaginal or cesarean section. The Random Forest (RF) yielded the most favorable outcomes, achieving an accuracy rate of 91.1%, a sensitivity of 90.0%, and an impressive AUCROC value of 96.7%.

The objective of the research [2] was to pinpoint the key factors that impact the standard of care during emergency cesarean sections (ECSs) in relation to the time interval between arrival and delivery. Emergency cesarean sections are critical procedures with time-sensitive requirements. While various factors can impact the efficiency of the medical team, their specific significance remains unclear. The dataset comprised records from 2409 patients who underwent ECS procedures. Among the range of predictors considered, those related to the qualifications and experience of the medical team emerged as the foremost determinants of the arrival-to-delivery interval across all categories of ECS emergencies.

The study [3] compared the performance of Gaussian Naive Bayes (GNB), linear discriminant analysis (LDA), K-nearest neighbors (KNN), gradient boosting classifier (GBC), and logistic regression (LR). The ADAdaptive SYNthetic (ADASYN) method was utilized to balance the model, and the suggested HGSORF was compared to other classifiers and studies. HGSORF outperformed the competition with an accuracy rate of 98.33% on the PDHS dataset.

Research work [4] focuses at the demographic, antenatal care, geographical, and socioeconomic factors that are connected with the cesarean delivery procedure. To achieve this goal, information gathered through the National Family Health Survey was employed. According to the outputs of logistic regression models incorporating mother age, birth order, current age, and births in health care organizations, and regional inequalities are all strongly related to cesarean section deliveries in Kerala.

In comparison to their younger equivalents, older cohorts of women were shown to be at a higher risk of having a cesarean section.

The aim of this research [5] was to determine the factors affecting the decision of women to opt for elective cesarean sections. To achieve this objective, a descriptive cross-sectional survey approach was employed. Data were gathered from a sample of 78 women aged 18 years and older who had undergone cesarean sections. A comprehensive survey was conducted, involving a whole population sampling technique, where researchers visited respondents at their residences and administered a pretested questionnaire.

Research paper [6] illustrates the utilization of machine learning in conjunction with fetal heart rate signals to offer direct insights into the fetal condition. It serves as a means to mitigate the subjective judgments of medical professionals when employed as a decision support tool. The primary objective is to establish a proof-of-concept that showcases the practical application of machine learning in objectively identifying instances where medical intervention, such as a cesarean section, is warranted. By doing so, it aims to contribute to the prevention of avoidable prenatal deaths.

The research paper [7] examines whether there are valid reasons for imposing restrictions on Maternal Request Cesarean Sections (MRCSs). Childbirth holds immense physical and emotional importance in a woman's life, and it is crucial to acknowledge that women's preferences should not be disregarded. We need to assess whether depriving women of their agency is justified. To address this inquiry, the author first illustrates that obtaining MRCS in the UK is subjected to a degree of randomness. Additionally, the author contends that pregnancy does not inherently override the ethical principle of respecting patients' autonomous decisions, suggesting that there is no exceptional circumstance that justifies doing so.

The aim of this systematic review [8] is to outline the enduring advantages and disadvantages of cesarean delivery concerning the well-being of the mother, the baby, and subsequent pregnancies. Primary focus regarding maternal outcomes was pelvic floor dysfunction, while for infant outcomes, it was asthma. In comparison to vaginal delivery, cesarean delivery demonstrates a reduced occurrence of urinary incontinence and pelvic organ prolapse. However, it is essential to consider these benefits in light of the heightened risks it poses to fertility, future pregnancies, and the long-term health of children. This knowledge can serve as valuable information for counseling women on their choice of delivery method.

3 Proposed Model

We considered the cesarean dataset and preprocessed it. We train the data after the preprocessing steps are completed. On the input dataset, many machine learning approaches including J48, Logistic Model Trees, Random Forest, Random trees, and Multilayer Perceptron are used. Figure given below depicts the flow of work (see Fig. 1).

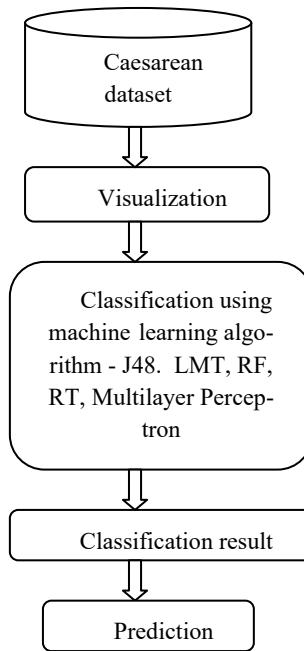


Fig. 1 Experimental framework

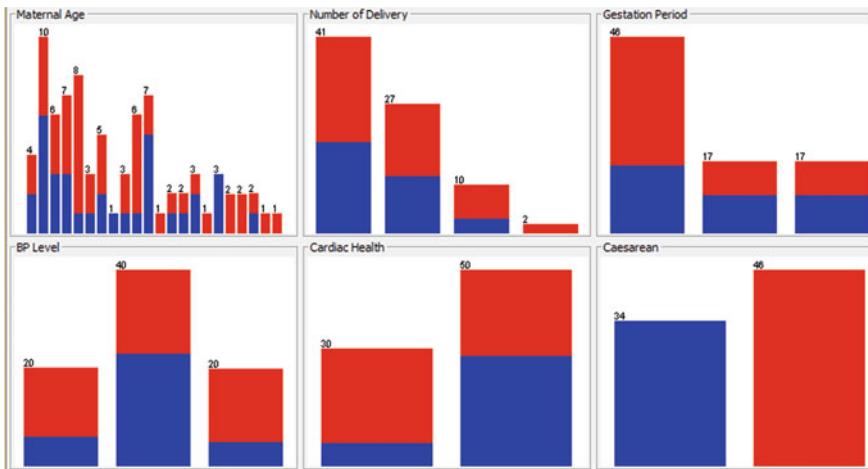


Fig. 2 Visualization of the features

3.1 Dataset

The experiment utilized a cesarean dataset comprising 80 instances and 6 attributes. This dataset was sourced from the UCI repository. Attributes include Maternal Age, Number of Delivery, Gestation Period, BP Level, Cardiac Health, and Cesarean. All the attributes are of nominal type.

3.2 Data Visualization

Data visualization is an important part of the data analysis procedure since it helps us to visually obtain insights, examine patterns, and comprehend the features of datasets. These visualizations help us in making sensible decision concerning preprocessing, selection of features, and building a model (Fig. 2).

4 Description of Algorithm Used

4.1 J48

The main principle behind J48 is to recursively divide the training data based on the attributes to create the most informative splits. A feature that significantly reduces entropy or impurities in a dataset is identified using the idea of information gain. The information gain of each attribute is determined, and the attribute with the greatest amount of information gain is decided to be the splitting attribute. Until a halting requirement is satisfied, this process iterates for each child node. After the decision tree is generated, it may be utilized for classification of unidentified instances by traversing through the tree's nodes from root to leaf according to the instances' attribute values. The leaf node's class label can be considered a representation of the instance's predicted class. J48 is renowned for being simple to comprehend and capable of handling both categorical as well as numerical attributes. It has been widely used in various domains due to its effectiveness in classification tasks.

4.2 Logistic Model Trees

LMTs combine the decision tree's hierarchical structure along with the logistical regression model. Recursive decision trees are constructed, with nodes representing feature-based decisions and leaves representing class labels. A logistic regression model is connected with each leaf of the decision tree, which analyzes the class probability for instances that fall into that leaf. LMTs preserve decision tree interpretability.

The tree's hierarchical structure facilitates visualization and comprehension of the process of decision-making. Furthermore, the models using logistic regression at the leaves can reveal how particular features contribute to class probabilities.

4.3 Random Forest

The fundamental concept behind the Random Forest algorithm is to construct a "forest" of decision-making trees, whereas each of the trees is learned on a randomly selected subset of the model training data and employs a random selection of features. This randomness introduces diversity, which, in turn, helps in reducing over-fitting, leading to enhanced prediction accuracy and robustness. All in all, the Random Forest algorithm is a potent and influential method that leverages the collective wisdom of multiple decision trees to achieve accurate and reliable predictions. It is widely used in various domains, including health care, finance, and image recognition.

4.4 Random Tree

The "random" part of random trees stems from unpredictability being introduced throughout both the selection of training data and the attribute selection for particular trees in the ensemble. Randomized subsets of the initial training data have been utilized for training each decision tree, and a random selection of characteristics is examined for splits at each node of the tree. Once all of the particular decision trees have been trained, they can predict new data points. The Random Forest generates predictions in classification problems using majority voting, whereas it averages expected values in regression tasks.

4.5 Multilayer Perceptron

The Multilayer Perceptron (MLP) is foundational artificial neural network (ANN) architecture; this has a widespread application in the realm of machine learning applications, notably those incorporating pattern recognition, classification, as well as regression. Feedforward propagation is the mechanism by which information flows in an MLP. The network passes the input features through it, with the biases and weights of each neuron affecting the strength of the signal as it passes from layer by layer. After that, the activation function generates an output associated with each neuron. The existence of hidden layers enables MLPs to learn hierarchical data features and representations, allowing them to understand complicated relationships that might otherwise be missed by simpler models.

5 Results and Discussion

The performance with regard to accuracy is displayed below (see Fig. 3) and also comprehensive overview of various metrics including precision, recall, F-measure, and kappa for the implemented algorithms is represented (see Table 1).

J48 demonstrates the highest accuracy rate compared to all other algorithms (see Fig. 3). Accuracy is a standard measure of performance for assessing the efficacy of an algorithm. It computes the proportion of accurately classified instances in the dataset with respect to the total number of instances. It provides a broader view of how well the model is doing.

Above table further underscores that J48 attains the highest values for precision, recall, F-measure, and kappa when compared to all other methods (see Table 1).

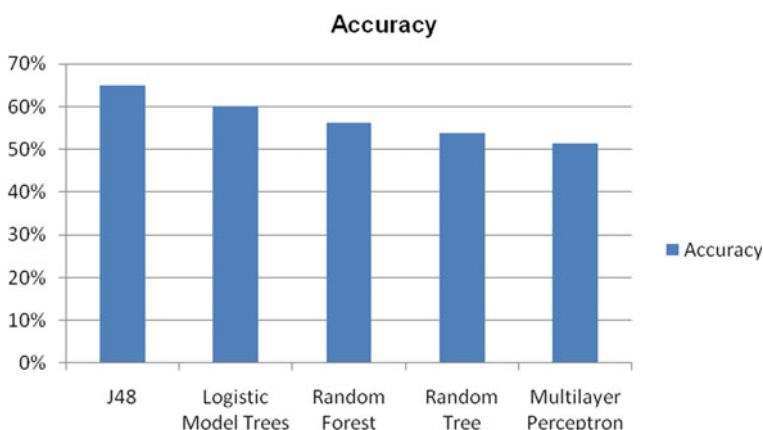


Fig. 3 Accuracy comparison of different classification algorithms

Table 1 Results generated for different algorithms

	J48	Logistic Model Trees	Random Forest	Random Tree	Multilayer Perceptron
Precision	0.698	0.597	0.568	0.536	0.518
Recall	0.65	0.6	0.563	0.538	0.513
F-Measure	0.646	0.598	0.564	0.537	0.515
Kappa statistic	0.3253	0.1753	0.115	0.0501	0.0139

6 Conclusion

This study explored advanced machine learning algorithms to predict the way of delivery during labor. The major goal was to determine the effectiveness of several different advanced algorithms in this important healthcare application. Through a thorough comparison of five different methods, viz. J48, Logistic Model Trees, Random Forests, Random Tree, and Multilayer Perceptron, we learned a lot about their prediction abilities. The findings of our study demonstrated the importance of using machine learning approaches to forecast the way of delivery. Notably, J48 has proved to be the most effective algorithm, with the highest accuracy, precision, recall, F-measure, and kappa values among all techniques. This demonstrates the utility of models based on decision tree for healthcare decision-making, particularly in critical tasks such as birthing delivery mode prediction.

While J48 was the best performer, it is important to point out that various algorithms succeed in various circumstances. Future study can look into combining techniques that leverage the capabilities of several models for more accurate predictions as the area of machine learning evolves. In essence, this work bridges the divide between advancements in machine learning and obstetric healthcare improvements and obstetric care, revealing the potential of complex algorithms in improving medical decision-making quality. The results of our research contribute to a growing repository of insights that provides ways for safer and more personalized delivery experiences, highlighting the potential for emerging technologies to impact the future of maternal and infant health care.

7 Future Scope and Summary

Future study can look into how existing machine learning models can be seamlessly integrated with electronic medical record systems. This would allow healthcare providers to have real-time access to prediction tools, improving decision-making during labor and delivery. To improve accuracy of predictions, consider using other data sources that include continuous monitoring data, imaging, and genetic information. Using multimodal data and enhanced fusion technologies to offer a broader understanding of the state of the patient could give a more comprehensive perspective of the health of the individual.

Given the importance of childbirth decisions, the use of comprehensible AI techniques is crucial. Future research should concentrate on establishing interpretable models that may provide clear insights into the elements influencing mode of delivery predictions. This can increase trust between healthcare providers and patients.

Predictive models that are tailored to particular patient profiles could be a potential route. Research can be conducted to develop personalized prediction models that take into account the medical records of a person, demographics, and genetic predispositions, resulting in more precise recommendations. Continued research and

innovation in such domains will help to develop machine learning algorithms for predicting mode of delivery, boosting maternal and neonatal healthcare experiences in the long run.

Our primary objective was to assess the effectiveness of various advanced algorithms in this critical healthcare application. We conducted a comprehensive comparison of five different methods, namely J48, Logistic Model Trees, Random Forests, Random Tree, and Multilayer Perceptron, to evaluate their predictive capabilities. Our findings underscored the significance of employing machine learning approaches to forecast the mode of delivery. Notably, J48 has emerged as the most effective algorithm, demonstrating superior accuracy, precision, recall, F-measure, and kappa values compared to all other techniques. While J48 excelled as the top performer, it is essential to acknowledge that different algorithms excel under varying circumstances. Future research can explore the synergistic potential of combining techniques that harness the strengths of multiple models to achieve even more accurate predictions, reflecting the dynamic nature of machine learning advancements.

In essence, our work bridges the gap between advancements in machine learning and enhancements in obstetric health care, showcasing the potential of sophisticated algorithms to elevate the quality of medical decision-making. The results of our research contribute to a growing body of knowledge, offering pathways to safer and more personalized delivery experiences. This work emphasizes the transformative potential of emerging technologies in shaping the future of maternal and infant healthcare.

References

1. Ricciardi C, Imrota G, Amato F, Cesarelli G, Romano M (2020) Classifying the type of delivery from cardiotocographic signals: a machine learning approach. *Comput Methods Programs Biomed* 196:105712
2. Andersen BR, Ammitzbøll I, Hinrich J, Lehmann S, Ringsted CV, Løkkegaard ECL, Tolsgaard MG (2022) Using machine learning to identify quality-of-care predictors for emergency caesarean sections: a retrospective cohort study. *BMJ Open* 12(3):e049046
3. Islam MS, Awal MA, Laboni JN, Pinki FT, Karmokar S, Mumnenin KM, Mirjalili S (2022) HGSORF: henry gas solubility optimization-based random forest for c-section prediction and XAI-based cause analysis. *Comput Biol Med* 147:105671
4. Sabu SP, Suresh Kumar S, Sajini BN, Anitha Kumari KR (2000) Caesarean section delivery in Kerala, India: evidence from a National Family Health Survey. *Soc Sci Med* 51(4):511–521. [https://doi.org/10.1016/S0277-9536\(99\)00491-8](https://doi.org/10.1016/S0277-9536(99)00491-8)
5. Diema Konlan K, Baku EK, Japiong M, Dodam Konlan K, Amoah RM (2019) Reasons for women's choice of elective caesarian section in Duayaw Nkwanta Hospital. *J Preg*
6. Fergus P, Hussain A, Al-Jumeily D, Huang DS, Bouguila N (2017) Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. *Biomed Eng Online* 16(1):1–26
7. Romanis EC (2019) Why the elective caesarean lottery is ethically impermissible. *Health Care Anal* 27(4):249–268
8. Keag OE, Norman JE, Stock SJ (2018) Long-term risks and benefits associated with cesarean delivery for mother, baby, and subsequent pregnancies: systematic review and meta-analysis. *PLoS Med* 15(1):e1002494

Comparative Study of Different Document Similarity Measures and Models



Anshika Singh and Sharvan Kumar Garg

Abstract Document similarity refers to an approach of measuring how two or more documents look alike in terms of their content or structure. Document similarity algorithms are used to determine the degree of resemblance or relatedness between various documents. Document similarity plays a pivotal role in a wide range of tasks involving natural language processing, information retrieval, recommender systems and duplicates detection. In this paper, we will be studying and compare the similarity score of documents using different document similarity measures and models like cosine similarity, Euclidean distance, Jaccard similarity, Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), Bidirectional Encoder Representations from Transformers (BERTs), etc.

Keywords Cosine similarity · Jaccard similarity · Euclidean distance · LSA · LDA · BERT · RoBERTa

1 Introduction

Text similarity refers to the assessment of how closely two or more different texts (words, phrases, or documents) resemble each other. Identifying the resemblance between words constitutes a fundamental aspect of textual similarity, serving as a crucial initial step in assessing the likeness between sentences, paragraphs, and documents [1]. It enables us to quantify the degree to which documents, phrases, or even individual words are similar. We can find patterns, detect duplicates, and group similar documents together by comparing texts, which leads to a wide range of applications in different fields like information retrieval, natural language processing, recommendation systems, duplicates detection, etc. Document similarity algorithms implement an approach of text similarity to measure how two or more documents look alike in

A. Singh (✉) · S. K. Garg

Department of Computer Science and Engineering, Subharti Institute of Technology and Engineering, Swami Vivekanand Subharti University, Meerut, India

e-mail: anshika.research@gmail.com

terms of their content or structure. They are used to determine the degree of resemblance or relatedness between various documents. They play a pivotal role in various areas like Information Retrieval, Text Classification, Text Summarization, Natural Language Processing, Duplicate Detection, Plagiarism Detection, Recommendation Systems, Machine Translation, Question Answering Document Clustering and Topic Modeling, Content Recommendation and personalization, etc. [1].

Search engines rely on document similarity to provide relevant search results. When a user inputs a query for search, the search engine compares the query to the documents in its database and ranks them based on their similarity. The most similar documents are then presented as search results. By improving the accuracy of document similarity measures, search engines can greatly enhance their effectiveness.

In information retrieval, there is a frequent need to categorize or classify documents into specific categories. This process involves assigning predefined labels to documents based on their content. For example, news articles can be categorized into topics such as politics, sports, or entertainment. One way to determine the category of a document is by measuring its similarity to previously categorized documents. By comparing the content and characteristics of different documents, we can determine which category an uncategorized document belongs to.

Detecting duplicate or similar documents is crucial for ensuring the accuracy and reliability of databases and search engine indexes. By determining document similarity, we can effectively identify and eliminate duplicate content. Document similarity is used in abstractive and extractive text summarization techniques to identify sentences or passages that are most similar to the main content of a document. This helps in generating concise and informative summaries. In machine translation, document similarity can be used to align source and target language sentences or phrases. Similar sentences in different languages can help to improve translation accuracy. Document similarity is essential in question answering systems, where the system must identify relevant passages or documents that contain answers to a user's query.

Document similarity is used to group similar documents into clusters in unsupervised learning tasks. It is also utilized in topic modeling algorithms like Latent Dirichlet Allocation (LDA) and Latent Semantic Analysis (LSA) to discover hidden semantic structures in large text corpora. In academia and content publishing, it is crucial to identify instances of plagiarism where one document closely resembles or copies portions of another. Document similarity is a fundamental tool in plagiarism detection systems.

Content-based recommendation systems use the similarity between documents to suggest items such as articles, products, or movies based on a user's preferences. These systems recommend documents that are similar to ones the user has previously engaged with. Collaborative filtering is another application of document similarity, where users with similar preferences or behavior patterns are identified. By analyzing the documents that these users have interacted with, personalized recommendations can be made based on the preferences of users who share similar tastes. In online platforms, such as e-commerce websites and news portals, understanding

document similarity helps in recommending similar products, articles, or content to users. Personalized recommendations not only enhance user engagement but also user satisfaction.

In this paper, we will be studying and compare the similarity score of documents using different document similarity measures and models like cosine similarity, Jaccard similarity, Euclidean distance, Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), Bidirectional Encoder Representations from Transformers (BERT), RoBERTa, etc.

2 Different Document Similarity Measures and Models

There are three major similarity measures:

Cosine similarity, Jaccard similarity, and Euclidean distance.

2.1 Cosine Similarity

Cosine similarity is a widespread measure for determining the similarity between two vectors, which is frequently employed in the context of document similarity or text analysis. The computation of the cosine of angle between two vectors yields the cosine value, which indicates the degree of similarity. In the realm of document similarity, each individual document is portrayed as a vector within a high-dimensional space, and cosine similarity assesses how similar the direction of these vectors is, regardless of their size [2].

Before applying cosine similarity, each document needs to become a numerical vector representation. During the vectorization process, techniques including TF-IDF and word embeddings like Word2Vec or BERT change textual data into quantitative feature vectors. Each word in the corpus has a corresponding dimension in the vector space. Point values in a multidimensional space characterize how well terms align with their relevance levels for individual documents.

The cosine similarity is determined by calculating the cosine of the angle between two vectors, which represent the documents. In mathematical terms, the cosine similarity between vectors A and B is computed by evaluating:

$$\text{Cosine similarity } (A, B) = (A \cdot B) / (\|A\| * \|B\|),$$

where $A \cdot B$ represents the dot product of vectors A and B.

$\|A\|$ and $\|B\|$ represent the magnitudes (Euclidean norms) of vectors A and B.

The cosine similarity ranges between -1 and 1, where:

A value 1 indicates, the vectors are parallel in direction.

A value 0 means the vectors have no similarity.

When opposing vectors are present, it indicates perfect dissimilarity (1).

High similarity scores (closely approaching 1) suggest the same content across document vectors and alignment in the vector space. With a close to -1 or 0 cosine

similarity score, dissimilarity implies. Despite document length, cosine similarity keeps its effectiveness since it concentrates on directing vectors instead of their intensity. With high-dimensional data, efficiently computational and works well. Though document structure and language particularities may have an impact on cosine similarity analyses because they ignore word order and semantic meaning. Cosine similarity is used in various applications, including information retrieval, document ranking, document clustering [3], recommendation systems like job recommendation system [4], natural language processing like online essay assessment [5], etc.

2.2 Jaccard Similarity

Comparing their intersection and union allows us to calculate the Jaccard similarity metric, which measures the similarity of two sets. Especially relevant to document similarity analysis is its usefulness when working with text data or datasets where word/term omission holds significance.

Mathematically, the Jaccard similarity between two sets A and B is calculated as:

$$\text{Jaccard similarity } (A, B) = |A \cap B| / |A \cup B|.$$

Set A and set B have in common those elements whose size is represented by $|A \cap B|$.

Total number of shared elements in set unions A and B is represented by $|A \cup B|$.

Jaccard similarity value covers a range from 0 (no shared features) to 1 (identical sets).

Jaccard similarity works depend on set size; as the set size increases, value tendency shifts and may drop. With text data, frequency plays significant role when working with terms. Jaccard similarity computational processes are easy to grasp and execute. Though dealing with large and diverse sets of data means that intersection and union sizes lose relevance. During document similarity analysis, Jaccard similarity measures the similarity between documents by treating each page as individual sets of words. Document content similarities are measured by calculating the Jaccard similarity. For example, suppose we have two documents:

Document 1: “cat, dog, pet”.

Document 2: “dog, pet, animal”.

$\text{Jaccard Similarity } (\text{Document 1}, \text{Document 2}) = |\{\text{cat, dog, pet}\} \cap \{\text{dog, pet, animal}\}| / |\{\text{cat, dog, pet}\} \cup \{\text{dog, pet, animal}\}| = 2/4 = 0.5$

$$\text{Jaccard Similarity } (\text{Document 1}, \text{Document 2}) = |\{\text{cat, dog, pet}\} \cap \{\text{dog, pet, animal}\}| / |\{\text{cat, dog, pet}\} \cup \{\text{dog, pet, animal}\}| = 2/4 = 0.5.$$

With a Jaccard similarity of 0.5, 50% of the terms in Document 1 overlap with those found in Document 2.

Jaccard similarity has its applications in information retrieval, document identification through duplication or near-duplication detection, document clustering, etc. With higher Jaccard similarity, groups of related documents are formed. High similarity between documents could mean that duplication is present. Recommendation systems like paper recommendation system where Jaccard similarity is used

to find the resemblance between the enquiries made by users (in terms of their characteristics) and the characteristics of the papers [6].

Key distinctions between cosine similarity and Jaccard similarity [7]:

1. Cosine similarity considers the entire sentence vector, whereas Jaccard similarity calculates the similarity based on the unique length of words in the set.
2. When it comes to evaluating the similarity between two vectors, cosine similarity is a suitable method, especially in cases where data duplication is not a concern. However, if data duplication is a significant factor, it is preferable to utilize Jaccard similarity. This approach allows us to accurately measure similarity without being influenced by duplicated data.

2.3 Euclidian Distance

Euclidean distance is a commonly used metric in analyzing document similarity and other fields. It computes the straight-line distance between two points in Euclidean space [8], typically utilized for comparing numerical or vector data. In the realm of document similarity analysis, Euclidean distance can be employed to compare documents represented as the vectors in high-dimensional space, like the term frequency-inverse document frequency (TF-IDF) space or word embedding space.

To analyze texts, each document is transformed into a high-dimensional vector. This vector reflects the presence and frequency of terms (words or phrases) across the entire collection of documents. The value assigned to each dimension is usually based on either the term's frequency within the document (in TF-IDF) or its real-valued representation as a vector (in word embeddings such as Word2Vec or GloVe).

To compare two documents, the Euclidean distance is calculated by treating their vectors as points in a high-dimensional space. The distance is then determined as the straight-line distance between these points. To determine the distance between two document vectors A and B, we rely on the Euclidean distance formula:

$$\text{Euclidean distance } (A, B) = \sqrt{(\sum (a_i - b_i)^2)}$$

where when it comes to interpreting the Euclidean distance, a smaller value indicates that two documents are closer in the high-dimensional space and therefore considered more similar. On the other hand, a larger value means that the documents are farther apart and therefore considered less similar.

Euclidian distance finds its applications in information retrieval, recommendation system based on content-based filtering, text classification, etc. One commonly used method for clustering similar documents together in algorithms like K-Means is through the use of Euclidean distance.

There are various models that we use to find document similarity like Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), Bidirectional Encoder Representations from Transformers (BERT), RoBERTa, etc.

2.4 Latent Semantic Analysis (LSA)

Latent Semantic Analysis (LSA) is a method utilized in the field of natural language processing technique that seeks to uncover hidden patterns of meaning within documents. By applying principles from linear algebra and dimensionality reduction, LSA can effectively identify latent semantic structures in large textual datasets. This approach proves valuable in revealing underlying connections and meanings that may not be immediately apparent. Here, we explore the fundamental principles behind LSA and its unique capability to unveil latent semantic structures. LSA uncovers meaningful patterns by examining the connections among terms, documents, and latent semantic concepts in a reduced dimensional space. It can be utilized for tasks like document clustering, document retrieval, topic modeling, natural language processing, and information retrieval.

LSA uses a high-dimensional vector space model to represent documents and terms. In this model, each document and term is assigned a specific vector in the space [9]. The vector space model (VSM) is a representation that shows the connection between documents and terms. The concept relies on statistical metrics such as term frequency-inverse document frequency (TF-IDF). LSA utilizes a term–document matrix, also known as the document–term matrix. In this matrix, terms are represented by rows, documents by columns, and the cells contain either term frequencies or TF-IDF weights. At the heart of LSA is Singular Value Decomposition (SVD), a technique for reducing dimensions. SVD factorizes the term–document matrix into three matrices: U , Σ (Sigma), and V . The left singular vectors (also known as the term-topic matrix) are denoted by U , while the diagonal matrix of singular values is represented by Σ . The right singular vectors, which form the document–topic matrix, are denoted by V .

LSA streamlines the term–document matrix by retaining solely the foremost significant singular values and their corresponding vectors. Usually, this results in a much smaller matrix dimension. Dimensionality reduction has the benefit of reducing noise and eliminating less relevant information by retaining only the most informative dimensions. By reducing the complexity of the term–document matrix, Latent Semantic Analysis (LSA) simplifies documents and terms into a lower-dimensional space. Each dimension in this space represents a hidden concept or topic known as “latent semantics”.

2.5 Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) is a popular technique in natural language processing and text analysis for topic modeling. It uses probabilistic modeling to uncover hidden topics within a set of documents. LDA has several applications in document similarity tasks. LDA can be used to uncover hidden topics within a collection of documents. It represents topics as distributions of words, while documents

are represented as mixtures of these topics. It helps in identifying the main topics in a document, providing a broad overview of the thematic content. This representation can then be used for analyzing document similarity.

Latent Dirichlet Allocation (LDA) is a technique which represents documents by assigning probability distributions to topics. These probability distributions indicate the presence and prevalence of each topic within a given document. This representation of a document captures the main themes or topics discussed within it, allowing for effective comparison between documents based on their thematic content. It has its applications in document clustering based on the topics, recommendation systems like drug recommendation system [10], text summarization, etc.

2.6 Bidirectional Encoder Representations from Transformers (BERTs)

Although Bidirectional Encoder Representations from Transformers (BERT) is not typically categorized as a traditional topic modeling technique, it is a cutting-edge, pre-trained transformer-based model that excels in natural language understanding and generation. Despite this distinction, BERT has demonstrated its remarkable relevance when applied to document similarity tasks and can be effectively utilized in various ways for these specific purposes.

BERT, a language model, creates contextualized word embeddings. This means that the representation of each word considers the surrounding sentence's context, allowing BERT to capture comprehensive semantic information. When it comes to tasks like measuring document similarity, BERT proves to be quite effective. It has the ability to generate contextual embeddings for whole documents, considering the connections between words and their positions within the document. This contextualization significantly enhances the quality of document representations.

To generate embeddings for entire documents, BERT can be utilized. This process involves tokenizing the text of the complete document and feeding it into the BERT model. From there, the embedding of the “[CLS]” token (which typically represents the entire sentence) is extracted. These document embeddings capture the full context and semantic meaning of the entire document, making them ideal for analyzing document similarity.

BERT-based embeddings can be utilized to improve document retrieval tasks, document clustering, recommendation systems like research-related recommendation system, etc. [11]. By ranking documents based on their semantic similarity to a given query, more precise information retrieval can be achieved. This allows for better accuracy in finding relevant documents.

BERT has the ability to be fine-tuned for specific tasks involving document similarity. This involves training BERT using labeled pairs of documents that are either similar or dissimilar, allowing it to generate document embeddings that are specifically optimized for the given similarity task. Fine-tuning BERT has the potential

to improve performance by enabling it to adapt to the specific characteristics of the document similarity dataset.

2.7 Robustly Optimized BERT Approach (*RoBERTa*)

RoBERTa, an advanced natural language processing model, was created by Facebook AI. Robustly optimized BERT approach (RoBERTa), a model based on the BERT model that was trained over a longer period of time and with more data [12]. It has shown significant advancements in various NLP tasks, such as document similarity.

Prior to being fine-tuned for specific tasks, models like RoBERTa undergo a process called pretraining. This involves exposing the model to an extensive amount of textual data and implementing a self-supervised learning technique. Through this process, the model learns to predict the missing words in sentences by considering contextual information from both directions—left-to-right and right-to-left. This enables the model to effectively capture bidirectional context and enhance its understanding of language.

Measuring document similarity is an important task, and RoBERTa's ability to capture semantic and contextual information makes it ideal for this purpose. By encoding documents into RoBERTa embeddings and using similarity measures like cosine similarity, you can accurately gauge the similarity between two documents. This enables efficient analysis of textual data and comparison of document content. Semantic understanding plays a crucial role in RoBERTa's contextual embeddings. These embeddings have the ability to capture not only the surface level meaning of words and phrases but also their deeper semantic nuances. By leveraging this power, RoBERTa can deliver more accurate and context aware measurements when applied to document similarity tasks. RoBERTa has undergone pretraining on multiple languages, making it an effective tool for tasks involving document similarity across different language pairs.

3 Methodology

Steps involved in finding the similarity between the documents using different document similarity measures and models are as follows.

Step 1:

1. Find similarity using cosine similarity, Euclidian distance, and Jaccard similarity.
2. Compare which measure works best.

Step 2:

1. Find embeddings using LSA, LDA, BERT, RoBERTa.
2. Find similarity using cosine similarity.

```

Contents of Document1: Harappa was discovered in 1921 by Dayaram Sahni. The excavations were done under the guidance of Sir John Marshal and Colonel Meke. Re
Contents of Document2: Harappan civilization was first identified in 1981 at Harappa in the Punjab region and then in 1922 at Mohenjo-daro (Mohenjodaro), nea
[42]: print("With different similarity measures:")
print("Similarity Percentage using Cosine Similarity: (similarity_percentage_cosine:.2f)%")
print("Similarity Percentage using Euclidean Distance : (similarity_percentage_euclidean[0][1] * 100,.2f)%")
print("Similarity Percentage using Jaccard similarity: (similarity_percentage_jaccard:.2f)%")
print("Cosine Similarity With different models:")
print("Similarity Percentage using Latent Semantic Analysis (LSA): (similarity_percentage_LSA:.2f)%")
print("Similarity Percentage using Latent Dirichlet Allocation(LDA): (similarity_percentage_LDA:.2f)%")
print("Similarity Percentage using BERT: (similarity_percentage_BERT:.2f)%")
print("Similarity Percentage using RoBERTa: (similarity_percentage_ROBERTa:.2f)%")

With different similarity measures:
Similarity Percentage using Cosine Similarity: 28.79%
Similarity Percentage using Euclidean Distance : 5.33%
Similarity Percentage using Jaccard similarity: 10.34%
Cosine Similarity With different models:
Similarity Percentage using Latent Semantic Analysis (LSA): 28.71%
Similarity Percentage using Latent Dirichlet Allocation(LDA): 4.15%
Similarity Percentage using BERT: 87.91%
Similarity Percentage using RoBERTa: 98.66%

```

Fig. 1 Document similarity using different measures and models

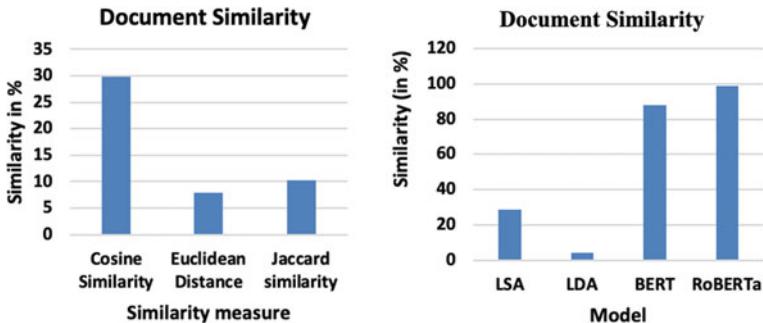


Fig. 2 Document similarity using different measures and models

Results (see Figs. 1 and 2).

4 Conclusion

In this paper, we studied various algorithms for finding similarity between the documents. Document similarity algorithms are used to determine the degree of resemblance or relatedness between various documents. Document similarity plays a pivotal role in various natural language processings (NLP), information retrieval tasks, recommender systems, duplicates detection, plagiarism detection, etc. There are various types of document similarity measures (cosine similarity, Jaccard similarity, Euclidean distance) and models like Latent Semantic Analysis (LSA), BERT, RoBERTa, etc. Cosine similarity is the most suitable measure for finding the document similarity. The results show that RoBERTa measure works most efficiently to find similarity between the given documents. BERT also performs better as compared to LSA and LDA.

References

1. Gomaa W, Fahmy A (2013) A survey of text similarity approaches. *Int J Comput Appl* 68(13):13–18
2. Han J, Kamber M, Pei J (2011) Data mining: concept and techniques, 3rd edn. The Morgan Kaufmann, India
3. Muflukkah L, Baharudin B (2009) Document clustering using concept space and cosine similarity measurement. In: 2009 international conference on computer technology and development, IEEE, Malaysia, pp 58–62
4. https://esource.dbs.ie/bitstream/handle/10788/4254/msc_jeevankrishna_2020.pdf?sequence=1&isAllowed=y. Accessed on 2023/09/10
5. Lahitani AR, Permanasari AE, Setiawan NA (2016) Cosine similarity to determine similarity measure: study case in online essay assessment. In: 2016 4th international conference on cyber and it service management, IEEE, Indonesia, pp 1–6
6. https://www.academia.edu/4041704/Content_Based_Recommendation_Systems. Accessed on 2023/09/10
7. Medium. Retrieved from <https://medium.com/analytics-vidhya/introduction-to-similarity-metrics-a882361c9be4>. Accessed on 2023/09/10
8. Wang J, Dong Y (2020) Measurement of text similarity: a survey. *Information* 11:421
9. Foltz PW (2001) Semantic processing: statistical approaches. In: Smelser NJ, Baltes PB (eds) International encyclopedia of the social & behavioral sciences, Elsevier, Pergamon, pp 13873–13878
10. Sripathi SR, Pradyumna NVS, Dhanush A (2022) Drug recommendation system using LDA. In: 2022 international conference on futuristic technologies (INCOFT), IEEE, Belgaum, India, pp 1–7
11. Yang N, Jo J, Jeon M, Kim W, Kang J (2022) Semantic and explainable research-related recommendation system based on semi-supervised methodology using BERT and LDA models. *Expert Syst Appl* 190
12. De Oliveira RS, Nascimento EG (2022) Analyzing similarities between legal court documents using natural language processing approaches based on transformers. ArXiv. /abs/2204.07182
13. Manalu DR, Rajagukguk E, Sirigoringo R, Siahaan DK, Sihombing P (2019) The development of document similarity detector by Jaccard formulation. In: 2019 international conference of computer science and information technology (ICoSNIKOM), IEEE, Medan, Indonesia, pp 1–4
14. Foltz P (1996) Latent semantic analysis for text-based research. *Behav Res Methods* 28:197–202
15. Mishra A, Panchal V, Kumar P (2020) Similarity search based on text embedding model for detection of near duplicates. *Int J Grid Distribut Comput* 13:1871–1881
16. Oduntan OE, Adeyanju IA, Falohun AS, Obe OO (2018) A comparative analysis of euclidean distance and cosine similarity measure for automated essay-type grading. *J Eng Appl Sci* 13:4198–4204
17. Machine learning mastery. Retrieved from <https://machinelearningmastery.com/a-brief-introduction-to-bert/>. Accessed on 2023/09/10
18. Hugging face. . Retrieved from https://huggingface.co/docs/transformers/model_doc roberta. Accessed on 2023/09/10
19. Gunawan D, Sembiring C, Budiman M (2018) The implementation of cosine similarity to calculate text relevance between two documents. *J Phys: Conf Ser* 978:012120

Schmitt Trigger Leakage Reduction Using MTCMOS Technique at 45-Nm Technology



Deepak Garg and Devendra Kumar Sharma

Abstract With the advent of FinFET-based circuits, the scope of Moore's law may be extended without the continuous scaling of CMOS devices. FinFET devices provide a viable method to create highly integrated, power-efficient Schmitt Trigger circuit for low-power digital applications due to its combination of enhanced flexibility and decreased short channel effects (SCEs). Schmitt Triggers are used in the design of Integrated Circuits (ICs) to produce digital signals from analogue signals in order to facilitate the implementation of short channel lengths and deliver exceptionally ultra-low power. The Schmitt Trigger Mechanism is being worked on, and the Schmitt Triggers are an important System on Chip (SoC) circuit. The simulations are carried out for various parameters such as leakage, parametric variation of power consumption, and dissipation using Cadence Virtuoso Tools at 45 nm technology. The MTCMOS technique is helpful in reducing the leakage parameters, and it is more appropriate than the conventional one. In the FinFET-based Schmitt Trigger, the leakage power is lowered up to a value of 27%, and the leakage current is reduced by 49% when compared to the conventional Schmitt Trigger.

Keywords FinFET · Schmitt trigger · MTCMOS · Leakage current · Leakage power

D. Garg (✉)

Department of Electronics and Communication Engineering, SRMIST, Delhi-NCR Campus,
Ghaziabad, India

e-mail: deepakgarg1985@gmail.com

D. K. Sharma

Department of Electronics and Communication Engineering, ABES Engineering College,
Ghaziabad, Uttar Pradesh, India

1 Introduction

In recent years, research has focused more on power usage as a crucial metric. Conventional metal–oxide–semiconductor (CMOS) circuits are the backbone of every digital integrated circuit (DIC). Power consumption is a key worry in VLSI design as a result of the continuous scaling in feature size of CMOS circuits and the corresponding rise in chip density and operation frequency. When any circuit consumes too much electricity, the chip gets too hot, the chip's lifespan shortens, and the integrated circuit's efficiency drops. As the density, practicality, reliability, and cost of integrated circuits (ICs) continue to rise, reducing power consumption becomes more crucial. By adjusting the input voltage, you can keep the circuit's power dissipation constant. However, power scaling has an impact on the circuit's performance. For this reason, businesses have been focusing on developing FinFET-based low-power, high-speed circuits [1]. To decrease leakage current and hence standby leakage power, the MTCMOS (self-controllable voltage level) technology is presented. The adjustment of supplied DC voltage of the load circuit during active operation, and reduce this voltage during standby is possible by using the MTCMOS method.

2 Consumption of Energy Using CMOS Technique

Compared to the standard Schmitt Trigger, which makes use of six transistors, the four-transistor (4T), and six-transistor (6T) versions need and consume significantly less power due to their shorter channel lengths. During operation, both the pull-up and pull-down networks are essential to a CMOS design circuit, as is common knowledge. Power dissipation must be under control if the supply voltage has been kept constant. The main controlling over the current that means to control over the leakage is supported by the FET due to its single controlling over the all channel using the double gate that are helpful for the controlling the flow of charge carriers [2]. Those phenomena are used in the double gate FET that helps to reduce the drawbacks of MOSFET at very low channel length 45 nm or less of them. In order to keep the same driving current and maximize performance, the transistor's threshold voltage must be kept constant. However, sub-threshold leakage current grows when the threshold voltage is reduced. Total power of the circuit shows the combination of static (P_{static}) and dynamic (P_{Dyn}) power. And it can express as:

$$P_{\text{Total}} = P_{\text{Dyn}} + P_{\text{static}} \quad (1)$$

where P_{Dyn} indicates the power lost during a transition in a logic gate's output signal as a result of charging and discharging capacitances. Sub-threshold leakage, gate direct tunneling leakage, and junction band-to-band tunneling leakage make

up the bulk of the leakage current and contribute significantly to the static power consumption represented by P_{static} [3].

Power consumption by DIC is:

$$P_{\text{Total}} = I_o \cdot V + \alpha CV_{\text{DD}}^2 \quad (2)$$

where I_o represents leakage current in any circuit and it defined by the diode equation, $I_s(e^{qv/kT} - 1)$. The leakage power is represented by the first component in the equation, while the dynamic switching power is represented by the second. Since V_{DD} has fallen along with the feature size, the transistor's threshold voltage (V_T) has also dropped. As a result, according to the diode equation, the V_T -dependent leakage current I_o grows. Additional information on the sub-threshold leakage expression [4]. Power usage at rest may be expressed as:

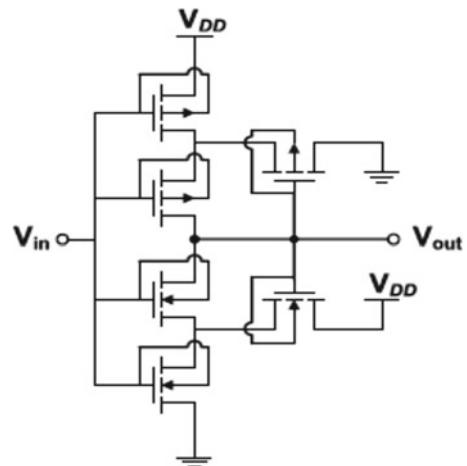
$$P_{\text{static}} = V_{\text{DD}} * I_{\text{static}} = V_{\text{DD}} * (I_{\text{gate}} + I_{\text{sub}}) \quad (3)$$

As technology advances from 180 to 45 nm, significant shifts are seen in both dynamic and static energy [5].

3 FinFET-Based Schmitt Trigger Circuit Description

In this research, a 45 nm, 120 nm-wide FinFET-based 6T Schmitt Trigger and V_{sine} pulse of amplitude 1 V and frequency 1 GHz is presented. Figure 1 depicts a FinFET-based Schmitt Trigger, which is comprised of four FinFET transistors and has the advantage of requiring a lower density chip space than competing technologies [6].

Fig. 1 Proposed FinFET-based Schmitt Trigger



A Schmitt Trigger is an electrical device used to cut down on background noise by modifying the shape of the input pulses between two predetermined voltage levels. In reality, it is a dual-inverter circuit, with positive feedback driving both inverters. The pull-up to pull-down ratio of the transistor determines the switching threshold of the design [7]. As the ratio grows, so does the switching threshold. Positive feedback helps generate steep slopes, which in turn decreases the direct current route and so leakage power [8].

4 Reduction of Leakage of a FinFET-Based Schmitt Trigger Using MTCMOS

In battery-powered devices, the processor designer must also consider power consumption. MTCMOS-based FinFET Schmitt Trigger was designed to reduce the power dissipation and increase the performance of read/write in memory to surmount this type of problem. The MTCMOS technique operates in two modes (active mode and standby mode) and the load circuit is used to execute the operation. This paper presents a FinFET-based MTCMOS Schmitt Trigger design.

FinFET-based Schmitt Trigger facilitates the demonstration of system operation control. Here, Shorted Gate DG FinFET is use, which means the double gate is shorted with each other in order to work in a low power mode, and the Independent Gate DG FinFET, which means the VG1 and VG2 are applied to the gate terminals separately and at different values in order to provide control over the threshold voltage (V_{TH}).

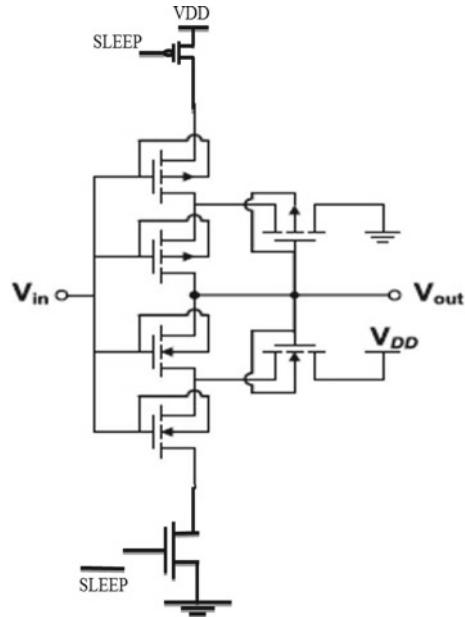
Since simple Schmitt Trigger is prone to sluggish data processing, it increases the system's launch time, whereas FinFET-based Schmitt Trigger is highly effective in addressing short channel issues [9].

MTCMOS Technique

MTCMOS is a multi-threshold CMOS technique. It is a technique where a high V_t (threshold voltage) transistors are inserted between the power supply and logic circuit or between logic circuit and ground or both. This results in creation of virtual supply or virtual ground respectively [10, 11]. The logic circuit blocks consist of low V_t transistors. The reason for imposing low V_t components in logic blocks is fast switching speed and high V_t header and footer to minimize the leakage. This technique is most effective in reducing standby leakage. The circuit diagram of proposed FinFET-based Schmitt Trigger using MTCMOS technique is shown in Fig. 2.

During normal mode both the high V_t PMOS and NMOS transistor are kept ON by applying signal to the gate of the transistors also denoted as SLEEP and SLEEP bar (sleep signal) [12]. The current flowing through this circuit will depend on the component with low threshold voltage and thus creates a virtual V_{DD} above the logic circuit block and a virtual ground below the logic block circuit whereas in case of

Fig. 2 Proposed FinFET-based Schmitt Trigger using MTCMOS technique



sleep mode both the high V_t PMOS and NMOS transistor are kept OFF and both virtual power supply rail and virtual ground rail does not exist any longer. Thus, the current that will flow through this circuit will depend on the lower of two currents as they are connected in series. Hence high V_t transistors have lower leakage current when the circuit is in standby mode. So in standby mode high V_t transistors are made OFF leading to smaller leakage current.

In minimizing the leakage parameters during the run time, the key factor is threshold voltage (V_t) which is defined as the minimum voltage at which current starts flowing. This indicates smaller V_t leads to faster operation and smaller delay [13]. To reduce the power dissipation supply voltage is scaled down and to maintain the high performance threshold voltage is scaled down.

$$P_D = V_{DD}^2 CF \quad (4)$$

5 Simulation Results

The working range and parametric analysis of the proposed circuit is verified by designing and drawing layout of circuit. To fulfill the simulation matrices requirements, which include knowing the behavior and comparing it to an idealized chart,

Fig. 3 Total power consumption at different supply voltage

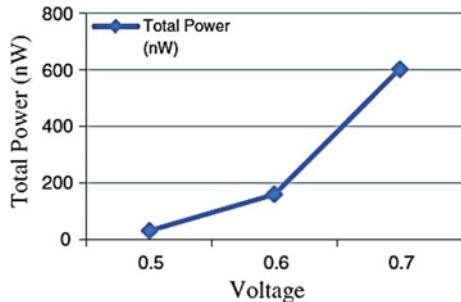


Table 1 Leakage current using various techniques in 6T Schmitt Trigger (pA)

Supply voltage (V)	Conventional circuitry (pA)	Using MTCMOS circuitry (pA)
0.5	4.58	1.895
0.6	4.07	1.564
0.7	3.67	1.48
0.8	3.023	1.351
0.9	2.09	1.04

is necessary in order to presume that our circuit is functioning properly at the front end level and that it may be sent to the foundry level.

In this paper, researcher is performing the 6T Schmitt Trigger circuit functionality and reducing the leakage current and power. MTCMOS also perform very good catalyzer for performing better in reducing the leakage parameters, but apart from that if Power Gating Technique is applied then this time the more reduced value of leakage parameters can get. The leakage parameter variation is shown in the Table 1 and Fig. 3.

6 Conclusion

In this article, Schmitt Trigger is presented which is a useful part of the System on Chip (SoC). This paper concludes that MTCMOS circuit designed for Schmitt Trigger with FinFET will use less power and have lower leakage parameters. Without having a major impact on the system's performance, this technique reduces standby leakage. Due to better device features of FinFET-based Schmitt Trigger in low voltage operation, this has become the most promising technique for lower VDD operation in FinFET technologies. Therefore, the power gating MTCMOS technique can be viewed superior to conventional approaches. These MTCMOS circuits are simulated using 45 nm technology node in Cadence Virtuoso Tool. The results show that the proposed Schmitt Trigger circuit has a 27% reduction in leakage power and a 49% reduction in leakage current when compared to the traditional Schmitt Trigger.

References

1. Pfister A (1992) Novel CMOS Schmitt Trigger with Controllable Hysteresis. *Electron Lett* 7(28):639–641
2. Wang Z (1991) CMOS Adjustable Schmitt Triggers. *IEEE Trans Instrum Meas* 40(3):601–605
3. Dokic BL (1984) CMOS Schmitt Triggers. *IEEE Proc G. Electron Circ Syst* 131:197–202
4. Weste NHE, Eshraghian K (1985) Principles of CMOS VLSI design: a systems perspective. Addison-Wesley Longman Publishing Co., Inc
5. Pedroni VA (2005) Low-voltage high-speed Schmitt trigger and compact window comparator. *Electron Lett* 41(22):1
6. Saxena A, Shrivastava A, Akashe S (2014) Design and performance evaluation of schmitt trigger for nanoscale CMOS. *Quantum Matter* 3(1):52–56
7. Mishra V, Akashe S (2015) Calculation of average power, leakage power, and leakage current of finfet based 4-bit priority encoder. In: 2015 Fifth international conference on advanced computing & communication technologies. IEEE, pp 65–69
8. Mishra V, Akashe S (2015) Calculation of power delay product and energy delay product in 4-bit finfet based priority encoder. In: Advances in optical science and engineering: proceedings of the first international conference, IEM OPTRONIX 2014. New Delhi: Springer India, pp 283–289
9. Al-Sarawi SF (2002) Low power Schmitt trigger circuit. *Electron Lett* 38(18):1
10. Khandelwal S, Raj B, Gupta RD (2013) Leakage current and dynamic power analysis of FinFET based 7T SRAM at 45 nm technology. In: Proceeding of the international arab conference on information technology, Khartoum Sudan
11. Garg D, Sharma DK (2023) Towards the evaluation from low power VLSI to quantum circuits. In: Quantum-dot cellular automata circuits for nanocomputing applications. CRC Press, pp 1–24
12. Katyal V, Geiger RL, Chen DJ (2008) Adjustable hysteresis CMOS Schmitt triggers. In: 2008 IEEE international symposium on circuits and systems. IEEE, pp 1938–1941
13. Saini S, Veeramachaneni S, Mahesh Kumar A, Srinivas MB (2009) Schmitt trigger as an alternative to buffer insertion for delay and power reduction in VLSI interconnects. In: TENCON 2009–2009 IEEE Region 10 Conference. IEEE, pp 1–5

A Review of Survey and Assessment of Facial Emotion Recognition (FER) by Convolutional Neural Networks



Sanyam Agarwal, Veer Daksh Agarwal, Ishaan Agarwal, Vipin Mittal, Lakshay Singla, and Ahmed Hussein Alkhayyat

Abstract Computer vision and the area of artificial intelligence (AI) both heavily rely on the detection of facial expressions. This article concentrates on operations based on face images. It demonstrates how visual articulations are most important data facilitates, despite the limitless possibilities of how FER can be analyzed by using various instruments. This essay provides a succinct analysis of recent FER research. However, theoretical FER structure designs and their initial evaluations are displayed close by conventional FER approaches. The presentation of numerous FER views using the “start to finish” learning permission through critical associating authorization follows. As a result, this study will help in connecting a convolutional neural network (CNN) for some LSTM components (long transient memory). This paper concludes with a short poll, evaluation assessment, findings, and standards that serve as a standard for measurable connections between all of these FER studies and experiments. For students in FER, this audit can serve as a succinct manual that provides pertinent details and evaluation for recent tests. Additionally, knowledgeable examiners are searching for promising paths for future work.

S. Agarwal

Department of Electronics and Communication Engineering, ACE College of Engineering and Management, Agray, India

V. D. Agarwal (✉) · L. Singla

Departmen of Computer Science Engineering, Thapar Institute of Engineering and Technology, Patialay, India

e-mail: veerdaksh2012@gmail.com

I. Agarwal

Department of Computer Science Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

V. Mittal

Department of Electronics and Communication Engineering, IIMT University, Meeut, India
e-mail: vipinmittal@iimtindia.net

A. H. Alkhayyat

Scientific Research Centre of the Islamic University, The Islamic University, Najaf, Iraq

Keywords Computer vision · Convolutional neural network · Facial emotion recognition · Sensors · Traditional FER

1 Introduction

A person's facial movements play a crucial role in conversation because they allow others to read their emotions behind their words. It is used everywhere for presentations, and the vocal tones are used to evaluate. Enthusiasm, such as happiness, sadness, burden, astonishment, is the different facial movements. Various reviews show [1] that the verbal part conveys 33% of human response, and the non-verbal part represents 66%. The significance of animation is one of the typical data distractions in social communication, along with two or three other non-verbal components. In this way, an unexpectedly lengthy time of study of face tilt, which is used in perception by science, the arts, and thrilling math and computer developments, has added to a thoughtful package [2]. Study of Personalized Facial Tilt Statements (Each treatise has different advanced kinds of FER for social events, as well as face tilt proof and appearance certificates [3]. As this study is consistent with the general one, the expression FER prompts the request for the facial tilt. A quickening of the development of deceptive, false techniques like Human–Computer Collaboration (HCI) [4] PC-Made Reality (VR) [5], Advanced Driver Right Hand Situations (ADAS) [6], Augmented Reality (AR) [7], and Entertainment. One can use various devices, but the camera is the most popular because it gives FER the most explicit cues and should not be worn. The research on modified FER has been divided into two social meetings in the first section of this treatise. The work's handcrafted nature, originality are indicators of this. The FER which is created using three fundamental steps in conventional FER approaches, as shown in Fig. 2.

- (a) Face and frontal area
- (b) Include look collection, feature extraction
- (c) Observing appearances

Early on, the facial image can be distinguished from the information picture, and the facial features (eyes, nose, etc.) can be used to determine success from the face portion. Second, the face region is used to derive spatial and temporal anomalies.

Thirdly, limits are used by tailored Face Emotions classifiers to produce forceful results, including supplementary vectors with machine processing (SVM), AdaBoost, and unstable back-wood regions (Figs. 1 and 2).

Using a fixed layout, facial localization and facial characteristics are identified from the input picture, and spatial and standard parts are removed from face. Based on fundamental facial expressions that have been established, the attitude is represented. (Frontal face pictures are being collected from the CK + collection) [8].

Meaningful learning has become a frequent method of monitoring AI in place of the traditional tactic of using carefully selected highlights. Because one has access to such vast knowledge, it receives the highest rating in all PC-Vision reviews [9].



Fig. 1 Compromises of pictures existing system in the Cohn-Kanade dataset

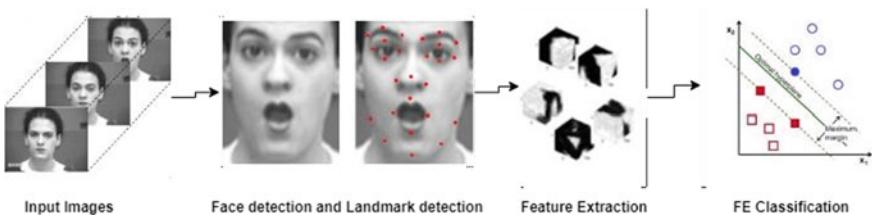


Fig. 2 The standard approach of the FER

As shown by the informational pictures, a crucial learning-based FER method is the actual science. It enables one to find a workable way to apply it in his pipeline “from start to finish” with less dependence on basic models and other feed-forward techniques [1]. The Convolution Mind Affiliate (CNA), a specific kind of crucial learning, is the most noteworthy linked model among the two or three main learning models that are currently accessible. In the CNN-based method, data pictures are gathered and immediately convolved to create a section map in the convolution layer. The facial aspects are then taken into account as a particular type of location based on the softmax outcome conclusion, and each sub-map is then linked to an entirely associated network. The CNNs method to the FER is shown in Fig. 2, where it is split into two subcategories. Bundles or film pictures are used to illustrate this [10]. Every time, the fixed FER entirely depends on the uniform facial portions acquired by truncating carefully chosen top-appearance accommodations in the picture group. Also spatial transients are included in the dynamic (video-based) FER to give the moving component a look approach. Because they provide additional inventory info, strong FERs are known to have a higher level of authentication than stable FERs, but they also have some disadvantages. For instance, based on the specific appearance, the apparent accents that have been removed have varying durations of change and partial honors. Additionally, the standardization process that is typically used to get a better grouping with the correct number of modifications might result in a brief absence of scale data (Fig. 3).

The input image is convoluted by adding a detour in the convolution layer. The integration guide is created from the convolution results, and the max-pooling layer (subsampling) reduces the spatial target of the specified partial map. CNN applies

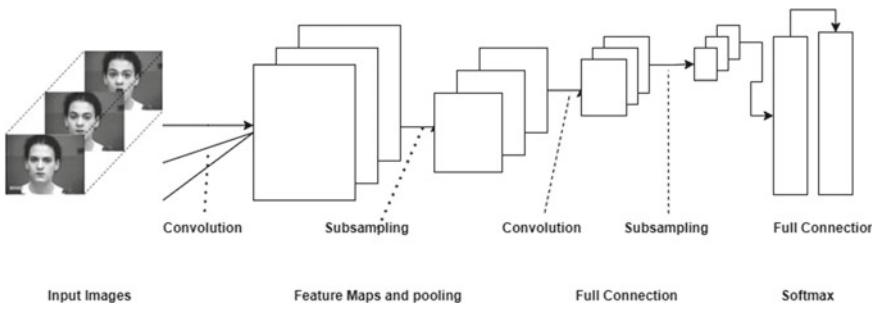


Fig. 3 CNN-based FER approach methodology

an entirely relevant network layer behind the convolutional layer, and if the softmax sequence is delayed, one will see a lonely facial appearance (the facial image is from the CK + dataset) [8].

1.1 Terminology

Record the sporadic communication anticipating the fundamental portion of FER studies below before considering any studies relating to FER:

- According to Ekman and Friesen [11] in 1978, the Facial Action Coding System (FACS) is a paradigm for describing changes in facial muscles and how they are used to express specific human feelings. A face muscle customization of facial emotions known as FACS encrypts the developmental unit. It offers a clearly discernible transient change in facial emotions [12].
- The facial spots of interest, which are obviously noticeable focus points in regions of the face like the edges of the nostrils, the side of the nose, and below the eyes and the lips, are depicted in Fig. 2b. As the intersection vector of FER, the paired points of all two power groups or neighboring power surfaces are used. The FL disclosure method can be divided into three categories: model time, z, and broadly as suggested by the model and CNN-based framework. The FL model is an organized model with differences in the rough set's look and form. At this stage, the fundamental shape keeps shifting into the primary position till it comes together [13].
- The seven fundamental human beliefs—satisfaction, astonishment, happiness, annoyance, dread, sadness, and an open mind—are known as basic beliefs (BE). Complex emotions (CE) are a synthesis of two key ideas. One and his coworkers [14] introduce 22 feelings along with 7 fundamental and 12 abstract views. Traditionally, people were the only ones who could spread all of this. Three more viewpoints were offered. (frustration, contempt, and surprise) (Fig. 4).

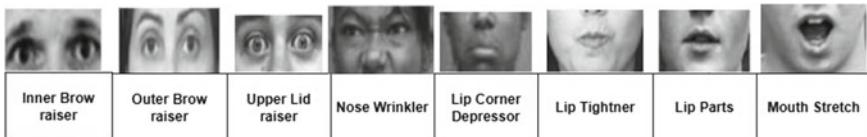


Fig. 4 Different facial views and AU test scenarios were used, including emotions (sad, frightened, disturbed), complicated emotions (euphoric, happily furious), catastrophic misery, and unrestrained mentality (facial feature images from the CK + dataset) [8]

- More cautious emotions indicate the more unrestricted, imperceptible face improvements that typically take place than regular articulations (MEs). They frequently disclose a person's true and most significant views in a brief period of time.
- The facial development unit (AU), as depicted in Fig. 3d, generates a lot of individual or communal muscular action (46 AU), which is typically observed when generating energy of a specific propensity [14] Encode. Individual AUs are identified for categorizing facial emotions, and reports are compromised based on the combination of AUs. The framework would suggest a “shocked” class if the picture had significance based on a grade of 1, 2, 25, or 26 AU, as shown in Table 1 (Fig. 5).

1.2 Contributions of This Review

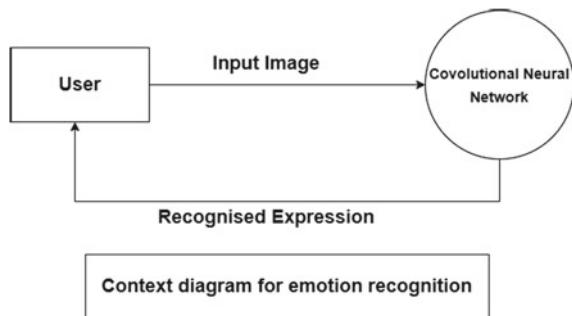
Whatever the case, there has never been a clear-cut method for researching FER in the lengthy history linked with it. A significant bias-based approach is not presented in certain evaluations [16, 17], which exclusively focus on conventional research. Recently, Ghayoumi [18] published a harsh research on Goliath learning in FER. In any event, what was fundamentally given was a summary of the focus between conventional procedures and a significant learning-based approach. The remainder of this paper focuses on a brief investigation from conventional FER to late altitude FER.

Many standard information bases incorporate actual picture, video for FER use, along with their traits and goals.

To the degree that accuracy and resources permit, essential differences between ordinary FERs and sizable learning-based FERs are considered. Most of the time, large learning-based FERs outperform regular FERs in terms of FER accuracy, but they must consider constraints like the optimal use of GPUs and CPUs. A large load is also necessary. Similar to this, different FER ratings are now applied in embedded frameworks, such as PDAs.

Table 1 Principal AU discovered in each class of basic and complex emotions preserved from [15]

Category	AUs	Category	AUs
Happy	12 and 25	Sad-disgust	4 and 10
Sorrow	4 and 15	Fear-anger	4, 20, and 25
Scared	1, 4, 20, and 25	Fear-surprise	1, 2, 5, 20, and 25
Anger	4, 7, and 24	Fear-disgust	1, 4, 10, 20, and 25
Surprise	1, 2, 25, and 26	Anger-surprise	4, 25, and 26
Disgust	9, 10, and 17	Disgust-surprise	1, 2, 5, and 10
Happy-sad	4, 6, 12, and 25	Happy-fear	1, 2, 12, 25, and 26
Happiness-surprise	1, 2, 12, and 25	Anger-disgust	4, 10, and 17
Happiness-disgust	10, 12, and 25	Awe	1, 2, 5, and 25
Sad-fear	1, 4, 15, and 25	Appalled	4, 9, and 10
Sad-anger	4, 7, and 15	Hate	4, 7, and 10

Fig. 5 Context illustration for emotion detection

1.3 Organization of This Review

The remaining review work is divided into three sections: a standard FER approach with an overview of agent action flows for FER structures and their standard evaluations, an advanced FER approach utilizing extensive research computation, the

open access FER information base, and test evaluation related to benchmark results. Finally, here are a few or three final thoughts on potential future projects.

2 Conventional FER Approaches

Various types of traditional methods have been considered for modified FER structures. A typical attribute of these strategies is to look at the facial area and limit the cross-collection of numerical parts and phenotypes. For numeric parts, the connections between face parts are utilized to generate segment vectors for planning [19, 20]. Two numerical components [20], taking account of the coordinates of the 52 facial Power-Points. Anyway, the point and Euclidean distance linking each pair of points of interest are within an utterly non-fixed range, then the distance and focus are deducted from the separation distance and focus of the video development header package. Two frameworks are presented for classifiers: multi-class AdaBoost with dynamic time travel or SVMs with maintained component vectors. Perhaps Happy et al. to use everything as a part and act as a framework [21] used a model (LBP) histogram nearly twice the size of the different squares from the entire facial area as a subvector and referenced different appearances using a substantive subtest (PCA). However, although this frame is always executed, the part vector cannot reflect the proximity map of the facial parts, which reduces the overall accuracy of the request. Under no circumstance, the importance of each facial area should not vary with the overall parts-based approach. For sample, the mouth and eyes possess greater information than shelters and cheeks. Ghimire et al. [22] removed the surfacing features of the area by uniquely inserting the whole facial area into the space near the area. The basic root is resolved with a consistent progressive tracking approach, reducing sub focal points and improving detection accuracy. For cross-arrangement units, some methods [15–24] investigate the lack of two methods of combining numerical and optical components, and in clear cases, can lead to unparalleled results. Video development uses a variety of structures [15–25] to investigate numerical improvements in facial performance between the current wrapper and past edges as regular parts and to characterize the appearance of spatial parts. The standard capacity between the still image and the FER of the video plan is that the performance of the last choice is tracked frame by frame, and the structure creates a new powerful aspect by moving between past and present accommodations. The relative display score is then used in the computer game schedule, as shown in Fig. 2. Using a high-speed camera to get a video unfold of the face to see the tiny look. Polikowski et al. [26] presented a detailed description of the face in a video plan captured by a high-speed camera at 200 frames per second (fps). These study areas face the district in a precise location and for some time (Table 2).

Table 2 A brief summary of FER databases present that are publicly available

Reference	Emotion analyze	Visual features	Decision methods	Database
Compound emotion [14]	7 emotions and 22 complex emotions	Allocation between each set of criteria	Closest average classifier and kernel subclass discriminant check	CE [14]
EmotioNet [15]	23 basic and complex emotion	Euclidean distance among normalized landmark, centers of gravity Gabor channel, centered on the center of power	Discriminant analysis of kernel subclass	CE [14] CK + [8], DISFA [27]
Ghimire and Lee [20]	7 emotions	Move linking landmarks in consecutive frames	Multi-class AdaBoost and SVM	CK + [8]
Local region specific feature [28]	7 emotions	Focus on the second element from a mathematically, standardized, explicit short distance	SVM	CK + [8]
3D facial expression[29]	6 prototype emotions	3D waveforms and 3D patch shaped through analyzing waveforms to form groups	Multiboosting and SVM	BU-3DFE [30]
Stepwise approach[31]	6 prototype emotions	Stepwise linear discriminant analysis (SWLDA)	Hidden conditional random fields (HCRFs)	CK + [8], JAFFEE [32], B + [33], MMI [34]

3 Deep Learning-Based FER Approaches

Over the years, there have been advancements and improvements in deep learning algorithms related to ANN, CNN, and recurrent neural network (RNN). The above deep learning-based algorithms were handed down for various machine learning purposes. The principal edge regarding CNN is that it permits “end-to-end” learning straight from the processed image, entirely terminating or remarkably bringing down the need to use physics-based replicas and additional preprocessing approaches [35]. For these reasons, CNN has won top marks in a variety of regions, for example, object identification, face acknowledgment, FER, etc. The CNN comprises of three sorts of heterogeneous layers, a convolutional layer, a most extreme pooling layer, and a completely associated layer. Being displayed in Fig. 3.

The convolution layer takes a picture or element map as info, accumulates it in a progression of channel banks utilizing the sliding window strategy, and gives a component map that establishes the facial picture's spatial plan. Feature map convolution filter weights are shared, and feature map level inputs are merged provincially [36]. The subsampling layer then brings down the structural resolution of the depiction by leveling or maximizing the specified input feature map to lower the dimensions. It ignores geometric deformation and small shifts in flexibility [36, 37]. At the end, there is a fully associated layer of the CNN structure that computes the merit for the whole authentic image. Nearly all procedures build on deep learning [37–40] adapts the CNN immediately after AU detection. The CNN perception technique to perceive models prepared on different FER datasets and were prepared in feeling acknowledgment on both, the whole dataset and different FER-related assignments [38]. Exhibit the functionality of the network Junge et al. [39]. Two distinct kinds of CNNs are used, the one draws out the worldly appearance attributes from picture successions, and the latter draws out the transient mathematical highlights from fleeting facial markers. These two replicas are merged utilizing the latest combination procedure to enhance the facial expression detection performance. The Deep Region and Multilabel Learning (DRML) is a district layer that can use anticipatory capacities to direct essential facial areas and secure facial underlying subtleties with learned loads [40]. The whole network is constantly trained and automatically learns the robust characterization of fluctuations inside the local region.

4 Result

The following table shows the difference between the deep learning-based methodology and the traditional one based on the CK + dataset [8] (Table 3).

5 Conclusion

This paper presents a simple survey of the FER methodology. As we have presented, such a method may be isolated on a traditional FER rule, including three levels, a traditional FER rule, including face and face revelation, and a traditional ferrule. A review of Get Together using a typical SVM, Adaboost, and sporadic back-wood. Of course, a vast learning remote approach is confident about the Fosterian science-based model and other preceding schemes by simply picking up the data images in the pipeline. As an important learning of a particular type, CNN provides an information image to support the model learned by various facial assets, and in the sense of unmistakable confirmation of liabilities, indicates the limit. In a way that CNN-based FER framework cannot reflect temporary combinations within face parts, cross-concert approaches are CNN for each accommodation space element and some accommodate normal parts. Some recent reviews have evaluated the CNN

Table 3 Contrast in FER advanced using traditional and deep learning methodology using CK + dataset [31]

Types	Authors	Method	Accuracy in %
Tradition approach	Ghimire et al. [3]	Local representation, LBP + NCM features	97.25
	Zhang, Zao and Lei [23]	LBP + SVM	95.24
	Pouraber et al. [41]	Texture and geometric feature	92.20
	Zhao et al. [22]	LBP-TOP + VLBP	96.26
	Zhao and Zhang [28]	LBP + KDIso map	94.88
	Zhi et al. [4]	Graph preserving sparse NMF	94.30
	Li et al. [42]	Geometric features dynamic Bayesian network	94.04
Deep Learning-Based Approach	Saeed et al. [43]	Geometric features	83.01
	Mollahosseini et al. [44]	CNN	93.2
	Lopes et al. [27]	CNN	96.76
	Mohammad pour et al. [29]	CNN	97.01
	Deepak Jain et al. [30]	CNN	93.24
	Yu et al. [30]	STC-NLSTM	99.8
	Liang et al. [33]	DCBiLSTM	99.6
	Cai et al. [34]	SBN-CNN	96.87

LSTM program in appearance can defeat the applied CNN approach in temporary averaging for social opportunities. Moreover a half and half arrangement has shown a staggering execution. Little articulations stay an actuating assignment to settle since they are more unconstrained and direct facial progressions that usually happen.

This paper also presents several remarkable instructive lists connected to FER, including video groupings and pictures. Within a common dataset, the looks of a human have been centered on utilizing either fixed 2D pictures or video plans. In any case, considering the way that a 2D-based assessment experiences issues managing giant collections in present and unassuming facial ways to deal with acting, consistent datasets have considered 3D spotlights on more speedily work with an examination of the fine fundamental changes normal for unconstrained verbalizations.

6 Future Scope

Besides, assessments of FER-based approaches which are known about giving standard assessments to relationships. Assessment in the area of confirmation, accuracy, precision, and review are basically utilized. Regardless, another evaluation methodology for seeing predictable looks, or applying limited scope verbalization confirmation for moving pictures, ought to be proposed. Disregarding the way that spotlights on FER have been passed all through the most recent ten years, In fact the introduction of FER was basically edited in combination with extensive learning evaluations. Since FER is a huge technology for infusing machines, it makes sense to advance various evaluations of its future applications. Suppose agitated collaborative large-scale learning calculations can be performed later and work well with additional Internet of Things sensors. In that case, FER will set the current statement rate, including an unconstrained formulation, to a limited extent. It is exceptionally expected that they can be significantly something very similar.

7 Summary

This paper gave a short outline of FER methodology, which is accumulated into two classes: standard FER approaches and critical learning-based FER methods. It combined a compact outline of probably the guideline FER information bases. Besides introducing a conversation that underlines the high rate via prepared experts, showing that machines will be more ready to disentangle feelings later on, construing that human-machine contact will become more common. FERs are possibly the most essential mean of passing on data about a person's vivacious state, all these can be obtained just by realizing the six crucial opinions. It conflicts with what is available in customary, ordinary presence, which has more confused sentiments. This will request that scholastics make more noteworthy enlightening records and substantial critical learning developments to see all vital and optional opinions later on in the future.

References

1. Walecki R, Rudovic O (2017) Deep structured learning for facial expression intensity estimation. *Image Vis Comput*
2. Kaulard K, Cunningham DW, Büthhoff HH, Wallraven C (2012) The MPI facial expression database—a validated database of emotional and conversational facial expressions. *PLoS ONE*
3. Chu WS, Torre FD, Cohn JF (2017) Learning spatial and temporal cues for multi-label facial action unit detection. In: Proceedings of the 12th IEEE international conference on automatic face and gesture recognition, Washington, DC, USA, 30 May–3 June 2017
4. Gunawan AAS (2015) Face expression detection on Kinect using active appearance model and fuzzy logic. *Procedia Comput Sci*

5. Hickson S, Dufour N, Sud A, Kwatra V, Essa IA (2017) Eyemotion: classifying facial expressions in VR using eye-tracking cameras. arXiv
6. Assari MA, Rahmati M (2011) Driver drowsiness detection using face expression recognition. In: Proceedings of the IEEE international conference on signal and image processing applications, Kuala Lumpur, Malaysia, 16–18 Nov 2011
7. Chen CH, Lee IJ, Lin LY (2015) Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Res Dev Disabil
8. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I, The extended Cohn-Kanade Dataset (CK+): a complete dataset for action unit and emotion- specified expression. In: Proceedings of the IEEE conference on computer vision and
9. Kahou SE, Michalski, V, Konda K (2015) Recurrent neural networks for emotion recognition in video. In: Proceedings of the ACM on international conference on multimodal interaction, Seattle, WA, USA, 9–13 Nov 2015
10. Kim DH, Baddar W, Jang J, Ro YM (2017) Multi-objective based Spatio-temporal feature representation learning robust to expression intensity variations for facial expression recognition. IEEE Trans Affect Comput
11. Ekman P, Friesen WV (1978) Facial action coding system: investigator's guide, 1st edn. Consulting Psychologists Press, Palo Alto, CA, USA
12. Hamm J, Kohler CG, Gur RC, Verma R (2011) Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders. J Neurosci Methods
13. Jeong M, Kwak SY, Ko BC, Nam JY (2017) Driver facial landmark detection in real driving situation. IEEE Trans Circ Syst Video Technol
14. Tao SY, Martinez AM (2014) Compound facial expressions of emotion. Natl Acad Sci
15. Benitez-Quiroz CF, Srinivasan R, Martinez AM (2016) EmotioNet: an accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 26 June–1 July 2016
16. Kolakowska A (2013) A review of emotion recognition methods based on keystroke dynamics and mouse movements. In: Proceedings of the 6th international conference on human system interaction, Gdansk, Poland, 6–8 June 2013
17. Kumar S (2015) Facial expression recognition: a review. In: Proceedings of the national conference on cloud computing and big data, Shanghai, China, 4–6 Nov 2015
18. Ghayoumi MA (2017) Quick review of deep learning in facial expression. J Commun Comput
19. Suk M, Prabhakaran B (2014) Real-time mobile facial expression recognition system—a case study. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, Columbus, OH, USA, 24–27 June 2014
20. Ghimire D, Lee J (2013) Geometric feature- based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines. Sensors
21. Happy SL, George A, Routray A (2012) A real time facial expression classification system using local binary patterns. In: Proceedings of the 4th international conference on intelligent human computer interaction, Kharagpur, India, 27–29 Dec 2012
22. Ghimire D, Jeong S, Lee J, Park SH (2017) Facial expression recognition based on local region specific features and support vector machines. Multimed Tools Appl
23. Siddiqi MH, Ali R, Khan AM, Park YT, Lee S (2015) Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEETrans Image Proc
24. Khan RA, Meyer A, Konik H, Bouakaz S (2013) Framework for reliable, real-time facial expression recognition for low resolution images. Pattern Recognit Lett
25. Torre FD, Chu WS, Xiong X, Vicente F, Ding X, Cohn J (2015) IntraFace. In: Proceedings of the IEEE international conference on automatic face and gesture recognition, Ljubljana, Slovenia, 4–8 May 2015
26. Polikovsky S, Kameda Y, Ohta Y (2009) Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor. In: Proceedings of the 3rd international conference on crime detection and prevention, London, UK, 3 Dec 2009

27. Mavadati SM, Mahoor MH, Bartlett K, Trinh P, Cohn J (2013) DISFA: a spontaneous facial action intensity database. *IEEE Trans Affect Comput*
28. Szwoch M, Pieniążek P (2015) Facial emotion recognition using depth data. In: Proceedings of the 8th international conference on human system interactions, Warsaw, Poland, 25–27 June 2015
29. Maalej A, Amor BB, Daoudi M, Srivastava A, Berretti S (2011) Shape analysis of local facial patches for 3D facial expression recognition. *Pattern Recognit*
30. Yin L, Wei X, Sun Y, Wang J, Rosato MJ (2006) A 3D facial expression database for facial behavior research. In: Proceedings of the international conference on automatic face and gesture recognition, Southampton, UK, 10–12 April 2006
31. Zhao G, Huang X, Taini M, Li SZ, Pietikäinen M (2011) Facial expression recognition from near-infrared videos. *Image Vis Comput*
32. Lyons MJ, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with Gabor wave. In: Proceedings of the IEEE international conference on automatic face and gesture recognition, Nara, Japan, 14–16 Apr 1998
33. B+. Available online: <https://computervisiononline.com/dataset/1105138686> 29 Nov 2017
34. MI. Available, online: <https://mmifacedb.eu/29> Nov 2017
35. Walecki R, Rudovic O, Pavlovic V, Schuller B, Pantic M (2017) Deep structured learning for facial action unit intensity estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 21–26 July 2017
36. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition. *Neural Comput*
37. Ko BC, Lee EJ, Nam JY (2016) Genetic algorithm based filter bank design for light convolutional neural network. *Adv Sci Lett*
38. Breuer R, Kimmel R (2017) A deep learning perspective on the origin of facial expressions
39. Jung H, Lee S, Yim J, Park S, Kim J (2015) Joint fine-tuning in deep neural networks for facial expression recognition. In: Proceedings of the IEEE international conference on computer vision, Santiago, Chile, 7–12 Dec 2015
40. Zhao K, Chu WS, Zhang H (2016) Deep region and multi-label learning for facial action unit detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 26 June–1 July 2016
41. Shen P, Wang S, Liu Z (2013) Facial expression recognition from infrared thermal videos. *Intell Auton Syst*
42. Wei W, Jia Q, Chen G (2016) Real-time facial expression recognition for affective computing based on Kinect. In: Proceedings of the IEEE 11th conference on industrial electronics and applications, Hefei, China, 5–7 June 2016
43. Tian Y, Luo P, Luo X, Wang X, Tang X (2015) Pedestrian detection aided by deep learning semantic tasks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, MA, USA, 8–10 June 2015
44. Deshmukh S, Patwardhan M, Mahajan A (2016) Survey on real-time facial expression recognition techniques. *IET Biom*
45. Dornaika F, Raducanu B (2007) Efficient facial expression recognition for human robot interaction. In: Proceedings of the 9th international work-conference on artificial neural networks on computational and ambient intelligence, San Sebastián, Spain, 20–22 June
46. Bartneck C, Lyons MJ (2007) HCI and the face: towards an art of the soluble. In: Proceedings of the international conference on human-computer interaction: interaction design and usability, Beijing, China, 22–27 July 2007
47. Zhan C, Li W, Ogunbona P, Safaei F (2008) A real-time facial expression recognition system for online games. *Int J Comput Games Technol*
48. Mourão A, Magalhães J (2013) Competitive affective gaming: winning with a smile. In: Proceedings of the ACM international conference on multimedia, Barcelona, Spain, 21–25 Oct 2013
49. Sandbach G, Zafeiriou S, Pantic M, Yin L (2012) Static and dynamic 3D facial expression recognition: a comprehensive survey. *Image Vis Comput*

A New Compact-Data Encryption Standard (NC-DES) Algorithm Security and Privacy in Smart City



Abdullah J. Alzahrani

Abstract A smart city is an innovative, urban, organized, and sustainable city. It mainly depends on the Internet of Things (IoT) technology which improves the excellence of living, safety, the operational efficiency of urban services, decision-making, government services, and social well-being of its people. Smart city means dealing smart in education, government, mobility, households, and e-health. IoT enables all smart devices to connect through the Internet, like sensors, detectors, actuators, wearable's, mobile phones, watches, and smoke detectors. The rapid increase of Internet of Things in most smart city applications defines new security hazards that threaten the confidentiality and safety of end devices. Therefore, it is significant to improve smart services and the information protection and privacy process. The IoT devices need wireless sensor links and Radio Frequency Identification (RFID) to benefit from IoT. These resource-limited require common authentication between the devices through the association of a novel device, where authentication and encryption of the data to be sent. This paper has three contributions. First, surveying the fundamental smart city privacy problems. Second, the paper proposes a firm, trivial, and energy-efficient security solution for IoT systems which is called; New Compact-Data Encryption Standard (NC-DES). Finally, a case study of the healthcare framework is taken as an example of IoT applications to provide secure transfer of the measured parameters. The proposed system proved that the IoT devices had been secured without wasting their limited resources.

Keywords Smart city · IoT · M2M · Constraint devices · NC-DES · RPMS

A. J. Alzahrani (✉)

Computer Engineering Department, College of Computer Science and Engineering, University of Ha'il, Hail, Saudi Arabia

e-mail: aj.alzahrani@uoh.edu.sa

1 Introduction

Recently, the ideas of smart cities have been widely spread. A smart city is the combination of infrastructure, data and interaction technology, and energy-efficient systems [1]. In general terms, a smart city relies on smart technology, embedded systems infrastructure, and the Internet of Things for a better living style. The main aim of a smart city is to enhance service features and citizen prosperity. To fulfill these aspirations, it is necessary to have sincere rules, intelligent governance, intelligent training for the workforce, intelligent citizens, digital asset switching, and the deployment of intelligent wireless sensor networks everywhere [2]. IoT is the backbone infrastructure for such smart cities, where billions of devices have been enabled to be communicated through the Internet [3].

The Internet of Things (IoT) transforms conventional products, including vehicles, buildings, devices, and machines, into intelligent and interconnected objects that can establish communication with individuals, applications, services, and other devices. However, this technological advancement also introduces novel risks and confidentiality threats, as noted in reference [4]. Wireless data communications are predominantly utilized, resulting in a higher susceptibility to eavesdropping when compared to wired networks. In the absence of encryption or implementation of a weak encryption method, unauthorized access to the message by an intruder is possible, resulting in a breach of confidentiality. The safeguarding of information security and data communication protection is imperative in the infrastructures of smart cities. This is due to the susceptibility of data assets and transmission to various forms of malicious attacks, as well as the lack of trustworthiness of both internal and external participants. As such, security serves as a crucial requirement for the acceptance of end-users [5]. To accurately determine the requirements and limitations, it is imperative to examine the potential smart cities to assess their achievements and deficiencies in the realm of cybersecurity.

As previously stated, the Internet of Things is recognized as the central pillar of smart city infrastructure. IoT devices are commonly characterized by resource limitations. The utilization of lightweight cryptographic algorithms is a highly suitable approach for protecting data communication in settings where minimal power consumption is required, while simultaneously ensuring that security is not compromised. The establishment of secure and dependable smart cities necessitates the acknowledgement of various technologies, with a particular emphasis on IoT. This paper presents a tripartite contribution, beginning with an extensive study of the core privacy risk. The proposed solution for practical data security in smart city is an application of an intelligent encryption algorithm known as New Compact-Data Encryption Standard (NC-DES), which is deemed suitable for IoT environment devices. This approach considers key influencing factors. The present work includes analysis of the applications of IoT to the healthcare system, specifically through a case study to evaluate the efficiency of our proposed approach. Today, IoT implementations are founded in so many areas of human life like healthcare. Due to many IoT applications, we shall have accurate information predication to make accurate decisions in

the perfect time and so that we can face any life-threatening and natural calamity. In the next section, we will deal with a healthcare system, which is the highest critical IoT applications of smart cities [5].

The paper shows a suggested secure remote health observing system based on lightweight NC-DES algorithm during the pandemic time, in particular the future effect of the pandemic on remote health monitoring and data protection and privacy consideration. The paper contribution is as following:

- The Arduino UNO microcontroller, which is connected to the five sensors BT sensor, HR sensor, SpO₂ sensor, ECG sensor, and BP sensor that measures the condition of a patient.
- The information is collected through sensors at the microcontroller then compare the results with the stored normal data.
- If the reading is normal, patient's data is encrypted employing NC-DES and forward by the connected Bluetooth module to the database server to store the data.
- If the collected data exceed or less than the normal range, patient's data is encrypted employing NC-DES and forward by the connected GSM and send an emergency message to doctor's mobile. Besides, the control signal will be sent to the buzzer and the abnormal reading value of the physical parameter is forward to Liquid Crystal Display (LCD).

2 IoT Security

IoT platform allows devices/objects to detect, observe, measure, identify, and recognize a situation, condition, state, circumstance, or the surroundings with the absence of the social factor. Therefore, IoT is way of a group of interconnected real objects as a network utilizing both wired and wireless communication technologies that can be accessed through the Internet. The entities encompassed by IoT that can be modeled as nervous system observation equipment for patients with integrated instruments. These entities are allocated an IP address and can transfer the information over the Internet, as noted by Sterbenz [4]. The incorporation of embedded technology within devices facilitates their ability to interact with both internal and external environments, thereby influencing decision-making processes [6]. From linked homes, enterprises, authorities, and smart cities to linked vehicles and machines to devices that keep track of a person's behavior and use the data gained for different kinds of services. Devices are deployed in any place, either public or private, to provide our requirements for well-being, care. Briefly, the key concept of IoT is the connection of any "thing" at any "time" from any "place".

The main challenges and concerns of information privacy are depended on collecting, storing, sharing, and analyzing IoT devices' enormous quantity of sensitive and large-scale data. The efficacy of smart city relies on successful usage of large data implementation because IoT systems in smart city generally capture massive information volumes for storage and further analysis [7]. Besides, data management,

network infrastructure, and sophisticated algorithms used to process big data of the smart cities is another concern [5]. Evolving big data issues along with IoT limitations into smart cities where data has been processed, analyzed, shared, transferred, and stored raises privacy concerns. Due to those concerns, people might not accept the idea of smart cities, especially when sensitive data is gathered, such as health records and financial information. Therefore, to get the maximum of an smart city's benefits, security, and confidentiality which involve each layer in the smart cities architecture [8].

IoT equipment can be functioned in diverse areas or sectors, and IoT is the cornerstone of the smart cities infrastructure. Researchers have had an interest in securing and protecting IoT data. Consequently, protection of knowledge desires to be fully demonstrated to guarantee the sustainability of essential facilities such as health services, maintenance, electricity, and sanitation in smart cities. In other hands, everybody is vulnerable to attacks while everything is linked [5]. The user identity of an entity is the capability to identify (or reveal) an individual depends on information collected for instance personal name, home address, or personal contact from a privacy context. The consequence of such a threat is that data protection breaches might be enforced as soon as the client is detected. Vulnerabilities make certain threats possible and aggravate them, for illustration, targeting and tracking people. It also helps information from various sources to relate to the identical stated objective. This data analysis will precisely represent the victim's history. For instance, programmers who can obtain client records might use deep learning methods toward deducing client data and learn the behavior of the victims, that then try to use client behavior to infect users with malware and targeted advertisements.

The widespread of IoT usages which used to enable the set of immense volumes of information utilizing IoT tools which be competent of processed and investigated outside the client's jurisdiction. Consequently, app identity is the primary obstacle to app protection. Increasing IoT implementations significantly increase the risk of user authentication and the hazards associated with it [9]. IoT solutions have advantages and disadvantages unique to themselves. RFID is commonly used for recognizing objects in Internet of Things scenarios, capturing metadata, and controlling specific targets via radio waves [10]. Pateriya et al. [11] shows that hacking, spoofing, traffic exploration, and denial of service attack are the issues of insecure tags. Such insecure tags are accessible to the unauthorized user without reasonable rights to entry. While the tag information might be secured by lightweight security mechanisms, tracking through tag replies is easy to recognize.

The utilization of various methods, involving face and voice recognition, is being introduced in approximately all mobile platforms. Surveillance systems also use integrated cameras and microphones to incorporate facial and voice recognition. An attacker may access a surveillance camera entirely and exploit other devices by sending data to legal server and copying himself. Clients are also a threat to privacy, which leads to other risks such as surveillance and tracking of utility services, approved by the clients Profiling [12] is about collecting and analyzing personal conduct data to measure or ascertain their interest within a particular field or for resolves of discrimination. Parker Higgins [13] tweeted advising customers to not

discuss the important subjects next to Samsung Smart TV. The consumers should be thorough in the collection of information and should always understand the potential scope of violence. Finally, control should be put in the hands of users to make appropriate decisions as to how their private data are processed and exchanged outside of their domain control.

From the previous discussion, IoT can be contemplated as a two-edged sword. If everything is linked to the Internet, providing security to gather information is very important [14]. To accomplish this requirement, NC-DES is proposed as a suitable security system which is going to gather information from IoT systems and then transmit and store the date in encrypted formats so any unauthorized person cannot read the information.

3 The Proposed Lightweight NC-DES Algorithm

NC-DES Algorithm is a proposed lightweight form of the conventional DES. It is presented as a robust, small, and competent encryption algorithm that is appropriate to be used in devices with limited resources, particularly IoT devices. NC-DES collects data from IoT devices, then transfers and stores the data in an encrypted format. NC-DES rely on the traditional DES algorithm. NC-DES employs compact lightweight S-boxes and excludes both the table of initial permutation (IP) and its reverse. NC-DES had proved its resistance to various types of attacks, for instance linear cryptanalysis and differential cryptanalysis outbreaks. NC-DES was designed to overcome earlier research weaknesses and restrictions, including S-boxes trapdoors, memory capacity, speed restriction, look-up records, P-box, key size, complexity, algorithm, and hardware implementations.

3.1 Structure of NC-DES

The symmetric key cryptography algorithm protects the IoT data that are transmitted over the Internet. It is used for encrypting and decrypting data with a 64-bit key. The proposed algorithm takes into account both confidentiality and simplicity. NC-DES is designed based on strengths from tested algorithms and improved to be suitable for constraint devices like that used in IoTs environments. Substitution boxes (S-boxes) of traditional DES were the main weakness in the algorithm that the presence of S-boxes trapdoors threaten the power of the system [9]. Shamir [13] succeeded in defeating the conventional DES using the differential cryptanalysis attacks. Also, Matsui [15] succeeded in reducing the period to crack it using linear cryptanalysis attacks. Rigid substitution boxes beside extreme nonlinearity properties that be exploited to fulfill uncertainty and replace the weak ones of DES. Permutation tables are used to diffuse the data in top of the Feistel structure. Conversely, the compact

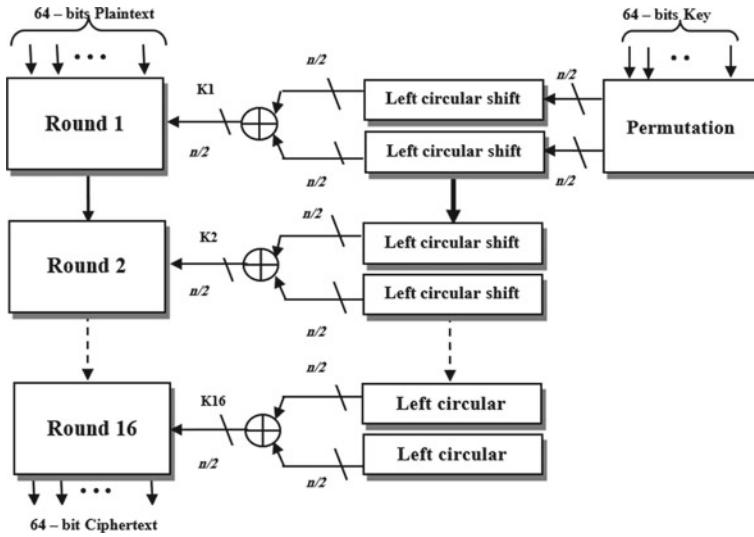


Fig. 1 Overall scheme for NC-DES encryption

size of replacement boxes and the employ of simple internal XORing make it simple to achieve.

3.2 NC-DES Encryption

The encryption function has two inputs: the encryption plaintext and the key as demonstrated in Fig. 1. The plaintext is 64-bits long, and the key also has a length of 64-bit. On the left side of the diagram, the procedure is done in a 16-round step, including Substitution, Permutation, XORing, and Swapping functions. The output of the final round is 64-bit, that is a feature of the ciphertext input plaintext and key. The right-hand side in Fig. 1 illustrates how the 64-bit key is used. The key is obtained through permutation functions. The result of a left circular move then applies XORing function and generates Sub-key K_i for each of the 16 rounds. Due to repeated changes in the key bits, different Sub-key occurs every round.

3.3 Details of Single Round

The organization of single round is divided into four steps:

- Initially, the input of 64-bit is split to dual 32-bit cuts, representing L_i and R_i , correspondingly.

- Second, the right part of the input R_{i-1} and the Sub-key K_i are altered in NC-DES function which is XORing binary operation as stated in Eqs. (1) and (2).

$$L_i = R_{i-1} \quad (1)$$

$$R_i = L_{i-1} \oplus f(R_{i-1} \oplus K_i) \quad (2)$$

- Third, output of the NC-DES function is forwarded to the S-boxes layer, then rearranged using the permutation layer.
- Fourth, the permuted output XORed with the left half L_{i-1} part and turn into the right side as an input of the subsequent round. The input R_{i-1} 's right half became the input's left half in the subsequent round.

The details of the inner design of the NC-DES single rounds are depicted in Fig. 2.

NC-DES was designed to overcome earlier research weaknesses and restrictions, including S-boxes trapdoors, memory capacity, speed restriction, look-up records, P-box, key size, complexity, algorithm, and hardware implementations. The primary

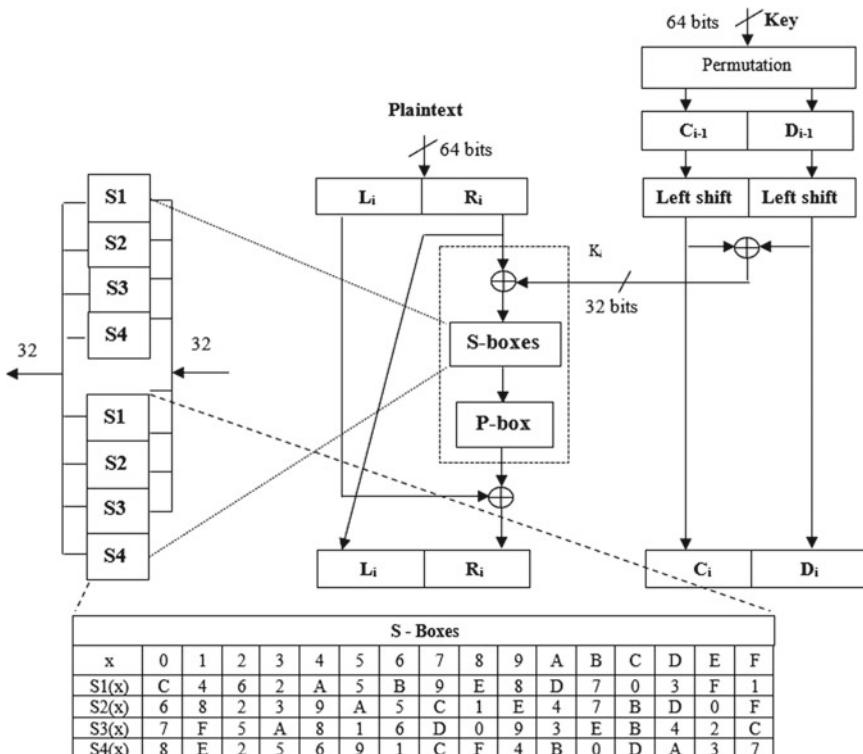


Fig. 2 Details design of a single round

Table 1 The S-boxes of NC-DES algorithm

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S1(x)	C	4	6	3	A	5	B	9	E	8	D	7	F	5	A	7
S2(x)	6	8	2	2	9	A	5	C	1	E	4	6	8	2	2	9
S3(x)	7	F	5	A	8	1	6	D	0	9	3	C	4	6	3	A
S4(x)	8	E	2	5	6	9	1	C	F	4	B	1	6	D	0	9

part in the NC-DES is the NC-DES function, which is composed of two basic layers: Substitution layer and Permutation layer.

Substitution Layer: The basic idea of NC-DES is to replace the eight traditional 6×4 substitution boxes by four compact boxes 4×4 . Each S-box is used twice per round. The new boxes save about 87.5% of memory spaces than the traditional DES. Therefore, the internal round structure requirements are being adapted. The replacement process leads to dropping the Expandable table (E-table) used in traditional DES [7]. E-table is used to expand data and uses both outer bits to select a certain row from the S-box and to replace the original data (before expansion) by a new pattern. This produces enhancement in the function speed.

NC-DES employs four strong, robust 44 S-boxes, labeled S1, S2, S3, and S4. These S-boxes possess the four fundamental strength properties for graphically “strong” S-boxes. These are bijection, nonlinearity, strict avalanche, and output bit independence. These S-boxes were evaluated prior to selection. In fact, a large number of S-boxes have been evaluated in order to choose the four highest protected S-boxes which meet the security measures. When selecting S-bases, protection against various forms of cryptanalysis was considered. Table 1 depicts the designated S-boxes (S1, S2, S3, and S4), respectively.

Permutation Layer: The definition of permutations is a procedure of a group of values that can be arranged in different ways. In this algorithm, permutations represent final phases of NC-DES functions, and are identical to the original DES [15]. The permutation boxes accept 32-bits of data, permutes it according to predetermined rules, and then outputs 32-bit data. The original permutation table of the traditional DES algorithm is utilized.

3.4 Key Generation

For 16 rounds and 64-bit blocks, it is necessary to generate 16 (32-bit) Sub-keys using the 64-bits encryption keys. The creation strategies are as follows. Initially, permutations are employed toward the key. Traditional DES utilizes the 56-bit P-table. Here, the same P-table is adapted to accept 64-bit governed by a Table 2, and then the resulting 64-bit key is allocated to two identical splits of 32-bit represented equally to C_{i-1} and D_{i-1} , correspondingly. Every round left turn of 1 or 2 bit is

Table 2 Permutation (P)

	49	41	33	25	17	9	1
58	50	42	34	26	18	10	2
59	51	43	35	27	19	11	3
60	52	44	36	63	55	47	39
31	23	15	7	62	54	46	38
30	22	14	6	61	53	45	37
29	21	13	5	28	20	12	4
64	56	48	40	32	24	16	8

subject to C_{i-1} and D_{i-1} independently, the same as classic DES [7]. The shifted values are used as input to the next round.

3.5 NC-DES Decryption

The decryption procedure resembles the encryption procedure in every way. NC-DES decipherment is accomplished through entering the ciphertext toward NC-DES algorithm, merely using a numerous variety of keys. The decipherment Sub-keys are obtained as of encrypting Sub-keys and employed to the opposite order.

3.6 Analysis of the Cipher

In cryptographic designs, especially S-boxes, it is significant to assess the power of all of the S-box elements. The required S-boxes properties reflect the immunity of the cryptographic algorithm against the several forms of attacks. Nonlinearity of S-boxes is the most valuable property. Maximum Difference Propagation Probability (DPP_{max}), Maximum Input–Output Correlation (IOC_{max}), and S-box Robustness are all required to accomplish this. To ensure the proposed technique is secure against differential cryptanalysis, linear cryptanalysis, and similar attacks [7, 13], these computations are necessary.

To examine the differential vulnerability of an S-box, the DPP_{max} needs to be calculated. The computation of the DPP_{max} is done by first building the XOR distribution table. S1's XOR distribution (differential distribution) is illustrated in Table 3. Then, we take the maximum value in the XOR distribution table that we've generated and divide it by the total elements 2^n where n values are the bit size of the elements. Therefore, $4/16 = 2^{-2}$ is the maximum DPP for the S1-box.

The probability bias or input–output correlation for each linear approximation denoted in Eq. (3) must be calculated, in this case X_i refers the i th bits of the value X to the S-box and Y_J which refers the J th bits of the outcome Y from the S-box,

Table 3 The differential distribution of S

I/P—O/P XOR difference	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	2	2	2	0	2	0	2	0	2	0	0	0	2	2
2	0	2	2	2	0	0	2	0	0	0	2	0	2	0	0	4
3	0	2	2	2	0	2	0	0	0	2	0	0	2	0	4	0
4	0	2	2	0	0	0	4	0	0	0	0	4	0	2	2	0
5	0	0	0	0	0	0	0	0	4	4	0	0	2	2	2	2
6	0	2	0	0	2	0	0	4	0	2	0	0	2	4	0	0
7	0	0	0	2	0	2	0	4	2	0	0	0	0	0	2	4
8	0	0	2	0	2	2	2	0	2	0	2	2	2	0	0	0
9	0	2	0	0	2	2	2	0	0	2	4	0	0	0	0	2
10	0	2	0	2	2	2	0	0	2	0	4	2	0	0	0	0
11	0	0	0	0	0	0	0	0	2	4	2	4	2	0	2	0
12	0	0	0	2	2	0	2	2	0	2	0	0	2	2	2	0
13	0	0	2	0	4	0	0	2	0	0	0	2	2	2	0	2
14	0	4	2	2	0	4	2	2	0	0	0	0	0	0	0	0
15	0	0	2	2	0	2	0	2	2	0	0	2	0	4	0	0

and \oplus represents bitwise XOR operation.

$$X_{i1} \oplus X_{i2} \oplus \dots \oplus X_{in} \oplus Y_{j1} \oplus Y_{j2} \oplus \dots \oplus Y_{jm} = 0 \quad (3)$$

Table 4 displays a full linear approximation of inputs and outputs, where each entry represents the probability bias of a unique linear combination of inputs and outputs bit on the scale from 0 to 2^n . Then the largest predetermine probability is

$$\pm 4/16 = \pm 2^{-2}.$$

The values of D_{\max} (XOR Dist. Table) are 4 and L_{\max} (Linear App. Table) is ± 4 for each NC-DES S-box. According to the results, all S-boxes have the same vales.

4 The Proposed Remote Patient Monitoring System (RPMS)

Preserving human life is widely regarded as the paramount imperative in the world. The vital need for hospital authorities is the enhancement of a real-time healthcare monitoring framework that can track patients' physical conditions, as stated in [15]. The Remote Patient Monitoring System (RPMS) development has utilized various

Table 4 The linear approximation of S1

8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2	-2	-2	-4	0	0	2	0	2	2	2	0	0	4	2	
0	4	2	0	0	2	0	-4	-2	2	0	2	2	0	-2	2	
0	4	0	-4	0	0	0	0	-4	0	0	0	0	0	0	-4	
0	2	0	0	0	-2	-2	2	-4	0	-4	-2	-2	2	0	2	
0	2	0	2	0	-2	0	-2	2	0	2	-4	2	4	2	0	
0	-2	2	-2	0	-4	-4	-2	0	-2	2	2	0	0	0	0	2
0	0	-2	2	0	2	-2	-4	0	-2	-2	0	-2	-2	4	0	
0	0	2	0	-4	2	-4	0	2	2	0	-2	-2	0	-2	-2	
0	-2	2	-4	0	0	2	-2	-2	2	0	-4	0	-2	2	0	
0	2	0	0	4	2	-2	2	0	0	4	-2	-2	-2	0	2	
0	-2	-4	-2	0	2	0	-2	-2	0	2	0	-2	4	-2	0	
0	0	0	2	-4	0	2	0	-2	-4	2	-2	0	-2	-2	2	
0	0	2	-2	0	2	2	0	4	-2	-2	0	-2	2	0	4	
0	-2	2	0	0	4	-2	2	-2	-2	0	0	4	2	2	0	
0	0	4	2	0	0	2	0	-2	0	2	2	-4	2	2	-2	

technologies, with a primary emphasis on the IoT. This approach facilitates the storage of patient information on cloud-based platforms. In a smart healthcare system, biomedical sensing tools are typically utilized to transmit patient medical data toward the centralized healthcare server control room (CCR) under ordinary circumstances. The health information is transmitted toward the physician through the employment of the Global System for Mobile Communication (GSM) module in situations that are deemed critical [16].

The transportation of personal medical information via insecure networks before its storage in insecure cloud services may result in the compromise of the privacy of sensitive data due to potential cyber-attacks. Furthermore, it is imperative to consider a crucial matter when it is transmitted via IoT devices that have limited resources [15]. The nature of these devices renders traditional algorithms unsuitable. According to [17], cryptographic algorithm is characterized as an effective security measure for IoT systems.

The purpose of the anticipated system is to consistently monitor the various physical conditions of patients that is a significant obstacle encountered by healthcare facilities. The system proposed uses an Arduino UNO microcontroller (μ c) to establish communication with different type of sensors, including Bluetooth (BT) sensor, heart rate (HR) sensor, oxygen saturation (SpO_2) sensor, electrocardiogram (ECG) sensor, and blood pressure (BP) sensor, for the persistence of observing a patient's medical status. Sensing devices possess the capability to produce health-related data of patients and transmit them to a healthcare repository through wireless network infrastructure or to a physician remotely by means of a GSM module.

The Pseudo-code of the anticipated RPMS system is shown in Algorithm 1.

Algorithm 1: Pseudo-code of the anticipated RPMS system

```

1. Initialize Memory, ADC, Input, Output ports, and variables
2. Initialize Bluetooth, GSM modules, and LCD
3. Read the patient's physical parameters
4. If readings for any abnormal conditions Then
    a. Issue alarm signal
    b. Display the abnormal value on LCD
    c. Send AT commands to GSM modem
    d. GSM modem operates in text mode
    e. Contact with the smartphone of the doctor
    f. Send a secure SMS to Doctor's mobile phone includes:
        i. Patient's name
        ii. Urgent parameter value
    g. Close Communication
    h. Return (2)
5. else
    a. Send the patient's physical parameters after the encryption method to CCR via the
    Bluetooth module
    b. CCR uploads the patient's parameters value to be accessed by authorized persons
    c. Close Communication
    d. Return (2)

```

The block diagram of the anticipated RPMS monitoring system is shown in Fig. 3. The anticipated system is composed of the Arduino UNO microcontroller, which is linked to the five sensors BT sensor, HR sensor, SpO₂ sensor, ECG sensor, and BP sensor that measures the condition of the patient. The information is collected through sensors at the microcontroller then match the findings with the stored normal data. If the reading is normal, the patient's data is encrypted employing NC-DES and forwarded by the connected Bluetooth module to the database server to store the data. If the gathered data exceeds or less than the normal range, patient's data is encrypted employing NC-DES and forwarded by the connected GSM and send an emergency message to doctor's mobile. Besides, the control signal will be sent to the buzzer and the abnormal reading value of the physical parameter is forward to Liquid Crystal Display (LCD).

The detail of the proposed RPMS system are described as following:

Sensing Element: The sensors are used to measure several physical parameters, then the values are transferred toward the microcontroller. Generally, several health parameters can be measured, but the anticipated system concentrates on measuring five important parameters. The following patient health sensors have been incorporated into the proposed system to monitor patients' health. **BT Sensor:** Body Temperature Sensor is a device precisely designed to change its own properties in compliance with temperature of the environment [18]. **ECG Sensor:** (Electrocardiogram sensor) the electrical cardiac rate is registered throughout a specified period of time, and the small electrical variations are observed throughout each heartbeat [18]. **HR Sensor:** (Heart Rate Sensor) is an electronic equipment used to detect heart beat pulse. Heart rate is the per-minute amount of heartbeats usually registered as BPMs [18]. **BP Sensor:** The sensor for measuring pulse and oxygen levels in the blood. The

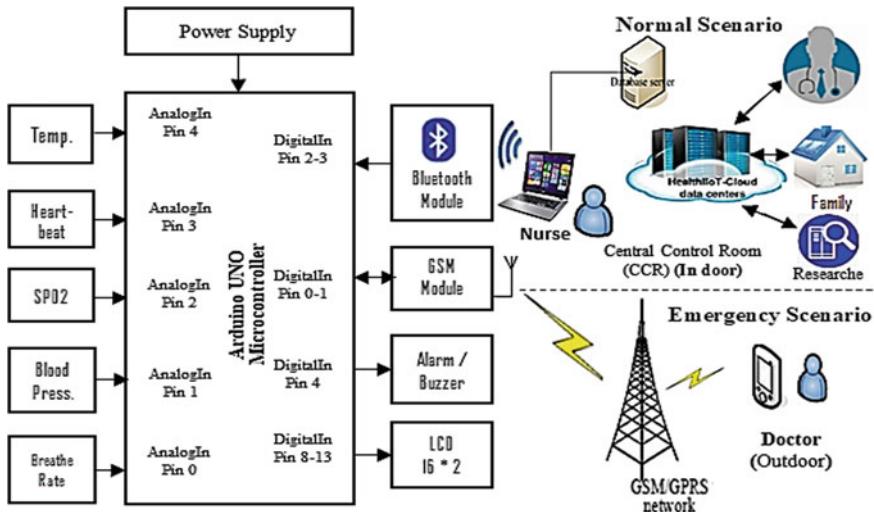


Fig. 3 Block diagrams of the anticipated RPMS system

technique of pulse oximetry is a straightforward method for monitoring the level of oxygen saturation in hemoglobin [18]. **SPO₂ Sensor:** (Pulse and oxygen in blood sensor) is an electronic equipment used to measure the saturation of oxygen in the blood [18]. There are other sensors to keep the sensors at their positions. Figure 4 illustrates the Body Area Network (BAN) and connect the sensors to the patients and the kit prototype which collects data from the sensors.

Control Element: The data gathered by the sensor network are forwarded to the controller that is the brainpower of the device, and it is accountable to make quick decision [19]. Controller receives the input signals from the sensing devices and compares it by the reference (normal) range. If the values are in the normal range, patient's information is encrypted using NC-DES and forwarded to the Bluetooth module. On the contrary, if abnormally changes of patient physiology condition occur (signs values exceed the normal range), security is also attainable by transmitting the data encrypted using NC-DES through the GSM. Besides, a control signal will be sent to the buzzer and the abnormal reading value of the physical parameter is forward to Liquid Crystal Display (LCD).

Controlled Element: In typical cases, the Bluetooth module will receive a control signal from the control element. Then, it will forward the secure patient's health information to CCR. Once CCR receives the encrypted message, the data sent to be stored in a database system where only authorized persons can retrieve the patient information at any moment [19]. CCR uploads the data to the server as well. In an emergency, the GSM module will receive a control signal from the control element. Then it will send an SOS message to a specialist physician.

Output Element: In an emergency, an alarm receives a control signal coming out of the controller then it produces continues ringing. Also, LCD receives a control

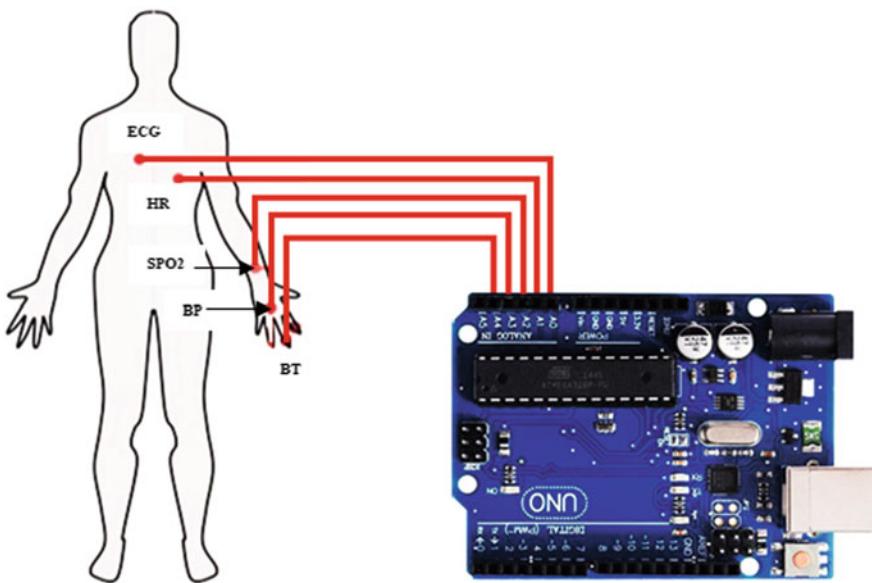


Fig. 4 Body area network and the prototype kit

signal coming out of the controller just before displaying the abnormal reading rate of the physical factor.

The proposed Remote Patient Monitoring System (RPMS) offers healthcare professionals the capability to monitor patient conditions in real-time. Aforementioned, information can be readily accessed by medical professionals at any given moment. The secure transfer of personal data to its intended destination and restricting access solely to authorized individuals constituted a significant challenge in conventional RPMS. Nevertheless, the anticipated system under consideration effectively addresses this issue by offering safeguarding measures for the conveyed data.

5 Conclusion and Future Work

The evolution, progression, and enhancement of intelligent urban areas are founded upon the technological paradigm shift, with a particular emphasis on IoT. The employment of intelligent medical systems is a crucial aspect of smart city improvement. Security concerns are among the most perilous challenges confronting the evolution of smart cities and remote observing systems for patients. This paper introduces a proposed system aimed at addressing security issues of smart cities. The proposed solution involves the implementation of a New Compact-Data Encryption Standard (NC-DES) to mitigate these concerns. Moreover, the healthcare sector

has been chosen as the domain to introduce a secure application for the point of monitoring patient health parameters. The proposal put forth a remote healthcare system, commonly represented as RPMS, which facilitates the secure transfer of measured parameters. The NC-DES algorithms have demonstrated their proficiency to offer security to Internet of Things and smart cities, especially in the perspective of healthcare systems.

References

1. Park E, Del Pobil AP, Kwon S (2018) The role of Internet of Things (IoT) in smart cities: technology roadmap-oriented approaches. *Sustainability* 10(5)
2. Meijer A, Rodríguez Bolívar MP (2016) Governing the smart city. *Int Rev Adm Sci* 82(2):392–408
3. Aboshosha BW, Dessouky MM, Ramadan RR, El-Sayed A (2019) LCA-lightweight cryptographic algorithm for IoT constraint resources. *Menoufia J Electron Eng Res* 15:28, pp 374–380
4. Sterbenz JPG (2017) Smart city and IoT resilience, survivability, and disruption tolerance: challenges, modelling, and a survey of research opportunities. In: 2017 9th international workshop on resilient networks design and modeling (RNDM). IEEE, pp 1–6
5. Weinberg BD, Milne GR, Andonova YG, Hajjat FM (2015) Internet of things: convenience versus privacy and secrecy. *Bus Horiz* 58(6):615–624
6. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. *Comput Netw* 54(15):2787–2805
7. De Montjoye YA (2015) Computational privacy: towards privacy-conscious uses of metadata. PhD diss., Massachusetts Institute of Technology
8. Mavragani A, Ochoa G (2019) Google Trends in infodemiology and infoveillance: methodology framework. *JMIR Public Health Surveill* 5(2)
9. Aboshosha BW, Dessouky MM, El-Sayed A (2019) Energy efficient encryption algorithm for low resources devices. The Academic Research Community Publication
10. Kim J, Yoo SK (2013) Design of real-time vital-sign encryption module for personal healthcare device. *J Inst Electron Inf Eng* (2)
11. Pateriya RK, Sharma S (2011) The evolution of RFID security and privacy. In: Conference on communication systems and network technologies. IEEE, pp 115–119
12. Huang GT (2004) Emerging technologies that will change your world. *Technol Rev* 107(1):32–40
13. Biham E, Shamir A (1991) Differential cryptanalysis of DES-like cryptosystems. *J Cryptol* 4(1):3–72
14. Romero-Mariona J, Hallman R, Kline M, San Miguel J, Major M, Kerr L (2016) Security in the industrial internet of things—the C-SEC approach. In: International conference on internet of things and big data, vol 2, pp 421–428
15. Matsui M (1994) Linear cryptanalysis of DES Cipher. In: Hellenseth T (ed) *Advances in cryptology—EUROCRYPT '93*, vol 765. Springer, pp 286–397
16. Standard, Data Encryption. “Federal information processing standards publication 46.” National Bureau of Standards, US Department of Commerce 23 (1977)
17. Kim J, Jin Lee B, Yoo SK (2013) Design of real-time encryption module for secure data protection of wearable healthcare devices. In: 35th international conference of the IEEE engineering in medicine and biology society, pp 2283–2286
18. Butt TA, Afzaal M (2019) Security and privacy in smart cities: issues and current solutions. In: *Smart technologies and innovation*. Springer, Cham, pp 317–323

19. Xu B, Da Xu L, Cai H, Xie C, Hu J, Bu F (2014) Ubiquitous data accessing method in IoT-based information system for emergency medical services. *IEEE Trans Ind Inf* 10(2):1578–1586
20. Sukanesh R, Vijayprasath S, Subathra P (2010) GSM based ECG tele-alert system. In: 2010 international conference on computing technologies. IEEE, pp 1–5

Design, Development, and Mathematical Modelling of Hexacopter



Vishwas Mishra, Priyank Sharma, Abhishek Kumar, and Shyam Akashe

Abstract The present achievement of Unmanned Aircraft System [UAS] innovation has led to it being the answer to every issue that arises in daily life. Hexacopter technology is now expanding quickly across a wide range of application areas. Development, design, and evaluation of hexacopters are carried out in the work using previously acquired actual elements. Drone autonomy, beyond visual line of sight control, and practical approaches for hexacopter platforms are all covered in depth, as are the mathematical modelling, controller, and command and control system design for multicopter drones in general, with an emphasis on hexacopter systems. The mechanical foundation of a physically constructed hexacopter, together with elements of matrix-based mathematical modelling principle, is presented, along with a video of the drone in flight.

Keywords Hexacopter · Material · Speed · UAV

1 Introduction

DRONE is a catch-all term for a wide range of airborne, terrestrial, aquatic, and subaqueous platforms [1]. Dynamic Remotely Operated Navigation Equipment represents one of the meanings given for the abbreviation DRONE, which originates from

V. Mishra (✉)

Department of EEE, Swami Vivekananda Subharti University, Meerut, India
e-mail: vishwasmishra88@gmail.com

P. Sharma

Department of ECE, MIET Meerut, Meerut, India
e-mail: priyank.sharma@miet.ac.in

A. Kumar

Department of ECE, NIT Jamshedpur, Jharkhand, India

S. Akashe

Department of ECE, ITM University, Gwalior, India
e-mail: shyam.akashe@itmuniiversity.ac.in

the English spoken language. The DRONE family includes the main types of vehicles listed below: unmanned ground vehicle (UGV), unmanned underwater vehicle (UUV), and unmanned aerial vehicle (UAV) are three examples [2]. Because of its many real-world uses in fields including firefighting, aerial imagery, agriculture, tracking, and rescue operations, unmanned aerial vehicles (UAVs) are growing in popularity in the scientific community. By the amount of propulsion motors, the multirotors included in the wing rotor group are divided into quadcopter, helicopter, hexacopter, tricopter, octocopter, etc., categories [3, 4]. Numerous researchers have studied topics including modelling of tricopters, autonomous helicopter control, control of quadcopter movement in terms of attitude and altitude, modelling and control of hexacopters, and octocopters' navigation. The PID controllers, back stepping, feedback linearization, sliding mode controllers, and neural network are the control types employed in this multirotor investigation [3]. Their widespread development and use have benefited a variety of civilian applications, including but not limited to: illegal hunting detection, law enforcement surveillance missions, landslide measurement, firefighter support, border and strategic target security, crowd incident tracking, large buildings and structures, pipelines for oil and gas, examination of large manufacturing plants, examination of continuous-flow machinery in mines, and many more. Beginning with the gimbal with three-axis stabilisation and the photo and video camera, this paper then moves on to the avionics elements for command and control and the radio control used by the operator on the ground to issue commands to the hexacopter. The second part of this article describes the mathematical relations presenting the drone's motions in three-dimensional space, forces, and aerodynamic moments during flight and adapts the mathematical modelling theory based on matrix formalisation to a hexacopter [4] drone with the rotors mounted in a plane parallel to the ground. The current research uses the rotational velocities of the hexacopter's blades as the sole controller [5, 6]

2 Mathematical Model of the Hexacopter Developed

2.1 *Structure of the Hexacopter*

Due to its under-actuation, dynamic instability, and six-degrees-of-freedom system, a multicopter drone needs flight stability control. Three-dimensional translation and rotation are included among the six degrees of freedom [7]. If the propeller is to move in a direction perpendicular to the pull of gravity, the thrust must be adjusted in both amplitude and direction. Fixed rotor blades may be rotated in a way that tilts the thrust vector by generating torques near the centre of rotation via independent adjustments to the propeller speed. The hexacopter is made out of carbon fibre and has a symmetrical frame with three sets of clockwise (CW) and counterclockwise (CCW) propellers attached to it, all of which are kept in place by six motors attached

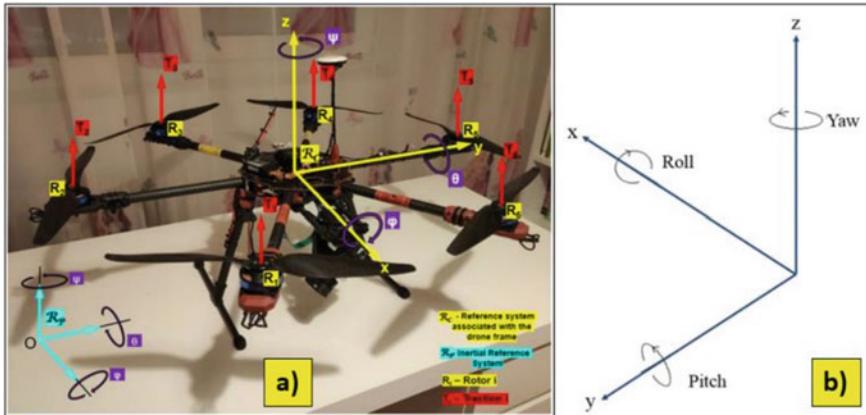


Fig. 1 **a** Drone's inertial and related coordinate systems; **b** rotational movement along the axes of x , y , and z

to six support arms grouped in groups of three (R_1, R_3, R_5, R_6) and at an angle of 120° to each other [8, 9].

2.2 Coordinate Systems of the Drone

Figure 1 from the paper Stamate et al. shows the two coordinate systems are used for expressing the orientation and position of the hexacopter in three dimensions during analysis: the ground-based (inertial) system and the drone-based (frame-based) system [10, 11]. Three rotors must revolve CW, while the other three rotate trigonometrically in order to maintain a stable position (hover) for the hexacopter [12].

2.3 Drone's Point of View and Axis of Rotation

The orientation of the drone with regard to its own inertial coordinate system is referred to as its attitude. It illustrates the movement of the drone in all three coordinate systems (x , y , and z). Figure 1b is an illustration of how the right-hand rule should be used, which ultimately results in the three basic aircraft motions of roll, pitch, and yaw.

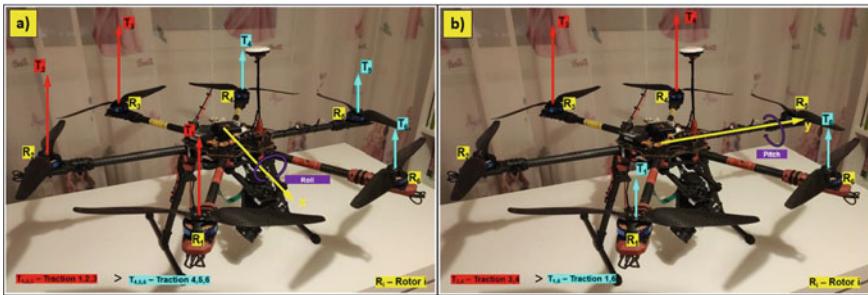


Fig. 2 **a** Action like rolling; **b** pitch motion

2.3.1 Roll

It is the result of changing the rotational speeds of rotors 1, 2, and 3 and rotors 4, 5, and 6 in unison to provide a rotating motion around the x -axis. This action induces an angular acceleration by producing a rotating torque around the x -axis. As can be seen in Fig. 2, the rolling motion's angular velocity is defined by and expressed in radians per second [13–15].

2.3.2 Pitch

To accomplish this rotational motion around the y -axis, the speeds of rotors 1, 6, 3, and 4 are individually and collectively altered. Rotor positions 2 and 5 have no effect on pitch since they are perpendicular to the y -axis. Figure 2b illustrates that the pitch angle, represented by, is quantified in radians per second.

2.3.3 Yaw

In other words, it has rotation around the z -axis. Each propeller in this movement generates a torque in the z -axis direction as it spins. Since the rotor rotates CCW, this torque moves in the other direction. A trigonometric rotation around the z -axis is generated if the propeller is spun clockwise. Rotors 2, 4, and 6 rotate at different rates than rotors 1, 3, and 5, which creates the spinning motion. Figure 3 depicts the angular velocity of the gyration motion as a rotational angle defined by and expressed in radians per second.

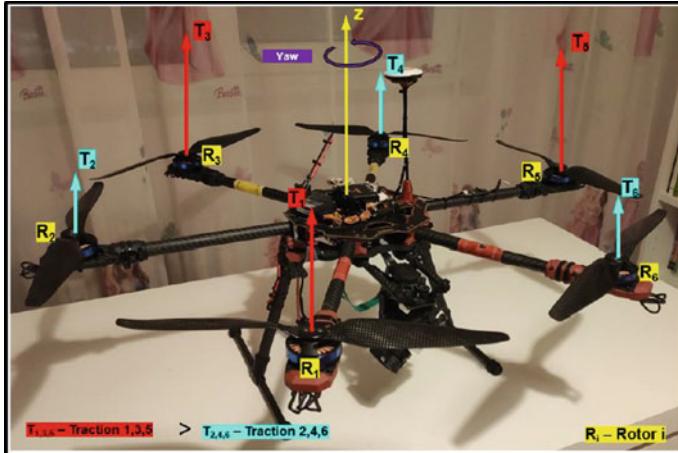


Fig. 3 Yaw motion (ψ)

2.4 Rotation Matrices

Defining the drone's motion with regard to a fixed object is necessary since the inertial coordinate system is a fixed reference frame. The sensors on an RC drone can only work if the drone is flown with the x -axis pointing forward, the y -axis pointing to the left, and the z -axis pointing upward. Drone frame (RC) coordinate rotations with respect to the ground (RP) inertial reference system will be represented using the rotation matrix, a popular transformation technique. A rotation matrix is a matrix that specifies the angle at which the drone will rotate around each of its three axes [16, 17].

2.4.1 Yaw Rotation Matrix (z-axis)

$$R_C^P(\varphi) = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

2.4.2 Pitch Rotation Matrix (y-axis)

$$R_C^P(\theta) = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}. \quad (2)$$

2.4.3 Roll Rotation Matrix (x-axis)

$$R_C^P(\Phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \Phi & \sin \Phi \\ 0 & -\sin \Phi & \cos \Phi \end{bmatrix}. \quad (3)$$

The rotation matrix for the drone's inertial frame is produced R_C^P by conducting the three rotations in the order described above.

$$R_C^P = R_C^P(\Psi)R_C^P(\theta)R_C^P(\Phi), \quad (4)$$

$$R_C^P = \begin{bmatrix} \cos \varphi & \cos \theta & \cos \varphi & \sin \theta \sin \Phi - \sin \varphi & \cos \Phi & \cos \varphi & \sin \theta & \cos \Phi + \sin \varphi & \sin \Phi \\ \sin \varphi & \cos \theta & \sin \varphi & \sin \theta \sin \Phi + \cos \varphi & \cos \theta & \sin \varphi & \sin \theta & \cos \Phi - \cos \varphi & \sin \Phi \\ -\sin \theta & \cos \theta & \sin \Phi & \cos \theta & \sin \Phi & \cos \theta & \cos \theta & \cos \theta & \cos \Phi \end{bmatrix}. \quad (5)$$

Since the inverse matrix of an orthogonal matrix (R_C^P) is identical to its transpose, this matrix may be used to transfer rotation matrices between the earth (inertial) frame and the drone frame.

$$(R_C^P)^{-1} = (R_C^P)^T = R_P^C. \quad (6)$$

2.5 Motion Equation of the Hexacopter

For the purpose of this discussion, the hexacopter will be modelled as a hard, symmetrical solid with the drone's centre of gravity at its geometrical centre. Keeping this in mind, the dynamics of a rigid solid subject to external aerodynamic forces and moments have been described using the Newton–Euler equation. Consistently, the following relationships will describe the forces \mathbf{F}_C [Nm] and moments τ_C (acting on the drone frame [Nm]). Both gravity (G) and the thrust force generated by the rotors' rotation (due to the entrainment of air currents) act on drones. Another aspect working against the multicopter's forward or upward motion is drag force, or friction with atmospheric air [15–18]. Force of gravity always acts downward and along the z -axis; its expression is as follows:

$$\mathbf{F}_{\text{gravity}}^C = R_P^C \begin{bmatrix} mg \sin \theta \\ -mg \cos \theta \sin \Phi \\ -mg \cos \theta \cos \Phi \end{bmatrix}. \quad (7)$$

2.6 Forces Acting on the Hexacopter

The hexacopter's ability to hover as well as fly horizontal is made possible by traction or lift force. This thrust force may be estimated by the following expression while performing the hover manoeuvre:

$$F_{\text{traction}}^C = b \sum_{i=1}^6 \Omega_i^2, \quad (8)$$

where b is traction constant, measured in Ns^2 .

A drone's frame experiences drag while in flight to prevent motion. The following equation explains the influence of this force on the x and y speed variations experienced during the constant-altitude flight manoeuvre:

$$F_{\text{Drag}}^C = \begin{bmatrix} -\mu u \\ -\mu v \\ 0 \end{bmatrix}, \quad (9)$$

where μ is constant which is calculated in kg/s .

2.7 Moments Acting on the Hexacopter

Roll, pitch, and gyration motions may be accomplished by altering the rotors' speed in order to get rotational moments along the axes of x , y , and z . Figure 4 illustrates how the motor support arms are positioned with respect to the drone's centre of gravity, or the distance between the rotor as well as the axis of rotation [16–18]. In this Fig. 4, $l[\text{m}]$ is a measure of how far out from the motor the propeller assembly extends, while $d (\text{Nm} \cdot \text{s}^2)$ denotes the drag factor.

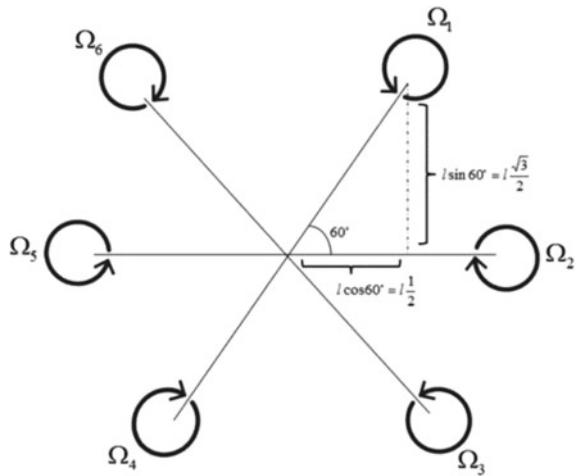
Increasing of $\Omega_4, \Omega_5, \Omega_6$ and decreasing of $\Omega_1, \Omega_2, \Omega_3$ will provide a favourable rolling moment.

2.8 Motor Speed Loss Calculation

Internal resistance in the electric motor windings drains some of the battery's power when the hexacopter is in the air [19]. As a result, the motors are less efficient and operate slower than usual. At a supply voltage of 14.8 V, the optimal relation yields the motor's maximum speed:

$$\text{rpm}_{\text{ideal}} = KV * U_i = 620 * 14.8 = 9176 \text{ RPM}. \quad (10)$$

Fig. 4 Rotor separations from the drone's centre of gravity



The connection that describes the motor speed in the actual world is:

$$rpm_{real} = KV(Ui - Up) = 620(14.8 \text{ V} - 1.746 \text{ V}) = 8082.32 \text{ RPM}. \quad (11)$$

Therefore, the loss in speed equals:

$$rpm_{real} - rpm_{ideal} = 9176 - 8082.32 = 1093.68 \text{ RPM}. \quad (12)$$

2.9 Airflow Velocity Calculation on the Rotor Blade Profile

Using above equation, $rpm_{real} = 8082.32 \text{ RPM} = rps_{real} = 134.7053 \text{ rot/s}$.
Calculation the angular velocity of the rotor:

$$\omega = rps_{real} * 2 * \pi = 134.7053 \text{ rot/s} * 2\pi = 846.3786 \text{ rot/s}. \quad (13)$$

2.9.1 Pitch Angle Calculation

The relationship is used to determine the appropriate pitch angle:

$$\begin{aligned} \varphi &= \text{atan}\left(\frac{H_{\text{pitch.angle}}}{2 \cdot \pi \cdot 0.75 \cdot R_{\text{blade}}}\right) \\ &= \text{atan}(0.1397 / (2 \cdot \pi \cdot 0.75 \cdot 0.1651)) = 10.1795^\circ. \end{aligned} \quad (14)$$

2.9.2 Actual Pitch Angle Calculation

The relation is used to determine the actual pitch:

$$H_r = 2 \cdot \pi \cdot r \cdot t_g(9.6) = 0.1113 \text{ m}, \quad (15)$$

$$\varphi = \text{atan} \left(\frac{H_r}{2 \cdot \pi \cdot 0.75 \cdot R_{\text{blade}}} \right) = 8.1417^\circ. \quad (16)$$

3 Material for Making Hexacopter

Knowing the materials utilised is crucial before researching the actions of hexacopter. In order to create the most accurate finite element model and analyse it to get a general understanding of material behaviour, the investigation of the materials utilised is the first and most fundamental stage in the study of materials [19, 20]. Composite materials make up the majority of unmanned aerial vehicles. Carbon T300/PR319 was the material used for the study of the hexacopter under consideration. Epoxy resin with a fibre volume percentage of 55% and 175 °C curing is used to strengthen the material. The composite possesses the mechanical features listed in Table 1 for composites [21]:

Table 1 Physical properties of composite materials for hexacopter

Properties	T300/PR319 (Toray)
Modulus of elasticity, Young, along the longitudinal axis (E_1)	123 GPa
Directional transverse Young's modulus (E_2)	5.72 GPa
Modulus of elasticity of young in the thickness direction (E_3)	5.72 GPa
Poisson's ratio (ν_{12})	0.35
Poisson's ratio (ν_{23})	0.55
Poisson's ratio (ν_{13})	0.35
Shear module (G_{12})	0.35 GPa
Shear module (G_{23})	1.85 GPa
Shear module (G_{13})	1.35 GPa
Density (ρ)	1565 kg/m ³

4 Conclusion

The dynamic features of the hexacopter built in this research are analysed and evaluated to further explore the possibilities of improving the constructional and functional aspects. Its equations of motion only calculate the moments and forces acting on the aircraft while it is hovering in place, regardless of the direction or speed of the wind. The paper also investigates a mathematical model based on matrix formalisation, which yields the equations of motion and rotor dynamics for the hexacopter. Results from numerical assessment may vary from an analytical result by as much as 15%. The model of the take-off and landing procedure was represented in the analytical computation as a tube constricted at one end with a weight at the other, which is one of the reasons why the calculations were inaccurate. The insufficient description of the content due to a lack of context is the root cause of the discrepancy in results. Due to the large number of input variables, it was found that the design of hexacopters may be enhanced, particularly in terms of mass, by reducing the number of laminate layers inside the composite material. Consideration of these criteria may guide future enhancements to the artificial structure.

References

1. Custers B (2016) Future of drone use. TMC Asser Press, The Hague, pp 3–20
2. Bershadsky D, Haviland S (2016) Electric multirotor propulsion system sizing for performance prediction and design optimization. In: AIAA SciTech, 4–8 Jan 2016, San Diego, California, USA, 57th AIAA/ASCE/AHS/ASC structures, structural dynamics, and materials conference. <http://arc.aiaa.org>, <https://doi.org/10.2514/6.2016-0581>
3. Biczyski M, Sehab R, Whidborne JF, Krebs G, Luk P (2020) Multirotor sizing methodology with flight time estimation. Hindawi J Adv Transp 14. Article ID 9689604. <https://doi.org/10.1155/2020/9689604>
4. Gatti M (2017) Complete preliminary design methodology for electric multirotor. J Aerosp Eng ASCE. ISSN 0893-1321. [https://doi.org/10.1061/\(ASCE\)AS.1943-5525.0000752](https://doi.org/10.1061/(ASCE)AS.1943-5525.0000752)
5. Dai X, Quan Q, Ren J, Cai KY (2018) An analytical design optimization method for electric propulsion systems of multicopter UAVs with desired hovering endurance. IEEE/ASME Trans Mechatron 1083–4435. <https://doi.org/10.1109/TMECH.2019.2890901>
6. Dai X, Quan Q, Ren J, Cai KY (2018) Efficiency optimization and component selection for propulsion systems of electric multicopters. IEEE Trans Ind Electron 0278–0046. <https://doi.org/10.1109/TIE.2018.2885715>
7. Balaji S, Prabhagaran P, Vijayanand R, Senthil Kumar M, Raj Kumar R (2020) Comparative computational analysis on high stable hexacopter for long range applications. In: Proceedings of UASG 2019: unmanned aerial system in geomatics 1. Springer International Publishing, pp 369–391
8. Fogelberg J (2013) Navigation and autonomous control of a hexacopter in indoor environments. Lund University, Department of Automatic Control
9. Lei Y, Cheng M (2019) Aerodynamic performance of hex-rotor UAV considering the horizontal airflow. Appl Sci MDPI (9):4797. <https://doi.org/10.3390/app9224797>
10. Stamate MA, Niculescu AF, Pupăză C (2017) Mathematical model of a multi-rotor drone prototype and calculation algorithm for motor selection. Proc Manuf Syst 12(3):119–128. ISSN 2067-9238

11. Stamate MA, Pupăză C, Nicolescu FA, Moldoveanu CE (2023) Improvement of hexacopter UAVs attitude parameters employing control and decision support systems. *Sens Spec Issue Adv Intell Control Robots* 23(3):1446. <https://doi.org/10.3390/s23031446>. IF 3.847, Q2
12. Courant R, Friedrichs K, Lewy H (1967) On the partial difference equations of mathematical physics. *IBM J Res Dev* 11(2):215234. <https://doi.org/10.1147/rd.112.0215>. Bibcode:1967IBMJ...11..215C, MR 0213764, Zbl 0145.40402
13. Artale V, Milazzo CLR, Ricciardello A (2013) Mathematical modeling of hexacopter. *Appl Math Sci* 7(97):4805–4811
14. Shelare S, Belkhode P, Nikam KC, Yelamasetti B, Gajbhiye T (2023) A payload-based detail study on design and simulation of hexacopter drone. *Int J Interact Des Manuf (IJIDeM)* 1–18
15. Sharipov D, Abdullaev Z, Tazhiev Z, Khafizov O (2019) Implementation of a mathematical model of a hexacopter control system. In: 2019 International conference on information science and communications technologies (ICISCT). IEEE, pp 1–5
16. Alaimo A, Artale V, Milazzo C, Ricciardello A, Trefiletti LUCA (2013) Mathematical modeling and control of a hexacopter. In: 2013 International conference on unmanned aircraft systems (ICUAS). IEEE, pp 1043–1050
17. Suprapto BY, Heryanto A, Suprijono H, Muliadi J, Kusumoputro B (2017) Design and development of heavy-lift hexacopter for heavy payload. In: 2017 International seminar on application for technology of information and communication (iSemantic), pp 242–247. <https://doi.org/10.1109/isemantic.2017.8251877>
18. Husaini H, Putra DIO, Syahriza S, Akhyar A (2023) Stress and strain analysis of UAV hexacopter frame using finite element method. In: AIP conference proceedings, vol 2613, no 1. AIP Publishing
19. Muminovic AJ, Saric I, Mesic E, Pervan N, Delic M (2019) Research about characteristics of designs from industrial designers and product designers. *Periodicals Eng Nat Sci* 7(2):860–869
20. Saric I, Mesic A, Delic M (2021) Hexacopter design and analysis. In International conference “new technologies, development and applications”. Springer International Publishing, Cham, pp 74–81

A Keypoint-Based Technique for Detecting the Copy Move Forgery in Digital Images



Kaleemur Rehman and Saiful Islam

Abstract A decade of research has been conducted on detecting copy-move forgeries (CMFD). Technology has enabled the manipulation of images, once the most authentic source of information. This paper proposes a copy-move forgery detection algorithm based on fused features to address issues such as time complexity and difficulty detecting forgeries in smooth regions. To extract descriptive features, a low-contrast threshold was used in conjunction with three detection methods, including scale-invariant feature transform (SIFT), speeded-up robust features (SURF), and accelerated KAZE (AKAZE). SURF and accelerated KAZE (AKAZE) are used in our keypoint-based CMFD technique. To detect manipulated regions efficiently, AKZAE, SURF, and SIFT can be used to extract major keypoints in smooth regions.

Keywords Communication networks · Performance enhancement · Resource scheduling · Scheduling methods

1 Introduction

Today, powerful image editing and processing software makes it easy to fake digital images. Copies (or copy-moves) are commonly used to fabricate digital images. An image is manipulated when copied, moved, or pasted from one location to another. Copy-move forgeries' detection (CMFD) has been developed several times over the past few years. Block-based and keypoint-based traditional methods can be distinguished. At least one region of the same image is copied and pasted into another region. During copy-move forgery, information within an image is concealed or suppressed. Copy-move forged documents are illustrated in Fig. 1. Images (a) and (b) present original images, while images (c) and (d) present fake images. A duplicate of the textured wall is shown in Fig. 1c, which was attached over the wall's base. In Figs. 1d, a tree enables a building to be hidden. In CMFD, overlapping

K. Rehman (✉) · S. Islam
ZHCET, AMU, Aligarh, UP 202001, India
e-mail: kaleemap14@gmail.com

blocks of fixed size are divided into CMFD blocks, and similar blocks are found in pairs. There is a significant difference between these methods in how the blocks are described. In [1], quantized DCT coefficients represent image blocks. Copy-move forgeries may be detected using DCT-based methods [2]. A reduced-dimension representation of the image block was created by Popescu and Farid [3] and Luo et al. [4] to calculate the average intensity of pixels in each RGB channel. Each block's feature representation was generated by calculating 24 blur-invariant moments [5]. In [6], block features were obtained through Fourier–Mellin Transforms (FMTs). The image was decomposed into four sub-bands using Discrete Wavelet Transforms (DWTs) and Singular Value Decomposition (SVD) [7]. FT correlation coefficients are used by Bravo-Solorio and Nandi [8] to measure the similarity between blocks. Upon reaching a threshold of entropy, the block is excluded. Using rotation-invariant uniform local binary patterns (LBPs) [9], features were extracted from circular blocks [10]. Detecting duplicated regions with Zernike moments was done by Ryu [11, 12]. Using the Patch Match algorithm [13] examined how the algorithm could be modified to handle rotation operations.

Block-based and keypoint-based detection methods are commonly used to detect copy-move forgery (CMFD). CMFD approaches commonly follow the pipeline shown in Fig. 2 in the literature. CMFD techniques are currently covered in four published surveys. CMFD feature extraction techniques are discussed. Their dataset was then used to evaluate the feature extraction techniques. A comparison of brute

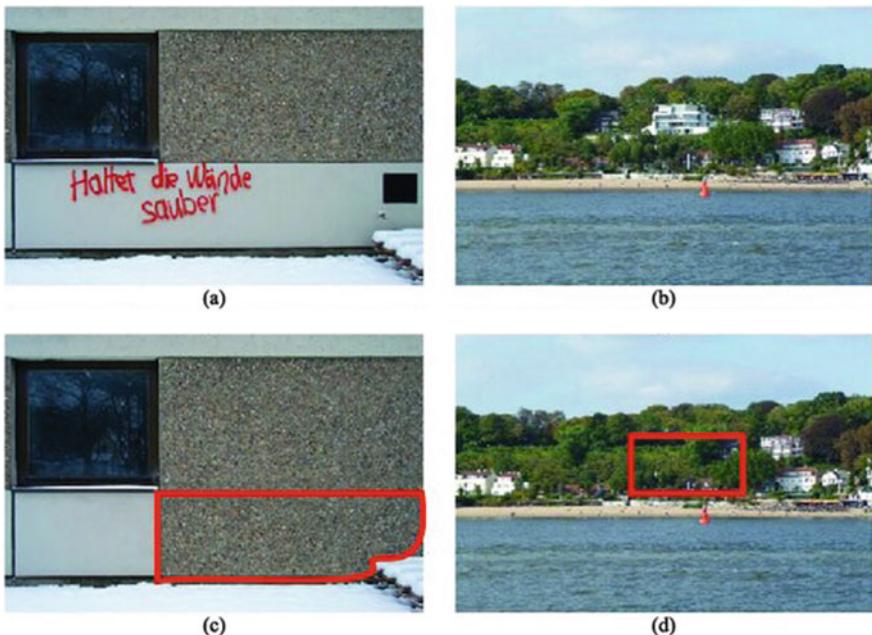


Fig. 1 Example of copy-move forgery **a, b** The original images. **c, d** The tampered images



Fig. 2 Detection of copy-move forged content using a common pipeline

force and block-based matching techniques in CMFD was done by Qazi et al. [14]. An overview of the entire CMFD process is discussed in the paper, including feature extraction and matching techniques. There are also categories of copied regions and domains associated with CMFD analysis, along with datasets and validations.

2 Literature Survey

Detection of tampered images is of utmost importance because tampered images can manipulate the authentic information shown by the image. Such manipulations can bring up devastating consequences by showcasing a false scene. Identifying whether an image has been tampered with using copy-move techniques can be difficult because the altered area has similar characteristics to the rest of the image. The manipulated image is also subjected to various complex operations to ensure no traces of forgery remain. Such operations applied over the forged images make it much more complicated to observe the authenticity of images through the naked eye. The original information can be hidden by copying and pasting tampered images over the original. Copy-move forged images can contain several duplicated objects within tampered images originally not present in the altered image. Consequently, several methods have been proposed to verify the authenticity of digital images in the field of digital image forensics. Copy-move forgery detection methods can be broadly divided into three types based on how they utilize different techniques: block-based, keypoint-based, and hybrid.

The image was divided using fixed dimensions and overlapping blocks and corresponding feature vectors for the image blocks were obtained through PCA. They have stored the feature vectors in the feature matrix. Further, they have employed lexicographical sorting over the feature matrix. The feature vectors present in neighbouring positions within the feature matrix are similar. The image blocks corresponding to similar feature vectors are utilized in the succeeding operations. Distance-based offsets allow identifying the tampered region by identifying multiple blocks with similar feature vectors at neighbouring locations. Author [4] forgeries involving copy-move are detected with a method proposed for detecting colour images. Squares overlapping each other have been used to divide the image. Feature blocks of image blocks are represented using RGB colour components and spatial directional intensity ratios. Feature vectors of dimension 1×7 have been utilized in their method. They have stored the block features in a feature matrix. Further, they have applied lexicographical sorting over the feature matrix. They have identified similar feature vectors by computing the difference between feature vector values. Feature vectors

with difference values less than the user-defined thresholds are similar feature vectors. Further, they have computed shift vectors for the image blocks corresponding to similar feature vectors. A user-defined threshold value is imposed over the shift vectors to identify a group of duplicated tampered blocks present in the different locations of the image.

Further, they have applied an opening operation to remove the isolated image blocks. They have utilized hole-filling operations to localize the tampered regions within forged images. Although tampered images can be manipulated by blurring, adding noise, or compressing them losslessly, their method remains sound. Author [15] an overlapping square block has been used to divide the image. An image block has been processed using FMT to extract features. Instead of lexicographical sorting, they have utilized a counting bloom filter to identify similar feature vectors. To identify the tampered blocks in the image, they computed shift vectors. The technique detects lossy compressions, scalings, and rotations of manipulated images. Author [16] have proposed an improved approach following Fridrich's method. A grayscale image has been created from a colour input image. An overlapping square block has been used to divide the input image. As a next step, features from images will be extracted using DCT. He has performed truncation over feature vectors.

Furthermore, they have used lexicographical sorting to locate neighbouring feature vectors in the feature matrix containing feature vectors corresponding to the image blocks. They have computed the shift vector corresponding to the similar image blocks. They have imposed a user-defined threshold over the shift vector counter to identify the tampered area in forged images. Despite attacks [17, 18] such as JPEG compression, blurring, and white Gaussian noise, the authors' method has demonstrated robustness.

An RGB to grayscale image conversion was performed by Armas Vega et al. [19]. Rather than dividing the grayscale image into individual blocks, they overlap them. They have applied DCT over each block. They have sorted the coefficients of image blocks in a zig-zag manner. Further, they have performed the truncation over the coefficient list to extract coefficients of higher significance. Feature vectors corresponding to image blocks are lexicographically sorted over the feature matrix. They have computed the similitude measure between nearest neighbour feature vectors. The feature vectors of image blocks having similitude and translation vector values less than the user-defined threshold values are considered identical. Their method shows robustness against small rotation, scaling, and mixed rotation-resizing attacks applied over the altered images (Table 1).

3 Proposed Methodology

Keypoint-based detection of copy-move forgeries is proposed in our paper. It improves time and space complexity and reduces false positives [30].

Table 1 Localization and CMFD methods

Authors	Methods used	Dataset	Performance	Remarks
[20]	An averaged sum of absolute differences (ASAD) is calculated using block-based DCT coefficients	Official Marks Card Dataset (OMCD), CoMoFoD,	For copy-paste forgery, the precision was 80.45%, the recall was 89.05%, F-score was 84.53%; the precision was 95.95%, the recall was 94.52%, and F-score was 95.22%	Two subjects have not been tested for the same marks
[21]	CenSurE, FREAK, kNN, agglomerative hierarchical clustering based on keypoints	Grip, CMFD, MICC-F220, MICC-F600, CoMoFoD	The F_1 -measure for the CMFD was 97.61, for the GRIP was 95.12, for coverage was 97.50, for MICCF600 was 97.14, for the MICC-F220 was 98.43, and for the CoMoFoD was 98.43	Highly smooth images will not yield good results using this method. A combination attack such as rotation followed by scaling also results in poor results
[22]	SURF, RANSAC, SIFT, and AHC	MICC-F220	92.5% recall and 91.7 F_1 -score	There are only 220 images in the dataset. We do not test any other standard datasets besides these
[23]	RANSAC, CLAE, SIFT, GORE, DBSCAN,	The MICC-F220 dataset consists of image manipulations	95% recall	There is no testing of combination attacks. There are many challenges associated with DBSCAN for high-dimensional data
[24]	Dilation, erosion, GLCM, SPT	CASIA, CoMoFoD	With no attacks, the FPR and FNR are 1%; the TPR and TNR are 99%	Attacks result in lower TPRs and TNRs. There is no testing of combination attacks. Noise was not added to the method
[25]	COMA-DFSCAN, SIFT, ADALAM, KBSSCAN, and Padding	Coverage, CoMoFoD, MICCF220	For CASIA, the FI-score is 0.714; for MICC-F220, the FI-score is 0.904; for coverage, the FI-score is 0.711	Detecting forgeries could have taken a long time because of the many computations involved. There is no testing of combination attacks

(continued)

Table 1 (continued)

Authors	Methods used	Dataset	Performance	Remarks
[26]	RANSAC, FBO, SIFT, DWT	A sample of 500 websites	96% accuracy	It is not easy to obtain sufficient samples for training and testing. No tests are conducted on attacks
[27]	Clustering, Free Feature Engineering, Emperor-Labelled Feature Points, Gabor filter, Turbo Pixels, RANSAC,	MICC-F600	95.06% E-Measure	JPEG compression testing and double attacks are not mentioned. The dataset contains only 600 samples, which is a very small number
[28]	A hierarchy of segmentations based on feature labels, SIFT, and feature label matching	GRIP, CMH, and IMD datasets	With CMH, we achieved 91.15% precision, 91.185% recall, and 91.50% F1-score, while with GRIP, we achieved 91.09% precision, 93.15% recall, and 92.11% F1-score	There was no testing of double attacks. The CASIA and COVERAGE benchmarks are not tested
[25]	KANN-DBSCAN, SIFT, AdaLAM, Convex Hull, and Padding	CoMoFoD, CASIA, MICCF220, COVERAGE	The CASIA FI-score is 0.714, the MICCF220 FI-score is 0.904, and the coverage FI-score is 0.711	It would have been possible to measure the time needed to detect forgery since a lot of computation is involved. There is no testing of combination attacks
[28]	The FAST, BRIEF, SIFT, g2NN, and linear spectral clustering algorithms	MICC-F8 multi-datasets	There are 0.98864 F1-score, 0.9023 precision, and 1 recall	We did not test double attacks. The benchmark datasets of other benchmarks are not tested
[29]	SIFT, g2NN, RANSAC, Ciratefi, and LBP-Rot	CMH, GRIP	CMH dataset accuracy of 97.88 and GRIP F-measure of 96	Two matching processes are involved in the method, which is complicated. We have not conducted a complexity analysis. There was no testing of double attacks. The data is not tested against other benchmark datasets

3.1 SURF

Author [31] reports that the URF can withstand illumination changes and affine transformations, as well as rotation and scaling. A Hessian matrix determinant is used to determine the scale and location of SURF. Equation (1) shows the Hessian matrix $H(x; \sigma)$ at scale.

$$H(x; \sigma) = \begin{bmatrix} C_{xx}(x, \sigma) & C_{xy}(x, \sigma) \\ C_{xy}(x, \sigma) & C_{yy}(x, \sigma) \end{bmatrix}. \quad (1)$$

$C_{xx}(x, \sigma)$ and $C_{yy}(x, \sigma)$ to the fact that they are Gaussian second-order derivatives of the image in $\partial^2 g(\sigma)/\partial x^2$ question, they are treated as similar.

Box filters are used to reduce computation time when simulating Gaussian second-order derivatives. $D_{xx}(x, \sigma)$, $D_{xy}(x, \sigma)$, and $D_{yy}(x, \sigma)$, convolution can be accelerated, and the complexity of the process can be reduced by improving calculation speed. As shown in Eq. (2), Hessian's determinant is approximated by:

$$\det(H_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xx})^2. \quad (2)$$

Box filters and integral images are used to construct image pyramid scale-spaces. The interest points are determined based on the neighbourhood-based non-maximum suppression method. A Hessian response threshold T_{SURF} is applied to the estimated Hessian matrix to determine the candidate points and the image opinions.

The main direction must be determined before obtaining rotation invariant descriptors. An eight-radius 6σ circular region with a centred circular region response determines the direction of the point. The circular region is processed with a Haar wavelet filter to obtain Haar wavelet responses. A fan-shaped area scans the entire circular Haar wavelet response from an angle of $3\pi/3$. Using the Haar wavelet response vector sum, a fan-shaped region is calculated. Among all vectors, the longest determines the main direction.

3.2 Accelerated-KAZE (AKAZE)

The AKAZE algorithm is particularly useful for CMF detection. It reduces noise without disrupting the image data by using nonlinear scale-space blurring. Nonlinear diffusion models an image's luminance by increasing its scaling levels. Partially differential equations control diffusion through the divergence of flow functions. With a nonlinear scale-space partial differential equation, the luminance of an image can be diffused [32]. It is possible to see nonlinear diffusion in the following equation:

$$\frac{\partial LI}{\partial t} = \text{div}(C(x, y, t)\nabla LI). \quad (3)$$

There are four parameters in the conductivity function: t , $\nabla L I_\sigma$ and $C(x, y, t)$, which is:

$$C(x, y, t) = G(|\nabla L I_\sigma(x, y, t)|). \quad (4)$$

The $L I_\sigma$ and $\nabla L I_\sigma$ represent the round Gaussian version and gradient $L I_\sigma$, respectively, of an image. Perona and Malik introduced two conductivity functions, and one is considered in the proposed study. The conductivity function G_2 represents the larger area support greater than the smaller area support in AKAZE:

$$G_2 = \frac{1}{1 + \frac{|\nabla L I_\sigma|^2}{\Omega^2}}. \quad (5)$$

The contrast factor Ω information about edges controls ω diffusion level. Edge information is retained more when the value Ω is lower. As an empirical value, 70% of the gradient histogram $\nabla L I_\sigma$ is considered. KAZE generates scale images at every scale using fast explicit diffusion (FED). In contrast to common explicit techniques, FED uses fluctuating time steps and is extensively faster. Scharr filters improve rotation invariance by determinants of the Hessian matrix in the AKAZE detector. Since MLDB is an effective algorithm, AKAZE's descriptor relies on it. As a result, AKAZE has been adopted for varying scales because it is scale- and rotation-invariant [32].

3.3 Scale-Invariant Feature Transform (SIFT)

A 128-dimensional feature vector is descriptorized by a SIFT descriptor. As described below, there are four main steps to follow:

- (a) **Scale-space extreme detection:** Differentiation of Gaussian (DoG) for detecting extreme scale-spaces. Using the following equation, you will be able to determine the scale-space of the image as $L(W, V, \sigma)$, which is the convolution between $Gf(W, V, \sigma)$ and input image $Y(W, Y)$.

A Gaussian function $Gf(W, V, \sigma)$ is convolution with the input image $Y(W, Y)$, and the scale-space of the image is defined as follows: $L(W, V, \sigma)$:

$$L(W, V, \sigma) = Gf(W, V, \sigma) * Y(W, V), \quad (6)$$

with

$$Gf(W, V, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(W^2+V^2)/2\sigma^2}. \quad (7)$$

A Gaussian function $Gf(W, V, \sigma)$ has standard deviation $Y(W, V, \sigma)$, where W and V are the coordinates of the input image. As shown in the following Eq. (8), the difference of the Gaussian (DoG) technique is used to identify the key points in the image:

$$D(W, V, \sigma) = (Gf(W, V, k\sigma) - Gf(W, V, \sigma)) * Y(W, V) = L(W, V, \sigma) - L(W, V, k\sigma) - L(W, V, \sigma). \quad (8)$$

This image is convolution with Gaussian blur with scale $k\sigma$ ($Y(W, V)$) and $Gf(W, V, k\sigma)$ to produce $L(W, V, \sigma)$. Detecting local maxima and minima requires comparing $D(W, V, \sigma)$ each pixel's value with its neighbours' values on the same scale and their neighbours' values on the opposite scale. Keypoints are determined by comparing all pixels and selecting the one with the lowest or highest value.

- (b) **Keypoint localization:** Some keypoint candidates are unstable when scale-space extreme detection is applied. A pixel is interpolated based on its neighbors, which gives more precise results. The keypoints are more accurately located by interpolating a Taylor expansion of second order. As long as the candidate keypoint serves as the origin, $D(W, V, \sigma)$ can be expanded as follows:

$$D(w) = D + \frac{\partial D^T}{\partial w} w + \frac{1}{2} w^T \frac{\partial^2 D^T}{\partial w^2} w. \quad (9)$$

A candidate keypoint $w = (W, V, \sigma)$ is used to calculate D and its derivatives, whereas offset is the offset from $w = (W, V, \sigma)$. A derivative of this function is taken and set to zero, as shown below, to determine the extreme \hat{w} :

$$\hat{w} = -\frac{\partial^2 D^{-1}}{\partial w^2} \frac{\partial D}{\partial w}. \quad (10)$$

If the offset \hat{w} in any dimension is exceeding 0.5, the extreme lies closer to another candidate's keypoint. An interpolation is performed in case of a change in the candidate's keypoint. It is necessary to add the offset \hat{w} to the candidate keypoint's position to determine the interpolated position of the extreme. Keypoints with low contrast are removed from the offset \hat{w} when the second-order Taylor expansion $D(W)$ is evaluated, and if any extreme has a lower intensity than a threshold, the key points are excluded. Localizing and filtering keypoint candidates involve eradicating keypoints with low contrast in keypoint localization [32].

- (c) **Orientation assignment:** A gradient-based orientation assignment is performed in a third step. The local image properties assign an orientation to each key point in this step. In the first step, we calculate the Gaussian image L . A magnitude gradient m and an orientation θ are calculated in the following way:

$$m(W, V) = \sqrt{(L(W+1, V) - L(W-1, V))^2 + (L(W, V+1) - L(W, V-1))^2}, \quad (11)$$

$$\theta(W, V) = \tan^{-1} \left(\frac{L(W, V+1) - L(W, V-1)}{L(W+1, V) - L(W-1, V)} \right). \quad (12)$$

Gradient orientations of your sample points are mapped to an orientation histogram, and the orientation of each point is further mapped based on its peaks.

- (d) **Keypoint descriptor:** Keypoint descriptors are generated using the local gradient data above. There are 16 sub-blocks of 4×4 size surrounding a 4×4 key point. In the central point, 16 4×4 sub-blocks are laid out. This results in 128 bin values in total. Keypoint descriptors are formed with this vector. Additionally, several measures address rotation, noise addition, and scaling.

4 Results and Analysis

The CMFD results are presented in this section. Experimental results were compared to benchmark datasets based on state-of-the-art schemes. MATLAB R2020a, Windows 10 and Core i5-6200U CPU (2.30 GHz), 8 GB memory, Python, and OpenCV were all used in the experimental environment [33, 34]. This experiment tests our algorithm's effectiveness using the NB-CASIA dataset. In this experiment, no postprocessing was done on any forged images. Figure 3 illustrates a few examples of detected results. This method outputs colour-coded maps identifying duplicated regions and forgeries based on colour lines connecting the matching points. Colour lines provide an easy method for identifying the tampered region, allowing us to detect it objectively despite it not being able to be pinpointed down to the pixel level. The authentic image can be seen in Fig. 3. Detections of rotation, scaling, and horizontal reflection are illustrated in Fig. 3b–d, indicating that our method is effective in exposing copy-move forgeries resulting from geometric transformations. We could detect stable results more reliably than SIFT algorithms, especially when horizontal reflections are present.

The AKAZE, SURF, and SIFT are applied to the input image. A feature descriptor is then extracted. CMFD methods based on keypoints do not detect duplicate regions in most cases. Furthermore, most methods that obtain keypoints from smooth tampered regions cannot be executed with the default parameters. This image shows a camouflage, cat, duck, and coin with the same image that has been tampered with, as shown in Figs. 4 and 5, respectively. AKAZE, SIFT, SURF, and KAZE cannot extract features in the tampered region using their default parameters. However, AKAZE and SURF with small thresholds can achieve those points. Keypoint-based CMFD relies on obtaining sufficient numbers of points.

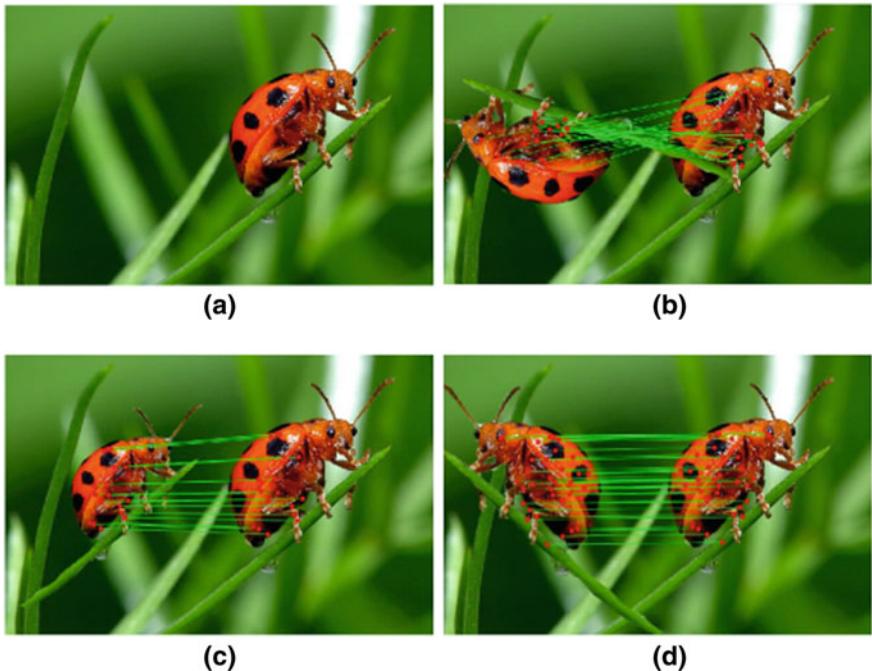


Fig. 3 Examples of detected results using the AKAZE method

Experimental comparisons were performed by performing the proposed method described in the feature extraction and its comparison with other keypoint-based CMFD approaches in Fig. 6. A comparison of the proposed method with existing CMFD methods revealed that it had a higher F -score [35], was more accurate, and recalled more data than other methods. Moreover, the proposed method is tested for handling multiple copies-moves. A comparison of the proposed copy-move forgery detection method with other relevant methods [36, 37] in terms of processing time is presented in Fig. 7.

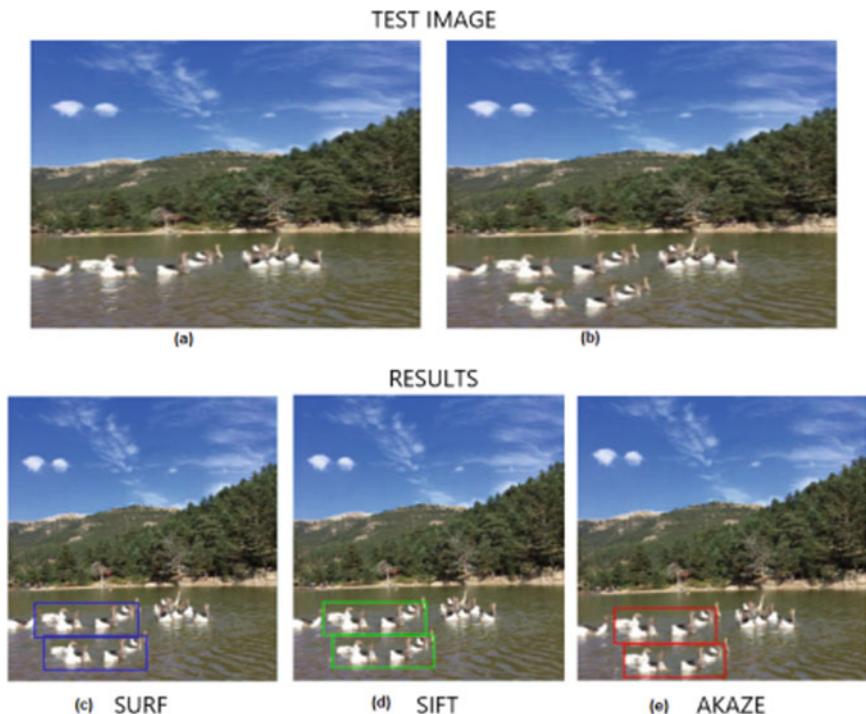


Fig. 4 Results of different keypoint detectors used to detect camouflage images: **a** unique image; **b** meddled image; **c** SURF detector; **d** SIFT detector; **e** AKAZE detector

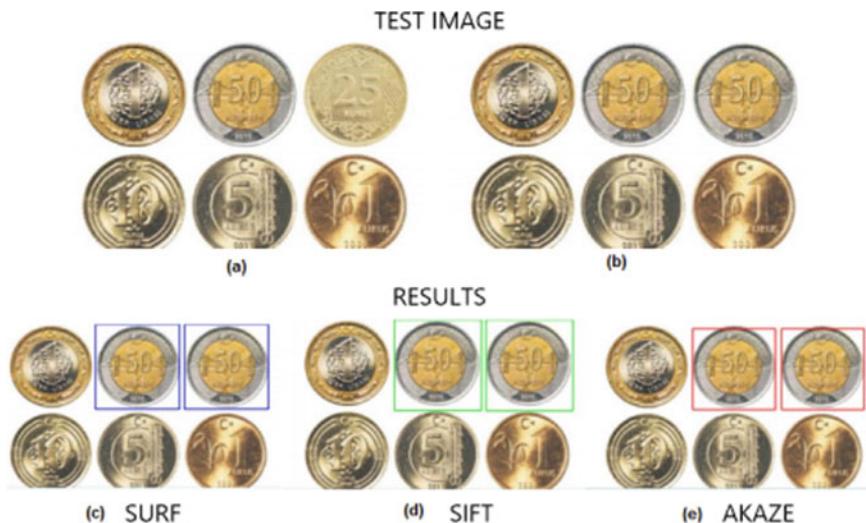


Fig. 5 Results of different keypoint detectors used to detect camouflage images: **a** unique image; **b** meddled image; **c** SURF detector; **d** SIFT detector; **e** AKAZE detector

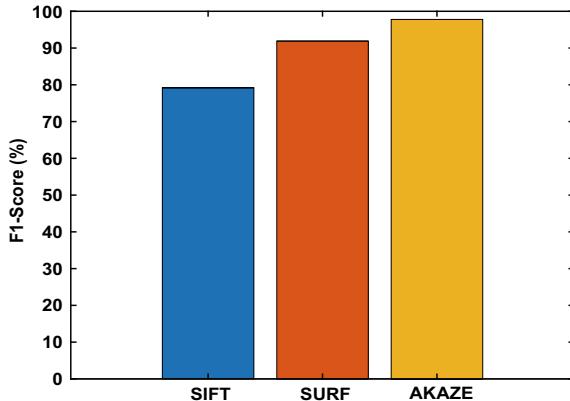


Fig. 6 Performance analysis of SIFT, SURF, and AKAZE in terms of F_1 -score (%)

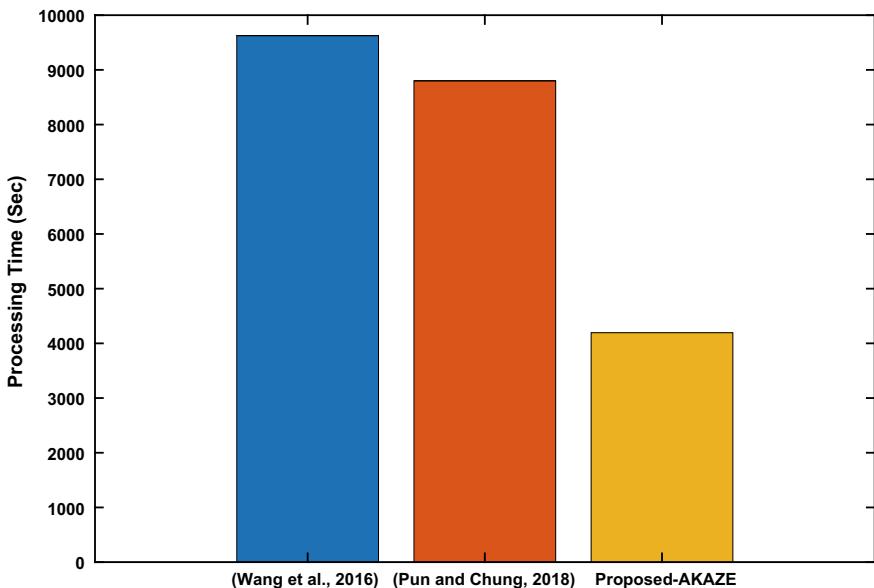


Fig. 7 Processing time (sec) of SIFT SURF and AKAZE

5 Conclusion

Copy-move forgeries hide the traces of tampering by copying and pasting parts of an image to hide tampering without appearing obvious. Forensics is concerned about the authenticity of images due to these types of tampering attacks. With powerful software development over the past few years, various techniques have been proposed to manipulate images. Many research papers have studied forgery detection techniques.

Insufficient keypoints also prevent keypoint-based methods from detecting smooth forged regions. In this work, we discuss a technique for detecting copy-move forgery. There is a proposed method for identifying cloned regions in tampered images no matter what geometric changes have been applied to them, for instance, rotations, scaling, noise, etc.

References

1. Fridrich J (2003) Detection of copy-move forgery in digital images. In: Proceedings of digital forensic research workshop, 2003
2. Huang Y, Lu W, Sun W, Long D (2011) Improved DCT-based detection of copy-move forgery in images. *Forensic Sci Int* 206(1–3):178–184
3. Popescu AC, Farid H (2004) Exposing digital forgeries by detecting duplicated image regions
4. Luo W, Huang J, Qiu G (2006) Robust detection of region-duplication forgery in digital image. In: 18th International conference on pattern recognition (ICPR'06), IEEE, pp 746–749
5. Mahdian B, Saic S (2009) Using noise inconsistencies for blind image forensics. *Image Vis Comput* 27(10):1497–1503
6. Bayram S, Avcibaş İ, Sankur B, Memon N (2005) Image manipulation detection with binary similarity measures. In: 2005 13th European signal processing conference. IEEE, pp 1–4
7. Li G, Wu Q, Tu D, Sun S (2007) A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. In: 2007 IEEE international conference on multimedia and expo. IEEE, pp 1750–1753
8. Bravo-Solorio S, Nandi AK (2011) Exposing duplicated regions affected by reflection, rotation and scaling. In: 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 1880–1883
9. Li L, Li S, Zhu H, Chu S-C, Roddick JF, Pan J-S (2013) An efficient scheme for detecting copy-move forged images by local binary patterns. *J Inf Hiding Multim Signal Process* 4(1):46–56
10. Li L, Li S, Zhu H, Wu X (2014) Detecting copy-move forgery under affine transforms for image forensics. *Comput Electr Eng* 40(6):1951–1962
11. Ryu S-J, Lee M-J, Lee H-K (2010) Detection of copy-rotate-move forgery using Zernike moments. In: International workshop on information hiding. Springer, pp 51–65
12. Ryu S-J, Kirchner M, Lee M-J, Lee H-K (2013) Rotation invariant localization of duplicated image regions based on Zernike moments. *IEEE Trans Inf Forensics Secur* 8(8):1355–1370
13. Cozzolino D, Poggi G, Verdoliva L (2015) Efficient dense-field copy–move forgery detection. *IEEE Trans Inf Forensics Secur* 10(11):2284–2297
14. Qazi T et al (2013) Survey on blind image forgery detection. *IET Image Process* 7(7):660–670
15. Bayram S, Sencar HT, Memon N (2009) An efficient and robust method for detecting copy-move forgery. In: 2009 IEEE international conference on acoustics, speech and signal processing. IEEE, pp 1053–1056
16. Bayram S, Sencar HT, Memon N (2009) An efficient and robust method for detecting copy-move forgery. In: 2009 IEEE international conference on acoustics, speech and signal processing. IEEE, pp 1053–1056
17. Hussain N, Rani P, Chouhan H, Gaur US (2022) Cyber security and privacy of connected and automated vehicles (CAVs)-based federated learning: challenges, opportunities, and open issues. In: Yadav SP, Bhati BS, Mahato DP, Kumar S (eds) Federated learning for IoT applications. EAI/Springer Innovations in Communication and Computing. Springer International Publishing, Cham, pp 169–183. https://doi.org/10.1007/978-3-030-85559-8_11
18. Hussain N, Rani P (2020) Comparative studied based on attack resilient and efficient protocol with intrusion detection system based on deep neural network for vehicular system security. In: Distributed artificial intelligence. CRC Press, pp 217–236. Accessed: Oct 19 2023.

- (Online). Available: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003038467-13/comparative-studied-based-attack-resilient-efficient-protocol-intrusion-detection-system-based-deep-neural-network-vehicular-system-security-naziya-hussain-preeti-rani>
- 19. Armas Vega EA, González Fernández E, Sandoval Orozco AL, García Villalba LJ (2021) Copy-move forgery detection technique based on discrete cosine transform blocks features. *Neural Comput Appl* 33(10):4713–4727. <https://doi.org/10.1007/s00521-020-05433-1>
 - 20. Darem A, Al-Hashmi A, Javed M, Abubaker A (2020) Digital forgery detection of official document images in compressed domain. <https://doi.org/10.22937/IJCSNS.2020.20.12.12>
 - 21. Diwan A, Sharma R, Roy AK, Mitra SK (2021) Keypoint based comprehensive copy-move forgery detection. *IET Image Process* 15(6):1298–1309. <https://doi.org/10.1049/ipr2.12105>
 - 22. Sunitha K, Krishna AN (2020) Efficient keypoint based copy move forgery detection method using hybrid feature extraction. In: 2020 2nd International conference on innovative mechanisms for industry applications (ICIMIA), pp 670–675. <https://doi.org/10.1109/ICIMIA48430.2020.9074951>
 - 23. Hegazi A, Taha A, Selim MM (2021) An improved copy-move forgery detection based on density-based clustering and guaranteed outlier removal. *J King Saud Univ Comput Inf Sci* 33(9):1055–1063. <https://doi.org/10.1016/j.jksuci.2019.07.007>
 - 24. Babu ST, Rao CS (2022) An optimized technique for copy–move forgery localization using statistical features. *ICT Express* 8(2):244–249
 - 25. Yue G, Duan Q, Liu R, Peng W, Liao Y, Liu J (2022) SMDAF: a novel keypoint based method for copy-move forgery detection. *IET Image Process* 16(13):3589–3602
 - 26. Uma S, Sathy PD (2022) Copy-move forgery detection of digital images using football game optimization. *Aust J Forensic Sci* 54(2):258–279
 - 27. Gan Y, Zhong J, Vong C (2022) A novel copy-move forgery detection algorithm via feature label matching and hierarchical segmentation filtering. *Inf Process Manag* 59(1):102783
 - 28. Fatima B, Ghafoor A, Ali SS, Riaz MM (2022) FAST, BRIEF and SIFT based image copy-move forgery detection technique. *Multimed Tools Appl* 81(30):43805–43819
 - 29. Tahaoglu G, Ulutas G, Ustubioglu B, Ulutas M, Nabiye VV (2022) Ciratefi based copy move forgery detection on digital images. *Multimed Tools Appl* 81(16):22867–22902
 - 30. Yadav SP, Jindal M, Rani P, De Albuquerque VHC, Dos Santos Nascimento C, Kumar M (2023) An improved deep learning-based optimal object detection system from images. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-023-16736-5>
 - 31. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: European conference on computer vision. Springer, pp 404–417
 - 32. Prakash CS, Panzade PP, Om H, Maheshkar S (2019) Detection of copy-move forgery using AKAZE and SIFT keypoint extraction. *Multimed Tools Appl* 78:23535–23558
 - 33. Rani P, Singh PN, Verma S, Ali N, Shukla PK, Alhassan M (2022) An implementation of modified blowfish technique with honey bee behavior optimization for load balancing in cloud system environment. *Wirel Commun Mob Comput* 2022:1–14
 - 34. Rani P, Verma S, Yadav SP, Rai BK, Naruka MS, Kumar D (2022) Simulation of the lightweight blockchain technique based on privacy and security for healthcare data for the cloud system. *Int J E-Health Med Commun IJEHMC* 13(4):1–15
 - 35. Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. *Comput Electr Eng* 105:108543
 - 36. Pun C-M, Chung J-L (2018) A two-stage localization for copy-move forgery detection. *Inf Sci* 463–464:33–55. <https://doi.org/10.1016/j.ins.2018.06.040>
 - 37. Wang X-Y, Li S, Liu Y-N, Niu Y, Yang H-Y, Zhou Z (2017) A new keypoint-based copy-move forgery detection for small smooth regions. *Multimed Tools Appl* 76(22). Art. no. 22

Correction to: Machine Learning Approach to Lung Cancer Survivability Analysis



Srichandana Abbineni , K. Eswara Rao , Rella Usha Rani , P. Ila Chandana Kumari, and S. Swarajya Lakshmi

Correction to:

Chapter “Machine Learning Approach to Lung Cancer Survivability Analysis” in: D. K. Sharma et al. (eds.), *Micro-Electronics and Telecommunication Engineering, Lecture Notes in Networks and Systems 894*, https://doi.org/10.1007/978-981-99-9562-2_33

In the original version of chapter 33, the following belated correction has been incorporated: The affiliations for the authors K. Eswara Rao and Rella Usha Rani have been changed as follows. From: K. Eswara Rao, Department of CSE (AI&ML), CVR College of Engineering, Hyderabad, India. Rella Usha Rani, Department of CSE, Aditya Institute of Technology and Management, Tekkali, India. To: K. Eswara Rao, Department of CSE, Aditya Institute of Technology and Management, Tekkali, India. Rella Usha Rani, Department of CSE (AI&ML), CVR College of Engineering, Hyderabad, India.

The updated version of this chapter can be found at
https://doi.org/10.1007/978-981-99-9562-2_33

Author Index

A

- Abdullah J. Alzahrani, 769
Abdulzahra, Lina Shugaa, 331, 585
Abha Kiran Rajpoot, 103
Abhinav Vidwans, 571
Abhishek Gupta, 673
Abhishek Kumar, 785
Abinav, K., 125
Adnan Shakeel Ahmed, 113
Akash Kumar, 67
Akila, D., 457, 481, 497, 509
Al-Hamami, Ahmed Alaa, 347
Al-Hamami, Mohammed Alaa, 347
Alkhayyat, Ahmed Hussein, 357, 389, 631, 755
AL Raheim Hamza, Lena Abed, 87
Al Rashid, Sura Zaki, 87
Aman Agarwal, 197
Amber Khan, 113
Amina Atiya Dawood, 43
Amitav Saran, 263
Amit Prakash Singh, 549
Anandha Lakshmi, R., 509
Ananta Sekia, 75
Ankit Kumar, 715
Ankit Shukla, 57
Anmol Tyagi, 197
Annapurani, Panaiyappan K., 161
Anshika Singh, 737
Anubhav Thakur, 57
Anuj Kumar Singh, 593
Anum Kiyani, 447
Anup Rana, 509
Anuradha Chug, 549
Archana, M. A., 357, 389, 631

Ashutosh Pandey, 653

- Ayan Sar, 685
Ayushi Gupta, 549
Ayushi Jain, 297

B

- Balwinder Raj, 29
Banchhanidhi Dash, 365
Bikramjit Sarkar, 497, 509
Bipasha Shukla, 309
Boovarahan, N. C. A., 357, 389, 631

C

- Chirag Agarwal, 197

D

- Danish Raza Rizvi, 113
Deepak Garg, 747
Deepti Kakkar, 29
Devendra Kumar Sharma, 673, 747
Dev Tyagi, 197
Dhaarna Singh Rathore, 437
Dinesh Kumar, T., 357, 389, 631
Dinesh Prasad, 113
Divya Sharma, 653

E

- Eswara Rao, K., 397

G

- Gaurav Kumar, 715

Gayathri Karthick, 447

George, Loay E., 183

Gopikrishnan, S., 137

Gunjan Chhabra, 437

H

Hasan, Zainab Ghayyib Abdul, 275

Hassib, Mustafa Dh., 377

Hussein, Raghad I., 469

I

Ila Chandana Kumari, P., 397

Ishaan Agarwal, 755

J

Jaideep Singh, 497

Jay Prakash, 617

Jeyalakshmi, S., 481

Jnyana Ranjan Mohanty, 409

Jyoti Maini, 1

Jyoti Srivastava, 539, 617

Jyoti Upadhyay, 249

K

Kabir Choudhary, 67

Kaleemur Rehman, 797

Kalpana Singh, 309

Kamini Lamba, 149, 289

Kamred Udhamp Singh, 715

Kanneboina Ashok, 137

Kaushal Kishor, 57, 67

Kavita Chaudhary, 309

Kavitha, K., 457

Ketan Kotecha, 523, 685, 715

Khadeer Dudekula, 161

Kothai, G., 13

Kousik Midya, 653

L

Lafta, Hussein Attia, 87

Lakshay Singla, 755

Lakshmi, S., 357, 631

Ling, Soonleh, 447

M

Madhumathi, K., 497

Mainuddin, 75

Malathi, V., 641

Mandeep Singh, 29, 249

Manisha Vashisht, 319

Manoj Ramaiya, 571

Martin Sagayam, K., 561

Mary Havilah Haque, B., 561

Masoody Al, Wasan Hashim, 377

Mays Ali Shaker, 43

Maysara Mazin Alsaad, 235

Mirtha Silvana Garat Marin de, 685

Mirza Tariq Beg, 75

Mohammed, Abulkareem Z., 183

Mohd Ashraf, 75

Mohini Preetam Singh, 197, 219

Mudhafer Al, Rusul A., 601

Muthukumaran, D., 631

Muzammil, S., 13

Mythili, R., 125

N

Naga Raju, B., 13

Nakkina Sai Teja, 29

Nalinda Somasiri, 447

Narayananamoorthi, R., 419

Neha Mittal, 297

Neha Singh, 539

Nidhal K. El Abbadi, 601

Nitin Jain, 661

O

Obaid, Ahmed J., 331, 585

Omkumar, S., 357, 389

Onaizah, Ameer N., 469

P

Partha Sarathi Ghosh, 263

Parthasarathy, C., 389

Poornima, G., 631

Porkodi, S. P., 173

Prabakaran, S., 389

Prabhas Bhanu Boora, 13

Prasant Kumar Patnaik, 365, 409

Priteshe Pathak, 481

Priti Verma, 437

Priya Kumari, 661

Priyank Sharma, 785

R

Rahul Singh, 197, 219

Raj Kumar Parida, 509

Ramesh Kumar Sunkaria, 249

Rebecca Jeyavadhanams, B., 447

Rella Usha Rani, 397

Richa Choudhary, 699

Richa Pandey, 437

Srichandana Abbineni, 397

Surtipragyan Swain, 365

Suresh Krishna, S., 125

Susheela Dahiya, 699

Swarajya Lakshmi, S., 397

S

Sahai, D. N., 673

Saiful Islam, 797

Saikat Maity, 481

Samiappan Dhanalakshmi, 419

Sandeep Kumar, 593

Sanjai K. Dwivedi, 75

Sanjive Tyagi, 235

Sankalp Goel, 103

Santosh Kumar, 437

Sanyam Agarwal, 755

Sarada, V., 173

Saraswathi, K., 357

Saurabh Adhikari, 481, 497, 509

Saurav Adhikari, 457

Sayantan Panda, 419

Seema Garg, 653

Seema Gupta Bhol, 409

Selvi, H., 497

Senthil, D., 457

Shagun Chandrvanshi, 219

Shalli Rani, 1, 149, 289

Sharvan Kumar Garg, 235, 727, 737

Shivam Sharma, 219

Shubham Nain, 297

Shyam Akashe, 785

Snigdha Parashar, 523

Sourav Kumar Singh, 125

Souvik Pal, 481, 497, 509

T

Tanupriya Choudhury, 523, 685, 699, 715

Tarun Chaudhary, 29, 249

Teekam Singh, 715

Thiyagarajan Chenga Kalvinathan, 263

Thyagaraj, M., 457

Tridha Bajaj, 523, 685

U

Umapathy, K., 357, 389, 631, 641

Umasankar Das, 263

Umesh Chandra Jaiswal, 539

V

Vaishali Bhargava, 727

Vanitha, D., 641

Veer Daksh Agarwal, 755

Venkata Subhash, L., 13

Vineet Kumar Singh, 593

Vipin Mittal, 755

Vipul Vashisht, 319

Vishwas Mishra, 785

Z

Zaiter, Mohammed Joudah, 377