

Lecture Notes in Networks and Systems 563

S. Smys
Khaled A. Kamel
Ram Palanisamy *Editors*

Inventive Computation and Information Technologies

Proceedings of ICICIT 2022

 Springer

Lecture Notes in Networks and Systems

Volume 563

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

Okay Kaynak, Department of Electrical and Electronic Engineering,
Bogazici University, Istanbul, Turkey

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).

S. Smys · Khaled A. Kamel · Ram Palanisamy
Editors

Inventive Computation and Information Technologies

Proceedings of ICICIT 2022

 Springer

Editors

S. Smys
Department of Computer Science
Engineering
RVS Technical Campus
Coimbatore, Tamil Nadu, India

Khaled A. Kamel
Department of Computer Science
College of Science and Technology
Texas Southern University
Houston, TX, USA

Ram Palanisamy
Department of Business Administration
The Gerald Schwartz School of Business
StFX University
Antigonish, NS, Canada

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-19-7401-4

ISBN 978-981-19-7402-1 (eBook)

<https://doi.org/10.1007/978-981-19-7402-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

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

ICICIT 2022 is highly privileged to dedicate this book to the young and aspiring researchers in the “Computation and Informatics” domain. We also dedicate this book to all of the reviewers, authors, and committee members who contributed actively during the whole conference programme.

Preface

This Conference Proceedings volume contains the written versions of most of the contributions presented during the conference of ICICIT 2022. The conference provided a setting for discussing recent developments in a wide variety of topics including Cloud Computing, Artificial Intelligence and Fuzzy Neural Systems. The conference has been a good opportunity for participants coming from various destinations to present and discuss topics in their respective research areas.

This Conference tends to collect the latest research results and applications on Computation Technology, Information and Control Engineering. It includes a selection of 68 papers from 312 papers submitted to the conference from universities and industries all over the world. All of accepted papers were subjected to strict peer-reviewing by 2–4 expert referees. The papers have been selected for this volume because of quality and the relevance to the conference.

We would like to express our sincere appreciation to all authors for their contributions to this book. We would like to extend our thanks to all the referees for their constructive comments on all papers, especially, we would like to thank to

organizing committee for their hard working. Finally, we would like to thank the Springer publications for producing this volume.

Dr. S. Smys
Professor
RVS Technical Campus
Coimbatore, India

Dr. Khaled A. Kamel
Professor
Department of Computer Science
Texas Southern University
Houston, USA

Dr. Ram Palanisamy
Professor
Department of Business Administration
The Gerald Schwartz School of Business
StFX University
Antigonish, Canada

Contents

Detecting Wildlife Trapped Images Using Automatically Shared Nearest Neighbouring Pixels (ASNNP)	1
S Anantha Babu, V. Manikandan, M. Jaiganesh, M. John Basha, and P. Divya	
A Meta-Classifer Link Prediction Model for False Profile Identification in Facebook	15
S. Saranya, M. Rajalakshmi, S. Devi, and R. M. Suruthi	
Comparison of Various CNN Models for Image Classification	31
S. Sony Priya and R. I. Minu	
Enhancing Customer Prediction Using Machine Learning with Feature Selection Approaches	45
R. Siva Subramanian, B. Maheswari, S. Nikkath Bushra, G. Nirmala, and M. Anita	
Domain Name System Resolution System with Hyperledger Fabric Blockchain	59
Khoi Tuan Huynh Nguyen, Hy Khang Vuong Nguyen, Dung Ngoc Vo Lam, Tuan Anh Dao, Nhan Trong Van Pham, Huong Hoang Luong, and Hai Thanh Nguyen	
Parametric Analysis of Resource-Constrained Devices for Internet of Things Using FreeRTOS	73
Partha Pratim Ray	
A Text Analytics Approach of Exploratory Visualization of Legal Parameters of Dowry Death Cases	85
Souraneel Mandal, Sajib Saha, and Tanaya Das	
Okra Cultivation Using IoT and Disease Detection	97
D. Abishek, S. Hareesh Chandran, S. Mayukha, and B. Sabitha	

Identifying Pitfalls and Solutions in Parallelizing Long Short-Term Memory Network on Graphical Processing Unit by Comparing with Tensor Processing Unit Parallelism 111
Aswathy Ravikumar, Harini Sriraman, S. Lokesh, and P. Maruthi Sai Saketh

Supply Chain Management in Blockchain Using Hyperledger 127
Arun Kumar Kushwaha, Soumya Singh, Anshay Rastogi, Vipin Deval, Pravesh Singh, and Vikas Goel

Jyoti: An Intelligent App for Women Security 139
Pooja Pathak and Parul Choudhary

Performance Analysis of Machine Learning Algorithms in the Systematic Prediction of Chronic Kidney Disease on an Imbalanced Dataset 149
M. Revathi, G. Raghuraman, and J. Visumathi

Robotic Process Automation Powered Admission Management System 169
Abhishek Gupta, Pranav Soneji, and Nikhita Mangaonkar

Implementation of Advanced High Performance Bus to Advanced Peripheral Bus Bridge 179
Chinta Sai Manasa, Navya Mohan, and J. P. Anita

The Influence of Digital Financial Literacy on Interest in Investing in the Capital Market of Millennial Generation Students 189
Ford Lumban Gaol, Christopher, Gerry Dermawan, Muhammad Resnandya Rafif, Rully Rubianto, and Tokuro Matsuo

Security Enhancement in IoT with an Application of Water Monitoring 207
Jui Sanjay Berde and Shirish S. Kulkarni

Decision Support System for Weeding Out Drug Seeking Behavior from Emergency Clinics 223
Rugved V. Deolekar, Sunil Wankhade, and Mayur Wanve

Analysis of Machine Learning Algorithm to Predict Symptoms of Diabetes—A Review 233
M. Sivaraman and J. Sumitha

Sound Mind: Detecting Depression and Anxiety in Humans 243
S. Revathi, K. Keerthanaa, A. R. Ranjitha, and Priyanka Balli

A Survey of Different Identification and Classification Methods for Medicinal Plants 273
Shashank M. Kadiwal, Venkatesh Hegde, N. V. Shrivathsa, S. Gowrishankar, A. H. Srinivasa, and A. Veena

SARS-CoV-2 Mortality Risk Prediction Using Machine Learning Algorithms to Aid Medical Decision-Making 293
 Manish Sewariya and Sanjay Patidar

A Comprehensive Study on Analysis and Prediction of Pollution Data Using Machine Learning 307
 R. D. Aishwarya, C. Sahana, V. J. Deepa, J. Durgashree, S. Gowrishankar, and A. Veena

Utilization of Waste Banks System in the COVID-19 Pandemic 319
 Riyan Pradanang, Bima Aditya, Adhitya Rayhan, Michael Adrian Winarlie, Ford Lumban Gaol, and Tokuro Matsuo

None Shall Pass: A Blockchain-Based Federated Identity Management System 329
 Shlok Gilda, Tanvi Jain, and Aashish Dhalla

Efficient Big Data Clustering Using Adhoc Fuzzy C Means and Auto-Encoder CNN 353
 Shreehari Harthi Sreedhara, Vineet Kumar, and Subia Salma

Covid Analysis Using Machine Learning Technique 369
 Rohit Gussain, Nitin Kumar, Minshul Sharma, and Pooja Dehraj

Skin Lesion Classification Using Machine Learning 385
 J. Vindhya, C. Pooja, Manisha H. Dongre, S. Gowrishankar, A. H. Srinivasa, and A. Veena

IOT-Based Weather Monitoring and Weather Prediction Using ML 399
 Brijesh Vadalía, Nisarg Thakkar, Shalin Tilva, Pratham Thakkar, and Ritiksha Modi

EverTrack: A System to Track History of Used Cars Using Blockchain 415
 Rex Rony Jacob, Rohit Mahesh, Shannon Gomez, Syed Sahabuddin, and Vijitha Robinson

Identification of Respiratory Diseases and Diabetes by Non-invasive Method Using IoT 425
 S. Suthagar, G. Mageshkumar, K. Hemalatha, Saggurthi Prabhakara Rao, R. Mahesh, and S. M. Kural Eniyavan

Use of E-Wallet as a Substitute for Physical Money in Transactions at Malls 441
 Achmad Rizky Kurniawan, Rafif Imam, Rifqi Farhan Zhalifunnas, Armando Rama Putra, Ford Lumban Gaol, and Tokuro Matsuo

Predictive Study and Classification of Diabetes Using Machine Learning Techniques 451
 Krishan Kumar and Sanjay Patidar

Review on Modern Techniques Behind Virtual Cloth Try-On 461
Prajakta Joglekar and Vinaya Gohokar

Identification and Representation of Spectral Anomalies in an Abandoned Quarry by Remote Sensing 479
C. Gambardella and R. Parente

A Naïve Bayes Approach for Predicting the Skin Allergy Diseases 495
B. Geluvaraj, K. Santhosh, T. Sandhya, V. Akshay Reddy, S. V. Bhaskar, and N. Sasikala

Internet of Things with Deep Learning Driven Disease Detection and Classification Model 507
C. Nithyeswari and G. Karthikeyan

OCR-Based Number Plate Identification Using MATLAB 521
P. Pooja, G. Maha Lakshmi, R. S. V. S. Vinith, and Siva Sankar Yellampalli

Machine Learning-Based Classification Between Block Cipher and Stream Cipher 531
Shivank Kumar Dadhwal and Girish Mishra

Comparative Study of Various Machine Learning Algorithms with MDLDPTS for Plant Leaf Disease Analysis 543
N. Venkatakrishnan and M. Natarajan

Arithmetic Optimization Algorithm with Deep Learning-Based Medical X-Ray Image Classification Model 563
T. Kumar and R. Ponnusamy

Lightweight Encoder-Decoder Model for Semantic Segmentation of Hand Postures 579
Siddhant Kumar Das, Rishabh Lahkar, K. Antariksha, Ayanmani Das, Aryamaan Bora, and Amrita Ganguly

Flexi-Pay System for Public Transport Based on Passenger Comfort Using Global Positioning System 593
G. Mageshkumar, S. Suthagar, D. Maher Shalal, S. Muruganantham, and T. Manoji

An Improvised Algorithm for a Dynamic Key Generation Model 607
D. V. Guru Saran and Kurunandan Jain

A New Authentication Protocol for Hardware-Based Authentication Systems in an IoT Environment 629
Rohit Pilla and Kurunandan Jain

Hybrid Cell Selection Mechanism for V2X Handover 641
Faiza Al Harthi and Abderezak Touzene

Utilization of Augmented Reality Technology in the Campus Environment 657
Ford Lumban Gaol, Mufti Ikhsan Kamil,
Aria Muhammad Iswardhana, Steven Gusda Firnandes, Fazri Fahrezi,
Tokuro Matsuo, and Fonny Hutagalung

FPGA-Based Implementation of Time-To-Digital Converter 669
H. Mangalam, S. Meishree, S. Mageshwari, and B. Prathiksha

Implementation of Smart Poultry Farm Using Internet of Things 679
Bharati Ramteke and Snehlata Dongre

Design and Implementation of a Solar Powered Floating Device for Water Quality Monitoring in Inland Aquaculture Farms Using LoRa Wireless Communication 687
G. Mageshkumar, Saggurthi Prabhakara Rao, S. Suthagar,
K. Hemalatha, K. Gowtham, and T. Hariharan

Afriendly: M-Health App as Strategy to Avoid Bullying 707
Genesis Dayana Pinto Almeida, María Cristina Páez Quinde,
Carlos Alberto Ramos Guaña, and Carlos Fabián Martínez Vásquez

Machine Learning Techniques for Automated Nuclear Atypia Detection in Histopathology Images: A Review 717
Jithy Varghese and J. S. Saleema

Security Tradeoff in Network Virtualization and Their Countermeasures 741
Nayeem Ahmad Khan and Mohammed Al Qurashi

Current Waste Management Issues in Developing Countries and IoT-Based Solutions in Prospect of Smart City Development 751
Mohd Anjum, Sana Shahab, Ummatul Fatima, and M. Sarosh Umar

Image Forgery Detection: A Review 769
Preethi Vijayakumar, Elizabeth Mathew, M. Gayathry Devi,
P. T. Monisha, C. Anjali, and Jisha John

Two Stage Deep Learning Channel Estimation Scheme for Massive MIMO Systems 779
B. Pradheep T. Rajan and N. Muthukumar

Gamification as a Tool in Psychomotor Development: Case of Study Elementary School 789
Carlos Fabián Martínez Vásquez, Genesis Dayana Pinto Almeida,
Evelyn Paulina Rovalino Ortega, and José David Sillagana Torres

The Implementation of Octave-S in the Risk Management Measurement of Information System in the Department of Communication, Informatics, and Statistics 801
 Oggsa Sukardi, Ford Lumban Gaol, and Tokuro Matsuo

Data Protection Using Scrambling Technique 811
 Pothuka Madhuri and E. Prabhu

Impact of Industrial Revolution 4.0 on Vietnamese Ethnic Minority Human Resources 823
 Hien Thu Thi Ta, Trung Tran, and Phuong Thuy Thi Nguyen

Fine-Tuning MobileNet for Breast Cancer Diagnosis 841
 Huong Hoang Luong, Nghia Trong Le Phan, Toai Cong Dinh, Thuan Minh Dang, Tin Tri Duong, Tong Duc Nguyen, and Hai Thanh Nguyen

Identification of Leaf Disease Using Machine Learning Algorithm—CNN 857
 P. V. Raja Suganya, A. R. Sathyabama, K. Abirami, and C. M. Nalayini

BIO-VR: Design and Implementation of Virtual Reality-Based Simulated Biology Laboratory Using Google Cardboard with an Emphasis on Virtual Education 867
 Towfik Ahmed, Afzal Un Nayeem Chowdhury, and Ziaul Hasan Mozumder

Multi-label Classification of Cervix Types with Image Size Optimization for Cervical Cancer Prescreening by Deep Learning 885
 Maryna Tomko, Mykhaylo Pavliuchenko, Ivan Pavliuchenko, Yuri Gordienko, and Sergii Stirenko

Faecal Image-Based Chicken Disease Classification Using Deep Learning Techniques 903
 S. Suthagar, G. Mageshkumar, M. Ayyadurai, C. Snegha, M. Sureka, and S. Velmurugan

Fake News Detection Using Supervised Machine Learning Classification Algorithms 919
 Akanksha Singh and Sanjay Patidar

The Effect of E-Commerce on Increasing Consumption Behavior During a Pandemic 935
 Ford Lumban Gaol, Sherina Gloria Tompunu, Jessica Adine Larissa, Arista Wijaksana, Ferdianto Prajnowira, and Tokuro Matsuo

Network Device Identity Management Using Cryptography 951
 Rohith Kapudasu and Kurunandan Jain

Change in Outlook of Indian Industrial OEMs Towards IIoT Adoption During COVID-19 965
Prasanna V. Rao Deshpande and S. Kumar Chandar

Author Index 973

Editors and Contributors

About the Editors

Dr. S. Smys received his M.E. and Ph.D. degrees all in wireless communication and networking from Anna University and Karunya University, India. His main area of research activity is localization and routing architecture in wireless networks. He serves as an associate editor of *Computers and Electrical Engineering (C&EE) Journal*, Elsevier, and a guest editor of *MONET Journal*, Springer. He is served as a reviewer for IET, Springer, Inderscience and Elsevier journals. He has published many research articles in refereed journals and IEEE conferences. He has been the general chair, session chair, TPC chair and panelist in several conferences. He is the member of IEEE and the senior member of IACSIT wireless research group. He has been serving as an organizing chair and a program chair of several international conferences and in the program committees of several international conferences. Currently, he is working as a professor in the Department of CSE at RVS Technical Campus, Coimbatore, India.

Dr. Khaled A. Kamel is currently a chairman and professor at Texas Southern University, College of Science and Technology, Department of Computer Science, Houston, TX. He has published many research articles in refereed journals and IEEE conferences. He has more than 30 years of teaching and research experience. He has been the general chair, session chair, TPC chair and panelist in several conferences and acted as a reviewer and guest editor in referred journals. His research interest includes networks, computing and communication systems.

Prof. Ram Palanisamy is a professor of Enterprise Systems in the Business Administration Department at the Gerald Schwartz School of Business, St. Francis Xavier University. Dr. Palanisamy teaches courses on Foundations of Business Information Technology, Enterprise Systems using SAP, Systems Analysis and Design, SAP Implementation, Database Management Systems and Electronic Business (Mobile Commerce). Before joining StFX, he taught courses in Management at the Wayne

State University (Detroit, USA), Universiti Telekom (Malaysia) and National Institute of Technology (NITT), Deemed University, India. His research interest includes enterprise systems (ES) implementation; ES acquisition; ES flexibility, ES success; knowledge management systems; and healthcare inter-professional collaboration.

Contributors

K. Abirami Department of Information Technology, Velammal Engineering College, Chennai, India

D. Abishek Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India

Bima Aditya School of Information System, Bina Nusantara University, Alam Sutera, Tangerang, Indonesia

Towfik Ahmed Department of Computer Science and Engineering, Leading University, Sylhet, Bangladesh

R. D. Aishwarya Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

V. Akshay Reddy Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

Faiza Al Harthi Department of Computer Science, Sultan Qaboos University, Muscat, Sultanate of Oman

Mohammed Al Qurashi Faculty of Computer Science and Information Technology, AlBaha University, AlBaha, Saudi Arabia

Genesis Dayana Pinto Almeida High School “Nuevo Mundo”, Ambato, Ecuador

S Anantha Babu Department of Computer Science and Engineering, Jain Deemed to-be University, Bangalore, India

J. P. Anita Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

M. Anita Department of CSE, S. A. Engineering College, Chennai, Tamil Nadu, India

C. Anjali Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

Mohd Anjum Department of Computer Engineering, Aligarh Muslim University, Aligarh, India

K. Antariksha Assam Engineering College, Guwahati, Assam, India

M. Ayyadurai SRM Institute of Science and Technology, Chennai, India

Priyanka Balli NHCE, Bengaluru, India

Jui Sanjay Berde Department of Instrumentation Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, India

S. V. Bhaskar Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

Aryamaan Bora Assam Engineering College, Guwahati, Assam, India

Parul Choudhary Department of Computer Application and Engineering, GLA University, Uttar Pradesh, Mathura, India

Afzal Un Nayeem Chowdhury Department of Computer Science and Engineering, Metropolitan University, Sylhet, Bangladesh

Christopher School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

Shivank Kumar Dadhwal Maharaja Surajmal Institute of Technology, Delhi, India

Thuan Minh Dang FPT University, Can Tho, Vietnam

Tuan Anh Dao FPT University, Can Tho, Viet Nam

Ayanmani Das Assam Engineering College, Guwahati, Assam, India

Siddhant Kumar Das Assam Engineering College, Guwahati, Assam, India

Tanaya Das Adamas University, Kolkata, India

V. J. Deepa Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Pooja Dehraj Department of Computer Science and Engineering, Noida Institute of Engineering and Technology, Noida, India

Rugved V. Deolekar Vidyankar Institute of Technology, Mumbai, India

Gerry Dermawan School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

Vipin Deval KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

S. Devi Faculty of Information Technology, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

Aashish Dhalla Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, USA

Toai Cong Dinh FPT University, Can Tho, Vietnam

P. Divya Department of Computer Science, Loyola College of Arts and Science, Mettala Tamilnadu, India

Manisha H. Dongre Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Snehlata Dongre Department of Computer Science and Engineering, G H Raison College of Engineering, Nagpur, India

Tin Tri Duong FPT University, Can Tho, Vietnam

J. Durgashree Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Fazri Fahrezi Alam Sutera, School of Information System, Bina Nusantara University, Jakarta, Indonesia

Ummatul Fatima Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh, India

Steven Gusda Firnandes Alam Sutera, School of Information System, Bina Nusantara University, Jakarta, Indonesia

C. Gambardella Benecon University Consortium, Naples, Italy

Amrita Ganguly Assam Engineering College, Guwahati, Assam, India

Ford Lumban Gaol Computer Science Department, Binus Graduate Program—Doctor of Computer Science, Bina Nusantara University, Jakarta, Indonesia

M. Gayathry Devi Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

B. Geluvaraj Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

Shlok Gilda Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, USA

Vikas Goel KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

Vinaya Gohokar School of Electronics and Communication Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Shannon Gomez Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

Yuri Gordienko National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

S. Gowrishankar Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

K. Gowtham Kongu Engineering College, Perundurai, Erode, India

Carlos Alberto Ramos Guña Ez English School, Ambato, Ecuador

Abhishek Gupta Sardar Patel Institute of Technology, Andheri, Mumbai, India

Rohit Gussain Department of Computer Science and Engineering, Noida Institute of Engineering and Technology, Noida, India

S. Hareesh Chandran Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India

T. Hariharan Kongu Engineering College, Perundurai, Erode, India

Venkatesh Hegde Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

K. Hemalatha Kongu Engineering College, Perundurai, Erode, India

Fonny Hutagalung University of Malaya, Kuala Lumpur, Malaysia

Rafif Imam School of Information System, Bina Nusantara University, Alam Sutera, Jakarta, Indonesia

Aria Muhammad Iswardhana Alam Sutera, School of Information System, Bina Nusantara University, Jakarta, Indonesia

Rex Rony Jacob Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

M. Jaiganesh Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, India

Kurunandan Jain Center for Cyber Security Systems and Networks, Amrita School of Engineering, Vallikavu, Kerala, India;
Cybersecurity Systems and Networks, Amrita Vishwa Vidyapeetam, Kollam, Kerala, India

Tanvi Jain Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, USA

Prajakta Joglekar School of Electronics and Communication Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Jisha John Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

M. John Basha Department of Computer Science and Engineering, Jain Deemed to-be University, Bangalore, India

Shashank M. Kadiwal Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Mufti Ikhshan Kamil Alam Sutera, School of Information System, Bina Nusantara University, Jakarta, Indonesia

Rohith Kapudasu Amrita Vishwa Vidyapeetam, Kollam, Kerala, India

G. Karthikeyan Department of Computer Science, Periyar Arts College, Cuddalore, Tamil Nadu, India

K. Keerthanaa NHCE, Bengaluru, India

Nayeem Ahmad Khan Faculty of Computer Science and Information Technology, AlBaha University, AlBaha, Saudi Arabia

Shirish S. Kulkarni Department of Instrumentation Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, India

Krishan Kumar Department of Software Engineering, Delhi Technological University, New Delhi, India

Nitin Kumar Department of Computer Science and Engineering, Noida Institute of Engineering and Technology, Noida, India

T. Kumar Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

Vineet Kumar Ramaiah Institute of Technology, Bengaluru, Karnataka, India

S. Kumar Chandar School of Business and Management, CHRIST (Deemed to Be University), Bangalore, India

S. M. Kural Eniyavan Kongu Engineering College, Perundurai, Erode, India

Achmad Rizky Kurniawan School of Information System, Bina Nusantara University, Alam Sutera, Jakarta, Indonesia

Arun Kumar Kushwaha KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

Rishabh Lahkar Assam Engineering College, Guwahati, Assam, India

Dung Ngoc Vo Lam FPT University, Can Tho, Viet Nam

Jessica Adine Larissa School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

S. Lokesh School of Computer Science and Engineering, VIT, Chennai, India

Huong Hoang Luong FPT University, Can Tho, Vietnam

Pothuka Madhuri Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

G. Mageshkumar Kongu Engineering College, Perundurai, Erode, India

S. Mageshwari Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

G. Maha Lakshmi SRM University-AP, Andhra Pradesh, Mangalagiri, India

D. Maher Shalal Kongu Engineering College, Perundurai, Erode, India

R. Mahesh Kongu Engineering College, Perundurai, Erode, India

Rohit Mahesh Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

B. Maheswari Department of CSE, Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India

Chinta Sai Manasa Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Souraneel Mandal Adamas University, Kolkata, India

H. Mangalam Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Nikhita Mangaonkar Sardar Patel Institute of Technology, Andheri, Mumbai, India

V. Manikandan Department of Computer Science and Engineering, Jain Deemed to-be University, Bangalore, India

T. Manoji Kongu Engineering College, Perundurai, Erode, India

P. Maruthi Sai Saketh School of Computer Science and Engineering, VIT, Chennai, India

Elizabeth Mathew Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

Tokuro Matsuo Advanced Institute of Industrial Technology, Tokyo, Japan

S. Mayukha Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India

S. Meishree Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

R. I. Minu SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

Girish Mishra Defence Research and Development Organization, Delhi, Delhi, India

Ritiksha Modi Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Navya Mohan Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

P. T. Monisha Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

Ziaul Hasan Mozumder Department of Computer Science and Engineering, Leading University, Sylhet, Bangladesh

S. Muruganantham Kongu Engineering College, Perundurai, Erode, India

N. Muthukumaran Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

C. M. Nalayini Department of Information Technology, Velammal Engineering College, Chennai, India

M. Natarajan Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

Hai Thanh Nguyen Can Tho University, Can Tho, Vietnam

Hy Khang Vuong Nguyen FPT University, Can Tho, Viet Nam

Khoi Tuan Huynh Nguyen FPT University, Can Tho, Viet Nam

Phuong Thuy Thi Nguyen Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

Tong Duc Nguyen FPT University, Can Tho, Vietnam

S. Nikkath Bushra Department of IT, St. Joseph's Institute of Technology, Chennai, Tamil Nadu, India

G. Nirmala R.M.D. Engineering College, Kavaraipettai, Tamil Nadu, India

C. Nithyeswari Department of Computer Science, Periyar Arts College, Cuddalore, Tamil Nadu, India

Evelyn Paulina Rovalino Ortega High School "Madre Gertrudis", Ambato, Ecuador

R. Parente Benecon University Consortium, Naples, Italy

Pooja Pathak Department of Mathematics, Institute of Applied Sciences, GLA University, Uttar Pradesh, Mathura, India

Sanjay Patidar Department of Software Engineering, Delhi Technological University, New Delhi, India

Ivan Pavliuchenko National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Mykhaylo Pavliuchenko Zaporizhzhia State Medical University, Zaporizhzhia, Ukraine

Nhan Trong Van Pham FPT University, Can Tho, Viet Nam

Nghia Trong Le Phan FPT University, Can Tho, Vietnam

Rohit Pilla Cybersecurity Systems and Networks, Amrita Vishwa Vidyapeetam, Kollam, Kerala, India

R. Ponnusamy Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

C. Pooja Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

P. Pooja SRM University-AP, Andhra Pradesh, Mangalagiri, India

Saggurthi Prabhakara Rao Joginpally B R Engineering College, Hyderabad, India

E. Prabhu Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India

Riyan Pradanang School of Information System, Bina Nusantara University, Alam Sutera, Tangerang, Indonesia

Ferdianto Prajnowira School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

B. Prathiksha Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore, India

Armando Rama Putra School of Information System, Bina Nusantara University, Alam Sutera, Jakarta, Indonesia

María Cristina Páez Quinde Instituto Superior Tecnológico España, Ambato, Ecuador

Muhammad Resnandya Rafif School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

G. Raghuraman SSN College of Engineering, Chennai, India

M. Rajalakshmi Faculty of Information Technology, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

B. Pradheep T. Rajan Department of ECE, Francis Xavier Engineering College, Tirunelveli, India

Bharati Ramteke Department of Computer Science and Engineering, G H Raisoni College of Engineering, Nagpur, India

A. R. Ranjitha NHCE, Bengaluru, India

Prasanna V. Rao Deshpande School of Business and Management, CHRIST (Deemed to Be University), Bangalore, India

Anshay Rastogi KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

Aswathy Ravikumar School of Computer Science and Engineering, VIT, Chennai, India

Partha Pratim Ray Department of Computer Applications, Sikkim University, Gangtok, Sikkim, India

Adhitya Rayhan School of Information System, Bina Nusantara University, Alam Sutera, Tangerang, Indonesia

M. Revathi St. Joseph's Institute of Technology, Chennai, India

S. Revathi NHCE, Bengaluru, India

Vijitha Robinson Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

Rully Rubianto School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

B. Sabitha Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India

Sajib Saha Adamas University, Kolkata, India

Syed Sahabuddin Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala, India

C. Sahana Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

J. S. Saleema Department of Computer Science, Christ Deemed to Be University, Bengaluru, Karnataka, India

Subia Salma Ramaiah Institute of Technology, Bengaluru, Karnataka, India

T. Sandhya Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

K. Santhosh Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

D. V. Guru Saran Center for Cyber Security Systems and Networks, Amrita School of Engineering, Vallikavu, Kerala, India

S. Saranya Faculty of Computer Science and Engineering, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

N. Sasikala Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore, India

A. R. Sathyabama Department of Information Technology, Velammal Engineering College, Chennai, India

Manish Sewariya Department of Software Engineering, Delhi Technological University, New Delhi, India

Sana Shahab Department of Business Administration, College of Business Administration, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

Minshul Sharma Department of Computer Science and Engineering, Noida Institute of Engineering and Technology, Noida, India

N. V. Shrivathsa Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Akanksha Singh Department of Software Engineering, Delhi Technological University, New Delhi, India

Pravesh Singh KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

Soumya Singh KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh, India

R. Siva Subramanian Department of CSE, S. A. Engineering College, Chennai, Tamil Nadu, India

M. Sivaraman Department of Computer Science, Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, Tamil Nadu, India

C. Snegha Kongu Engineering College, Perundurai, Erode, India

Pranav Soneji Sardar Patel Institute of Technology, Andheri, Mumbai, India

S. Sony Priya SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

Shreehari Harthi Sreedhara Department of ECE, S J C Institute of Technology, Chikkaballapur, Karnataka, India

A. H. Srinivasa Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Harini Sriraman School of Computer Science and Engineering, VIT, Chennai, India

Sergii Stirenko National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

P. V. Raja Suganya Department of Artificial and Data Science, Velammal Engineering College, Chennai, India

Oggsa Sukardi Master of Information System Management, Binus Graduate Programs, Bina Nusantara University, Jakarta, Indonesia

J. Sumitha Department of Computer Science, Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, Tamil Nadu, India

M. Sureka Kongu Engineering College, Perundurai, Erode, India

R. M. Suruthi Faculty of Computer Science and Engineering, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

S. Suthagar Kongu Engineering College, Perundurai, Erode, India

Hien Thu Thi Ta VNU University of Education, Vietnam National University, Hanoi, Vietnam

Nisarg Thakkar Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Pratham Thakkar Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Shalin Tilva Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Maryna Tomko National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

Sherina Gloria Tompunu School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

José David Sillagana Torres High School “Suizo”, Ambato, Ecuador

Abderezak Touzene Department of Computer Science, Sultan Qaboos University, Muscat, Sultanate of Oman

Trung Tran VNU University of Education, Vietnam National University, Hanoi, Vietnam;
Vietnam Academy for Ethnic Minorities, Hanoi, Vietnam

M. Sarosh Umar Department of Computer Engineering, Aligarh Muslim University, Aligarh, India

Brijesh Vadalia Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Jithy Varghese Department of Computer Science, Christ Deemed to Be University, Bengaluru, Karnataka, India;
Christ Academy Institute for Advanced Studies, Bengaluru, Karnataka, India

Carlos Fabián Martínez Vásquez MecanicS.A./Ambato, Ambato, Ecuador

A. Veena Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

S. Velmurugan Kongu Engineering College, Perundurai, Erode, India

N. Venkatakrisnan Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

Preethi Vijayakumar Department of CSE, Mar Baselios College of Engineering and Technology, Trivandrum, Kerala, India

J. Vindhya Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

R. S. V. S. Vinith SRM University-AP, Andhra Pradesh, Mangalagiri, India

J. Visumathi Veltech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, India

Sunil Wankhade Rajiv Gandhi Institute of Technology, Mumbai, India

Mayur Wanve University of Mumbai, Mumbai, India

Arista Wijaksana School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

Michael Adrian Winarlie School of Information System, Bina Nusantara University, Alam Sutera, Tangerang, Indonesia

Siva Sankar Yellampalli SRM University-AP, Andhra Pradesh, Mangalagiri, India

Rifqi Farhan Zhalifunnas School of Information System, Bina Nusantara University, Alam Sutera, Jakarta, Indonesia

Detecting Wildlife Trapped Images Using Automatically Shared Nearest Neighbouring Pixels (ASNNP)



S Anantha Babu, V. Manikandan, M. Jaiganesh, M. John Basha,
and P. Divya

Abstract Among the most critical problems is minimising vehicle-animal accidents on highways, which cause environmental imbalances and huge public expenditures. This work covers the components of a detect and categorises the species of trapped picture with crop using bbox detect by automatically selecting the shared closest neighbour pixel to detect the large data set detected by MegaDetector. The model automatically selects the weighted average pixel using KNN regression to find the nearest neighbour of SNN density to group the minimal number of points. The ASNNP model crop the trapped image with high accuracy as well as, minimal loss identified in learning rate. The proposed and presented techniques are evaluated based on their ability to meet the mean average precision 8.02 MB (mAP) and detection speed with 94.2 in VGG16.

Keywords Automatic shared nearest neighbour · KNN · Bbox · MegaDetector · Boundary detection · Learning rate · MCRMSE · RMSE · mAP

S. Anantha Babu (✉) · V. Manikandan · M. John Basha
Department of Computer Science and Engineering, Jain Deemed to-be University,
Bangalore 562112, India
e-mail: ananthmtech@gmail.com

V. Manikandan
e-mail: v.manikandan@jainuniversity.ac.in

M. Jaiganesh
Department of Computer Science and Engineering, Karpagam College of Engineering,
Coimbatore 641032, India

P. Divya
Department of Computer Science, Loyola College of Arts and Science, Mettala
Tamilnadu 637407, India

1 Introduction

Technology and software improvements have lately provided computer users with value-added services. The relevance of human–computer interaction (HCI) systems is the development of certain systems that interact and react with human-like robots [1]. The picture obtained from traffic surveillance film is likely to feature individuals and other moving objects. As a result, properly categorising the required moving elements for further processing is crucial. Initially, there are two basic sets of strategies for identifying visual features: shape-based categorisation and movement-based categorisation. Nowadays, the capacity to continually follow or monitor any location is the emerging objective. It is difficult to analyse constant monitoring with the capacity to detect a thing, a person, and an individual [2].

The primary challenge is to develop a dependable, limited, zero-impact system design. Researchers aim to leverage computer vision to automatically extract animal data such as species, number, availability of young, and movement in this research. These activities can be challenging for people as well. Camera-trap photos are rarely perfect, and many of them show species that are either too far away, too close, or just partially observable. Further, changing lighting, reflections, and climate can make data extraction considerably more challenging [3].

Camera traps are used to address a wide range of environmental research and design goals, and also the species that go along with it. There are substantial variances in how cameras are positioned and setup, the sorts of photos collected, [4] why analysts see those images, the attribute data recorded from photos, and how the received data is processed to account for this diversity.

Section 2 begins with an examination of grouping the nearest neighbouring pixel, which serves as the foundation for our segmentation paradigm, as well as a review of automatic selection of multiclass classifier ML model. The suggested automatic selection of nearest pixel grouping introduced and addressed in Sect. 3. Section 4 covers the details of our trials as well as a data analysis. The article is concluded with a brief conclusion.

2 Background Study

Deep neural network-based approaches, which including ResNet [5] as well as its derivatives fast *R*-CNN [6] and faster *R*-CNN [7], have recently attained region object detection performance. Such methodologies typically have two critical parts: (1) image region concept, that searches the actual picture to obtain a set of selection visual features (or object classification) at various levels of scale that could contain additional the image features, and (2) classification techniques, that calculates whether or not these suggested portions are really the artefacts. In the instance of animal recognition utilising camera-trap photos, we discovered two major flaws: frequency and accuracy [8].

Machine learning enables technology to overcome problems despite explicitly programming them to use it. Modern approaches train computers through supervised learning. When categorising photos, for example, the computer is trained on a significant number of combinations of pictures and labels, where the representation is the source and the appropriate attribute is the outcome [9].

CNN was used by Lecun et al. in 1990 to detect arm immobilization causes. The CNN was modeled based on visual context cells of monkeys as researched by Huber and Weist in 1968, which demonstrated that visual cortex cells of a monkey are geographically near and receptive only to a subset of cells in the retina. This sort of cell structure demonstrated that the visual of any item is extremely local, and that despite changes in the peripheral visual, the object remains the same and identifiable by the brain to create their neural network for handwritten recognition, which comprised of a 16×16 px picture as a direct input to the network [5].

R-CNN that is faster: Shortly after the publication of fast R-CNN, the primary bottleneck was overcome utilising region proposal computation [10], which increased the map from 77–70.4% on VOC 2012 [11]. They also created a Region Proposal Network, which shared features with the detection framework and improved the detection pipeline's performance, resulting in an effective running time of only 10 ms.

Jaskó et al. [12] demonstrated a mechanism able to detect numerous huge wild creatures in traffic scenarios. Visual data was collected using a camera with monocular colour vision. The purpose was to examine the traffic scenario picture, identify feature points, and precisely categorise them in order to locate creatures that were on the pathway or could present a hazard. The intensity, colour, and orientation attributes of the traffic scene image were used to build a saliency map. The significant regions on this map were assumed to be points of reference. A massive library of photos of wild creatures was generated. Significant features were extracted from the collected wildlife image dataset and were used for training the classifier for distinguishing between the various classes.

Norouzzadeh et al. [13] explored a significant impediment to researchers and conservationists observing animals in the artificial environment. Applying key generation in recurrent neural networks in object recognition, a system for developing autonomous animal detection was developed in the wild which was proposed with the intention of creating an intelligent species surveillance system. Parham et al. [14] suggested five component recognition process for a wildlife detection method vision based. This technique produced a set of innovative attributes of value (Aov) labelled as species and perspectives. The purpose of this approach was to establish the accuracy and efficiency of species studies while also providing environmentalists with better wild environments.

Besteiro et al. [11] investigated the consistency between 2 distinct occurrence measuring: an infrared sensors device that utilised the device's digital data to perform observations and qualitative perceptions of movement in a sample of 50 weaning piglets on a farm located. The camera's location enables the recording of the main transverse motions with respect to the sensor's rotation, which improved the capability. Humans observed two sorts of behavioural activity: nutrition and recreation.

Gupta and Verma [15] suggested a dictionary learning approach for detecting visual wild animals in photos. Dictionary with illegal to discriminate characteristics training was used to learn racially biased aspects of favourable results containing creatures in the optimistic class and negligible images containing no animals in the target class. The system created category dictionaries and was capable of extracting features using image samples as learning.

Vehicle licence plate identification technique based on convolutional neural network (CNN) and k -means clustering. This approach consists of three primary steps: detection and classification using k -means clustering, and identification of the number plate using a CNN model [16].

The classification capsule and the main capsule are coupled, and the logarithmic shared perceptions are used in conjunction with a Softmax activation function to generate pairing values [17].

To conduct an experimental study, a quantitative computation using the post-hoc analysis and the Friedman experiment is necessary. Whenever a large number of data sets are involved, such analyses allow for a recommendable of analysis [18].

These automatically shared nearest neighbourhood pixel identified the matching pixel and classify minimalistic encoding of image samples in the recommended approach (creature class and noisy backdrop class).

3 System Model

3.1 Proposed Architecture

Figure 1 shows the working flow of automatically shared nearest neighbourhood pixel.

3.2 Theoretical Idea: Shared Nearest Neighbouring Pixel (SNN)

Graph-based clustering techniques are also identified as shared nearest neighbours cluster analysis methods. The rank of a node in the closest neighbour list of another node is represented by the edge among pair of nodes in these approaches. The classification is determined by a fundamental similarity metric. A measure of similarity or distance between data items is required to facilitate clustering, although clustering is then highly dependent on density and resemblance. When the number of dimensions rises, these notions become more difficult to describe. The Euclidean distance, often known as the L_2 norm, is the most commonly used distance metric in low dimensions. In high-dimensional data, even the classic Euclidean definition of density,

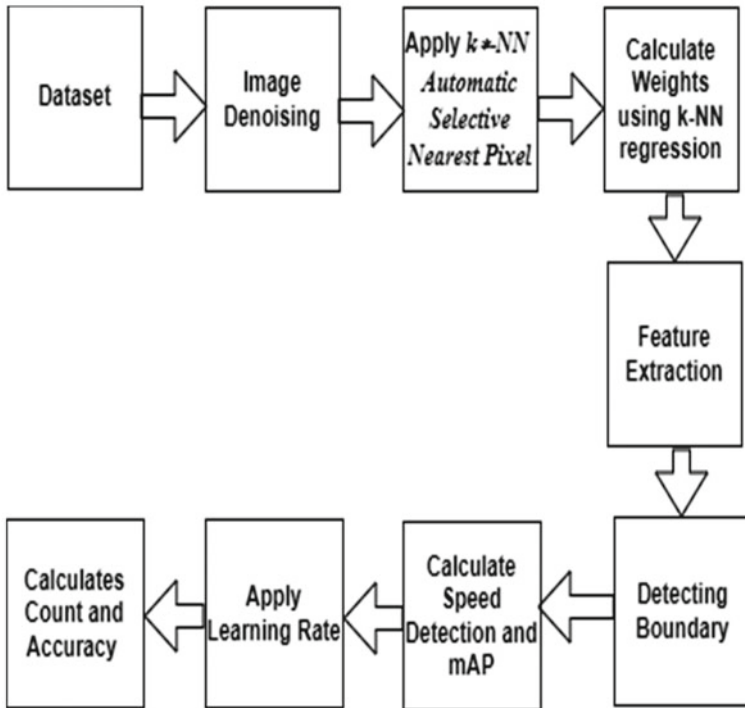


Fig. 1 Proposed architecture

which is the number of points per unit volume. The volume grows fast as the dimensions grow, and if the number of points does not expand with the dimensions, the efficiency approaches nil, as a result of low-density regions in higher dimensions. As a result, in low-density regions in high dimensions.

3.2.1 Generate Boundary Point Detection

The point p 's volume attractor is the position in its Eps-neighbourhood NEps with the maximum density (p). The saturation attractor of point p , for example, is shown in Fig. 2. We must emphasise the following: (1) When there are no other positions in NEps(p) besides p itself, then p is clearly a noisy source, thus we specify the volume fixation of p as p itself and set the threshold degree of p to the smallest value. (2) If NEps (p) contains numerous volume attractors, the first point found in NEps (p) is chosen as the density attractor of point p . A user-specified parameter, $\text{MinP } t_s$, is used to determine the core point, and the core points, that is all locations with an SNN density greater than $\text{MinP } t_s$, are examined. Figure 2 shows boundary detection point.

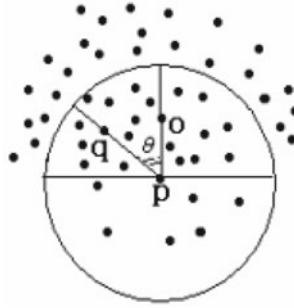


Fig. 2 Features of boundary detection point

3.3 *Automatically Selecting the Number of Nearest Neighbour*

Based on the test feature vector X , the method determines the number of clusters k or the kernel throughput h . To emphasise their dependency on x , we denote k as $k(x)$ and h as $h(x)$. In the case of KNN regression with coefficients.

Pseudocode for Selecting k * -NN-based Automatic Shared Nearest Neighbour

1. The test feature vector X that has been provided (x). The KNN is represented by $(X_{(i)}(x), Y_{(i)}(x))$ as the i th closest training data point to x inside the training data $(X_1, Y_1) (X_n, Y_n)$

$$\widehat{nk} - NN(x) = \frac{1}{k} \sum_{i=1}^k Y_{(i)}(x)$$

2. Generates a shared nearest neighbour graph for a specified k . The edge weights are determined by the number of shared k nearest neighbours (in the range of $[0, k]$)

3. Adjacent neighbours should be given greater weight (recognise that using neighbours closer to x seeks to minimise bias in the KNN regression prediction at x , as the answer)

minimise $\alpha_2 + \alpha^T \beta$, such that $\sum_{i=1}^n \alpha_i = 1, \alpha \in [0, 1]^n$

4. Determine the SNN density of each point, which is the number of points that are similar to their nearest neighbours

5. Identify the core points, which are any points with an SNN density larger than the minimal number of points

6. Apply boundary points to the clustered that have been constructed from the core points

7. Finally, utilising the dynamical issue, group the nearest pixels. to solve the low resolution

Fig. 3 k^* -NN algorithm choosing nearest neighbours

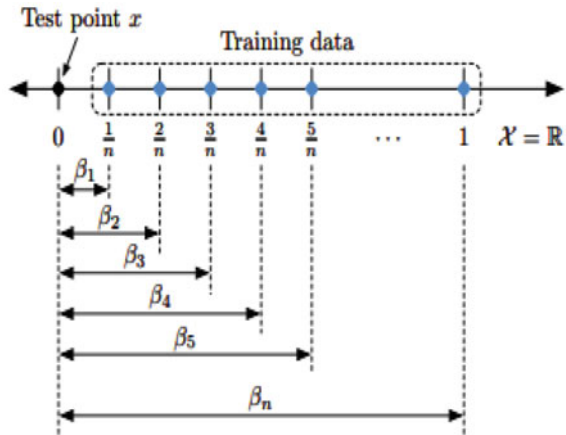


Figure 3 depicts how the KNN algorithm selects the number of nearest neighbours to match the pixel.

Pseudocode for Crop the image

1. The model begins by cropping the image based on the bbox detected by MegaDetector
 2. The correct solution labels are provided as labels in the training data, and we can use them to train a model
 3. Use the learned model to classify the cropped photos of the testing data
 4. We select the species of animals and counts of the image with the maximum count with many in the same image
-

4 Experimental Result and Analysis

Researchers used the Caltech Camera Traps (CCT) data set as our training data set, which comprises 13,553 camera-captured animal pictures from the tropics. For testing, we took 1712 data points from the same data set. The need for entire images or local fixes from large databases is escalating. The suggested automated selected closest pixel technique allows for the embedding of unknown properties while simultaneously assuring that the approach runs in linear time. Our technique's features are then utilised to perform visual search tasks in order to locate the most comparable objects in a repository. In terms of efficiency and accuracy, this method outperforms linear scan search.

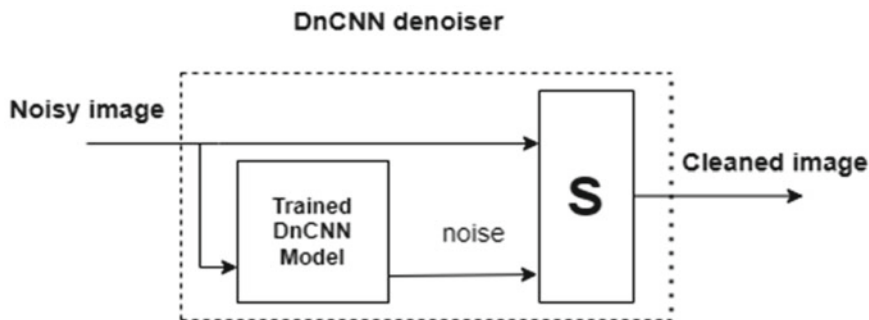


Fig. 4 DnCNN denoiser

4.1 Image Denoising Using CNN

Noise reduces pixel density, which might result in incorrect interpretation of key information. Evaluating chaotic photos is difficult, both programmatically and manually. Noise reduction techniques are often used in information communication systems to minimise the effect of noise on delivered images. In the proposed model, a group of pixels was reduced to one pixel using a nonlinear transformation. Pooling is a vast collection of ‘channels’ carrying a small quantity of data that are defined as the most conceptual ideas revealed from the original image after several iterations of image convolution. The information was then aggregated and sent to a regular, fully connected layers. The residual image is accurately predicted by DnCNN. With operations in the buried layers, it destroys the latent clean image. As a result, in an infocommunication system that sends a noisy image, DnCNN may be utilised to create a correction signal. Figure 3 shows the DnCNN denoiser. It eliminates the latent clean image with procedures in the hidden layers. As a consequence, in an information transmission system that delivers a messy image, DnCNN might be used to generate a correction signal. Figure 4 shows the DnCNN denoiser.

4.2 Feature Extraction

The characteristics were obtained using the following colour spaces: RGB, LAB, HSV, GRAYSCALE, and ENTROPY. The model examined all of the potential permutations of these space colour schemes without duplicating any of them, yielding the five components. Because the computing interface employs a variety of basic colours, the RGB colour space is the most often used for machine graphics. Each and every pixel also on console is made up of three red, green, and blue dots that have been triggered. In contrast, the RGB colour space is inadequate for dealing with shadows and darkness. We have decided to transform RGB image data to a separate homogeneous vector space.

$$E_j = -(P_j \log p_j)$$

where p_j indicates the possibility of the j th element where p_j denotes the probability of the j th piece occurring in a training set. The feature matrix M is used as a representation of the images to compute the entropy of each characteristic, for each membership encoding a pixel value. The M column was scanned from the top-left corners pixel to the corresponding bottom sector pixel, for each pixel acting as a reference pixel for a given region in the image.

4.3 Training and Testing Data

To maximise the most of our limited training data, we used a series of different variations to verify that our classifier never saw the same image twice. This improves the model's generalisability and prevents overfitting. Keras' ImageDataGenerator class may be used for this. Configure random alterations and normalisation techniques on image data throughout training using this class to create makers of reinforced image batches. It creates the layout of our deep neural network classifier employing automated closest pixel selection.

We can collect data for every wildcam picture in images values. Wildcam will capture many frames in a row. The value of the key 'seq num frames' is the number of frames, 'id' is the picture's id, and 'seq id' is the ID connected with the consecutively captured image. This 'seq id' corresponds to the 'Id' in the entry document. Using hundreds of thousands of bounding boxes from diverse ecosystems, this model is trained to recognise animals, humans, and cars in camera-trap photos. Table 1 presents detection level with label Id.

There are 263,504 detection data points, indicating that all train and test data has been processed. As a consequence, the model will not have to run MegaDetector to tune the MegaDetector's weights or re-estimate the findings, as shown in Fig. 5.

Table 1 Performance evaluation to detect the animal using bbox

Detections	Id	max_detection
[{'category': '1', 'bbox': [0.6529, 0.5425, 0...	905a3c8c	0.986
[{'category': '1', 'bbox': [0.0147, 0.0, 0.985...	905a3cfd	0.696
[]	905a3cee	0.000
[{'category': '1', 'bbox': [0.0, 0.4669, 0.185...	905a3cff	1.000
[{'category': '1', 'bbox': [0.0, 0.0494, 0.528...	905a3ccc	0.956

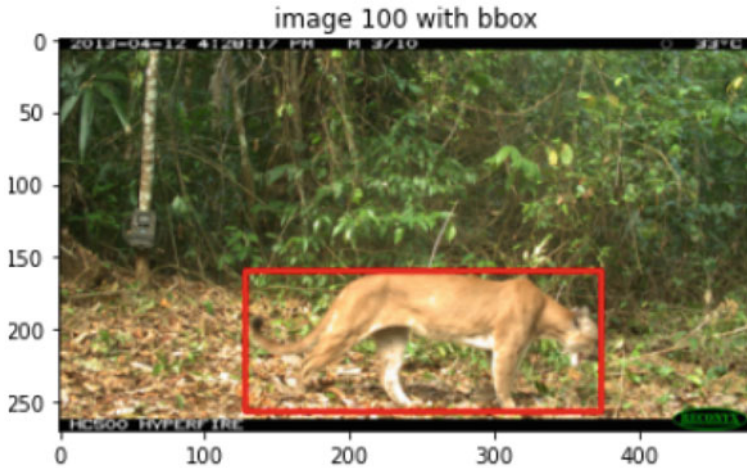


Fig. 5 Image with 100 bbox

4.4 Evaluation Metrics

Mean Columnwise Root Mean Squared Error will be used to evaluate the data (MCRMSE).

$$\text{MCRMSE} = \sum_{j=1}^m \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - y_{ij})^2}$$

where j is a species, I denotes a sequence, x_{ij} denotes the projected count for that species in that sequence, and y_{ij} denotes the ground truth count. This index measures the RMSE for each species and averages it across species. mAP is assessed using ‘Weighted-Distance-Cluster’, which outperforms the traditional ML model. (a) Every patch in Y_{ij} is correlated with several patch transformations learned from the training set in order to solve the poor resolution in SNN. (b) In X_{ij} , local associations between patches should be retained, and this should be the case in Y_{ij} as well. (c) Y_{ij} constrains neighbouring patches by overlapping them to ensure local smoothness and consistency. Using the wildcam2021 data set as a benchmark, Table 2 evaluates metrics. Compared to other classic ML models, our suggested technique appears to be more accurate.

Instead of a precise mAP, we choose to reduce segmentation error. As a result, we employ Dempster’s theory of evidence, which is in addition to the terms of the energy function. We use common benchmark photos to compare our method to others and show that our grouping outperforms others. In this section, we will build a model with the CNN learner function, which will automatically download and load the pre-trained weights. The fastai package implements a learning rate finder as stated in this article. This enables us to select the best learning rate for effective training.

Table 2 Evaluation metrics based on automatic selecting nearest pixel with traditional ML model

Model	E_j (Loss)	Speed (detection)	MCRMSE	RMSE	mAP@0.5
VGG16	0.48	94.2	74.2	192.3	8.02 MB
SqueezeNet	0.49	92.3	70.6	156.2	8.04 MB
MobileNet	0.50	92.0	58.2	168.0	8.09 MB
YoloV3 Lite MobileNet	0.51	91.4	60.3	170.2	8.39 MB
Tiny YoloV3 Lite MobileNet	0.52	90.5	62.4	212.4	5.19 MB
YoloV3 Xception	0.52	90.0	61.4	200.0	105.37 MB
ssd_mobilenet_v1_coco	0.52	89.4	54.2	117.3	108.05 MB

In a nutshell, the learning rate is modified throughout a single epoch, and the loss is displayed against it. When the loss drops the fastest, the learning rate is ideal. The ideal learning rate with loss of testing data is depicted in Fig. 6. Our suggested model achieves the learning rate of 2.6, which has been identified as the minimal loss of knowledge. One of the first and most critical factors of a model is the learning rate. It regulates the size of the leaps used to create the model and how quickly it learns after that. Neural networks train by executing gradient descent on given data over a network of vertices, resulting in a series of nodal weights that are predictive of the elements in the target of a particular label ID when aggregated. Researchers want to know how data is spread over time since animals vary their activity on a regular basis. Train data includes test data in terms of time point. Figures 7 and 8 show the number of animals identified as month data. The time date dataframe in the wildlife data set is mapped using a lambda function.

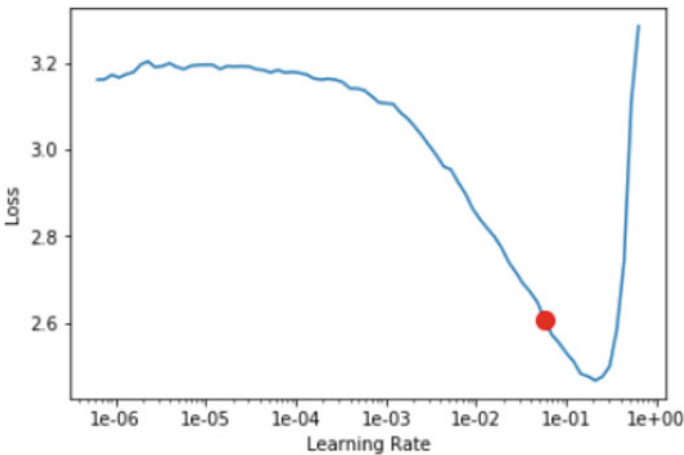


Fig. 6 Optimal learning rate

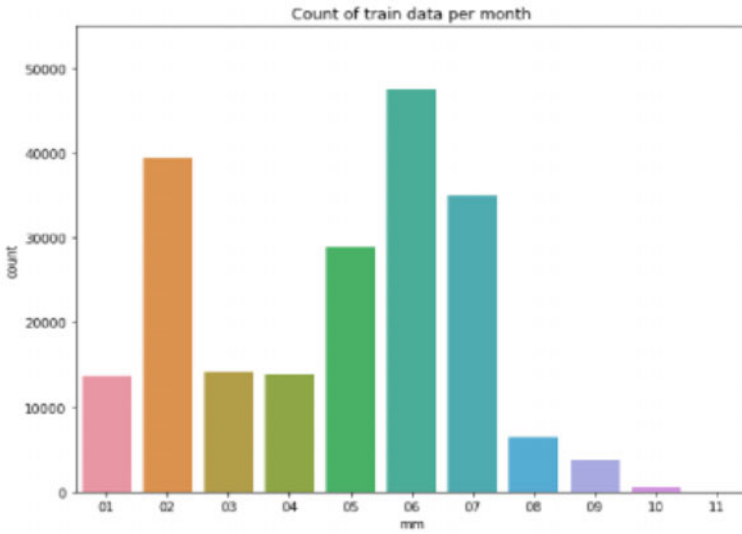


Fig. 7 Count test data per month

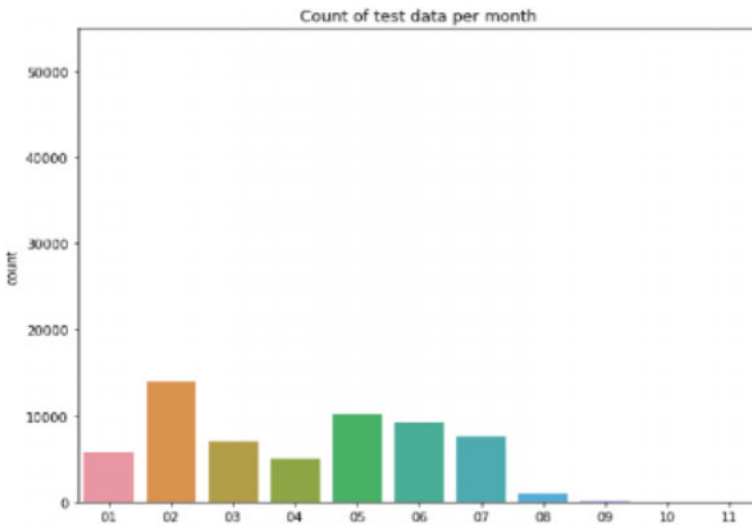


Fig. 8 Count train data per month

5 Conclusion

We revealed that computer vision object identification algorithms can be utilised efficiently to distinguish and classify creatures based on photo series recorded by

surveillance cameras in behaviour, which are notorious for their high levels of noise and chaos. This study investigates an image clustering procedure based on a shared closest neighbours strategy that allows for the handling of clusters of various sizes and shapes with automatically detect wildlife caught animal photos analysis with count.

References

1. Sundar SS, Bellur S, Oh J, Jia H, Kim HS (2016) Theoretical importance of contingency in human-computer interaction: effects of message interactivity on user engagement. *Commun Res* 43(5):595–625
2. Zuboff S (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile books
3. Schneider S, Taylor GW, Kremer S (2018) Deep learning object detection methods for ecological camera trap data. In: 2018 15th Conference on computer and robot vision (CRV). IEEE
4. Breckenridge RP, Dakins M, Bunting S, Harbour JL, White S (2011) Comparison of unmanned aerial vehicle platforms for assessing vegetation cover in sagebrush steppe ecosystems. *Rangeland Ecol Manage* 64(5):521–532
5. Hussain L, Saeed S, Awan IA, Idris A, Nadeem MS, Chaudhry QU (2019) Detecting brain tumor using machines learning techniques based on different features extracting strategies. *Curr Med Imaging* 15(6):595–606
6. Pal SK, Pramanik A, Maiti J, Mitra P (2021) Deep learning in multi-object detection and tracking: state of the art. *Appl Intell* 51(9):6400–6429
7. Zhu L, Tran D, Sevilla-Lara L, Yang Y, Feiszli M, Wang H (2020) Faster recurrent networks for efficient video classification. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, no 07
8. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vis* 115(3):211–252
9. Sen PC, Hajra M, Ghosh M (2020) Supervised classification algorithms in machine learning: a survey and review. *Emerging technology in modelling and graphics*. Springer, Singapore, pp 99–111
10. Srivastava S, Divekar AV, Anilkumar C, Naik I, Kulkarni V, Pattabiraman V (2021) Comparative analysis of deep learning image detection algorithms. *J Big Data* 8(1):1–27
11. Besteiro R, Arango T, Rodríguez MR, Fernández MD, Velo R (2018) Estimation of patterns in weaned piglets' activity using spectral analysis. *Biosyst Eng* 173:85–92
12. Jaskó G, Giosan I, Nedevschi S (2017) Animal detection from traffic scenarios based on monocular color vision. In: 2017 13th IEEE international conference on intelligent computer communication and processing (ICCP). IEEE
13. Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer MS, Packer C, Clune J (2018) Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc National Acad Sci* 115(25):E5716–E5725
14. Parham J, Stewart C, Crall J, Rubenstein D, Holmberg J, Berger-Wolf T (2018) An animal detection pipeline for identification. In: 2018 IEEE winter conference on applications of computer vision (WACV). IEEE
15. Gupta P, Verma GK (2017) Wild animal detection using discriminative feature-oriented dictionary learning. In: 2017 International conference on computing, communication and automation (ICCCA). IEEE

16. Chen JIZ, Zong JI (2021) Automatic vehicle license plate detection using K-means clustering algorithm and CNN. *J Electr Eng Autom* 3(1): 15–23
17. Manoharan JS (2021) Capsule network algorithm for performance optimization of text classification. *J Soft Comput Paradigm (JSCP)* 3(01):1–9
18. Pandian AP (2021) Performance evaluation and comparison using deep learning techniques in sentiment analysis. *J Soft Comput Paradigm* 3(2):123–134
19. Huber R, Ghilardi MF, Massimini M, Ferrarelli F, Riedner BA, Peterson MJ, Tononi G (2006) Arm immobilization causes cortical plastic changes and locally decreases sleep slow wave activity. *Nat Neurosci* 9(9):1169–1176

A Meta-Classifer Link Prediction Model for False Profile Identification in Facebook



S. Saranya, M. Rajalakshmi, S. Devi, and R. M. Suruthi

Abstract Nowadays, social networks have become part of our everyday routine to connect and communicate to diverse communities. The detection of abnormal vertices in these networks has become more vital in exposing network intruders or malicious users. In this paper, we have included a novel meta-classifier which will notice anomalous vertices in complicated interactive network topology by extracting useful patterns. We identify that a vertex with several link connections will incorporate a higher probability of being abnormal. Henceforth, we choose to apply the link prediction model on the Facebook dataset with high interconnectivity. In each step, we determine abnormal vertices are detected with minimal false positive estimates and better accuracy when compared to alternative prevailing methods. Hence, the method incorporated reveals the outliers in the social networks. We proposed an approach that can identify the false profiles with 93% accuracy and lower false positive rates with sustainable accuracy.

Keywords Social networks · Anomaly detection · Link prediction classifier · Random forest · *K*-nearest neighbor · Facebook

S. Saranya (✉) · R. M. Suruthi
Faculty of Computer Science and Engineering, Coimbatore Institute of Technology, Coimbatore,
Tamil Nadu, India
e-mail: saranya.s@cit.edu.in

R. M. Suruthi
e-mail: suruthi.rm@cit.edu.in

M. Rajalakshmi · S. Devi
Faculty of Information Technology, Coimbatore Institute of Technology, Coimbatore, Tamil
Nadu, India
e-mail: rajalakshmi@cit.edu.in

S. Devi
e-mail: devi.s@cit.edu.in

1 Introduction

In recent years, every person across the globe uses social networks as a platform to publicly share their knowledge and views, as an embodiment of their life. Fake profile identification is a salient problem within the field of social network analysis (SNA) and knowledge discovery (KDD). To maintain the network security and privacy in an interactive environment like social media, it is required to detect structural abnormalities that violate typical behavior of social networks. An outlier is a data point or a collection which deviates so much from the other observations as to arouse suspicions [1, 2]. Generally, normal data follow a salient pattern behavior differentiating it from anomaly. A general overview of graph-based anomaly detection methods is shown in Fig. 1. The anomaly in the graph can prevail as individual data outlier, in groups and disguises its nature on contextual analysis. Detection strategies calibers in using approaches based on statistics, classification, clustering, subspace, community detection, labeling/ranking, and so on.

Graph Based Anomaly Detection	Based on Data instances	Point outlier
		Collective outlier
		Contextual outlier
	Conditions for graph anomaly	abnormal vertex exists
		anomalous edge exists
		unexpected presence of vertex label
		unexpected presence of edge label
		missing vertex
	Detection Strategies	missing edge
		Statistical approach
Labeling and Ranking approach		
Classification		
Clustering		
	Subspace	

Fig. 1 General overview of anomaly detection in social network

Studies show that the vertices which deviate from the typical behavior of social networks offer necessary insights into a network. There is a vast amount of bogus id existing in social networks than legitimate users present in different communities. In order to identify the fake profiles which are a serious threat to the secure use of social media, it is highly mandated to detect the anomaly behavior. Graph data reveals inter-dependencies that should be accounted for during the outlier detection process. When compared to traditional data analysis, graph data processing is highly beneficial. They maintain significant interconnections between data in the long run and are flexible to change.

In this paper, anomalous vertices in graph topology are identified by two phases. The primary phase predicts the edge probability in the network by applying a link prediction classifier. Next phase produces the new features set as output. The contributions of the work:

- Deploy sampling of vertices and edges using positive and negative sample strategy.
- Generate a link prediction classifier to find anomalous vertices in a graph
- We adopt a Meta classifier model in the training and testing phase.

2 Related Work

Graphical data analysis has grown over the past years and hence has increased the need for research in social networks. In a graph, data structure substructure reoccur [3]. Hence, it was proposed that the anomalies are substructures that occur infrequently. To understand the problems associated with fake profile creation, several detection methods have been discussed [4]. Most research contributions in machine learning and deep learning prevail but still fake accounts thrive in social networks [5]. In network sampling, there is a comparison of two alternatives for sampling networks that have a predefined number of edges [6]. A random selection of edges with uniform distribution and the edges of the 1.5 ego-networks finds network hubs that are the nodes with high traffic. The hubs are used as they are showing new information due to their high degree. However, the use of hubs may result in major bottlenecks in the networks. Infiltration of fake users into social networks using random friend requests [7]. Feature extraction of the networks are done based on the degree of user, connections that exist midst the friends of the user, amount of communities the user is coupled to. However, this may result in high anonymity level and may fail in detecting the spammer profiles. Feature extraction uses principal component analysis [8]. Main components in the networks are evaluated by reducing the total number of variables and also reduces the dimensions of all observations based on the classification of alike observations. Communication structure among the variables is also found by PCA. However, another dilemma in the PCA method is based on the

selection of core components. Relational dependency networks accomplish collective classification of attributes of the designated variables and are assessed based on a joint probability distribution model. The technique should be trained with more unstructured multimedia data, databases, graph-based analytics, and conception for improved results.

Detection of strange nodes in bipartite graphs involved calculation of normality scores [9] created on the neighborhood relevance score where nodes with a minor normality score had a higher probability of being irregular. In machine learning, using simple heuristics or more refined algorithms is created that is based on the strength of the connection between the users [10]. The first evaluation method involves splitting of datasets hooked on training sets and testing sets. In the second evaluation method, the goal was to measure the classifiers recommendations average users precision. The connection-strength heuristic does not propose a general method of identifying which of the users are to be avoided.

Relational dependency networks (RDN) [11] can achieve collective classification when several attributes midst the selected variables are assessed along on the basis of a joint probability model. Multiple Gibbs sampling iterations were used to fairly accurate the joint distribution. The RDN results are nearly as good as the upper bound data-rich condition. The RDN is capable as it becomes possible to predict leadership parts with some degree of accuracy for circumstances in which very little specific data is known about the individual actors. The suggested method should be qualified with more unstructured data like multimedia data, graph data which will give imagining for improved results. In profile similarity technique [12] similar user profiles are grouped on the basis of profile attributes. User profiles are validated remotely. Using this strategy will drastically improve the search speed of a profile and gradually reduce the memory consumption.

Randomwalk [5, 13] extracts random path sequence from graph. Randomwalk is often used as a part of algorithms that create node embeddings. Deepwalk [14] builds upon a random walk approach which aids to learn latent meaningful graph representations. After the online clustering, anomaly detection is done. If the vertex or edges are far from all current clustering centroid points, then it will be claimed as abnormal. Deepwalk [15] is scalable [4]. Results show that when compared to spectral methods, this approach shows significantly better results for large graphs. This method is designed to operate for sparse graphs too. Netwalk [16], predicts anomalous nodes on dynamic traversal of nodes. As the graph network evolves, the network feature representation is updated dynamically. SDNE [16] and dLSTM [17] are deep network embedding models that facilitate learning vertex representations. It studies the network behavior using autoencoder and locality-preserving constraints.

A generic link prediction classifier [18, 19] aims to detect anomalous vertices by estimating the probability of edge absence in the graph. Based on features collected from graph topology, for each vertex the average probability of missing edges is calculated. Nodes with the highest average probability are predicted as anomalous nodes. This method was not applied on bipartite and weighted graphs [20, 21]. Label propagation algorithms (LPA) [22] is the fastest algorithm that involves labeling nodes in graph structure. The core idea is to handle highly influential nodes through label propagation. On edge traversal, we find the nodes with the more link connectivity are inferred as leader nodes or influential nodes. LPA is applicable to static plain graphs and it suffers from inconsistency and unstable behavior. Random forest algorithm [23–25] is a classification method to find link patterns in a given graph structure. It is handled on static plain graphs. This method performs probabilistic ranking to find irregular vertices with good accuracy prediction results. Node similarity communication matching algorithm [26] is proposed to identify the cloned profile in online social network (OSN). It monitors the behavior of profile and finds for similarity matching with recent activity with 93% detection accuracy. It outperforms well over K-nearest neighbor (KNN) and support vector machine (SVM) methods.

In this section, many researchers have contributed their machine learning strategies to solve the problem domain. Most of the strategies either use feature reduction or feature extraction to identify the characteristic behavior of the graph data. The analysis is focused to find patterns from node interconnections as tracing malicious behavior is problematic. Though KNN and random forest classification provides better prediction independently, the performance aspect on integrating the classifier to anomaly detection is targeted.

3 Proposed Methodology

The input of the proposed system is a graph dataset that represents the social networks in the form of nodes and edges. The work flow of the proposed methodology is shown in Fig. 2. The working model is categorized into three stages involving anomalous node identification processes from the original graph.

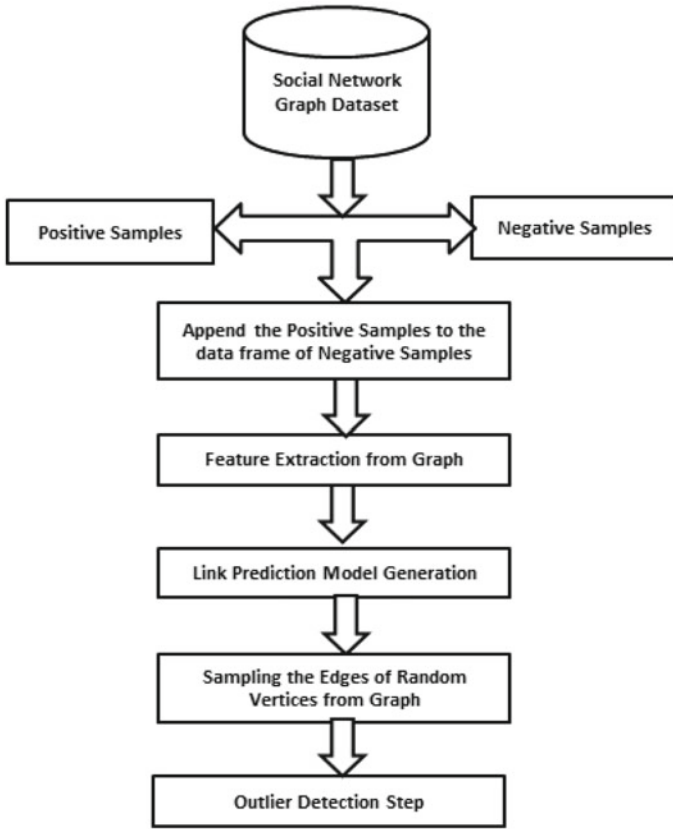


Fig. 2 Work model of the proposed methodology

3.1 Phase 1: Building a Classifier Model for Edge Connectivity

For every different graph manifest, a link prediction classifier is built which predicts a link between two unconnected nodes. For this, a training dataset is prepared from the graph that consists of negative and positive and samples. Negative samples are the major part of the dataset as they consist of the unconnected node pairs. Both Algorithm 1 and 2 shows steps in generating the positive and negative samples from graph input data, respectively.

Algorithm 1

Negative samples generation

Input: Graph G with node pairs

Output: Negative sample pairs

Get unconnected node-pairs

all_unconnected_pairs \leftarrow new list()

Traverse adjacency matrix

Initialize offset \leftarrow 0for each $i \leftarrow$ tqdm(range(G[0])) do for each $j \leftarrow$ range(offset, G[6]) do if ($(i \neq j)$ and ($\text{shortest_path_length}(G, i, j) \leq 2$)) and ($\text{adjacency}(i, j) = 0$) then all_unconnected_pairs \leftarrow append([node_list[i], node_list[j]])

end if;

offset = offset + 1

end for;

end for;

Algorithm 2

Positive samples generation

Input: Graph G with node pairs

Output: Positive sample pairs

Initialize

 initial_node_count \leftarrow len(G nodes)

#copy of nodes pairs in facebook graph dataframe

 fb_df_temp \leftarrow fb_df

#empty list to store removable links

omissible_links_index \leftarrow new list()for $i \leftarrow$ tqdm(fb_df.index.values) do

remove a node pair and build a new graph

 G_temp \leftarrow fb_df_temp.drop(index = i)

check if there is no splitting of graph and number of nodes is same

if (number_connected_components(G_temp) == 1) and (len(G_temp.nodes) == initial_node_count) then

 omissible_links_index \leftarrow append(i) fb_df_temp \leftarrow fb_df_temp.drop(index = i)

end if;

end for;

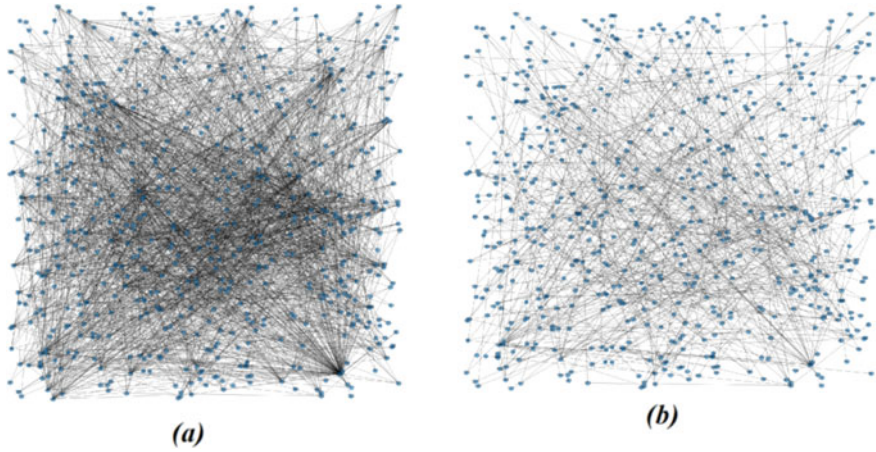


Fig. 3 **a** Original graph, **b** graph from link prediction classifier

Edges are randomly dropped from the graph in a condition that even in the process of dropping edges the nodes of the graph remain connected. These removable edges are appended to the data frame of unconnected node pairs. Feature extraction is performed on the graph after the dropping of the links. In the stage to obtain a trained dataset, jaccard coefficients/jaccard index is used to integrate the overlapping features. Let x, y be some graph vertex, the jaccard index is performed by comparing its neighbor nodes which is represented in Eq. 1.

$$\text{jaccard Index}(x, y) = \frac{|\text{neighborhood}(x) \cap \text{neighborhood}(y)|}{|\text{neighborhood}(x) \cup \text{neighborhood}(y)|} \quad (1)$$

After feature extraction to validate the model, the dataset is split for the training and testing of the model's performance. Random forest classifier is employed to generate the link prediction model based on the obtained features. Figure 3 represents the resultant graph after building the link prediction classifier.

3.2 Phase 2: Sampling Vertices from Graphs

In this phase, the samples are generated from a graph using a specific criterion. The threshold is set to be `min_friends` to satisfy the feature extraction process. The calculation of `min_friends` uses the knowledge of degree distribution to set the scale. Algorithm 3 shows the steps to perform the sampling of vertices and edges for a given graph G .

Algorithm 3

 Graph vertex sampling

Input: Graph G with node pairs, min_friends

Output: Selected edge list

```

Initialize
  selected_edges ← new set()
  list_nodes = list(G.nodes())
for each val ∈ list_nodes do
  if len(list(G.neighbors(val))) ≥ min_friends then
    Temp_edges ← new set()
    for node in list(G.neighbors(val)) do
      if len(list(G.neighbors(node))) ≥ min_friends then
        Temp_edges ← add(val)
        Temp_edges ← add(node)
      end if;
    end if;
    if len(Temp_edges) ≥ min_friends then
      selected_edges ← union(selected_edges, Temp_edges)
    end if;
  end for
end if;
end for;

```

3.3 Phase 3: Detection of Outlier from Resultant Graph

In this phase, the link prediction classifier will generate interesting features that enable it to perform outlier detection. This method is centered on the hypothesis that a vertex with many low variance link connections has a higher likelihood of being anomalous. Test set is obtained by performing random edge samples from the graph. The algorithm steps involved in selecting node links that are completely traversed and chooses neighbors with more than minimum friend neighbors. We configure the min friend to be set to three. We omit the nodes with very low neighbors connected as it provides poor network characteristics. These sets of vertices are considered for further evaluation of anomaly detection and the outliers are detected which is represented in Fig. 4.

4 Experimental Evaluation**4.1 Data Preparation**

The experiments are conducted on a system operating with a 64 bit windows platform with Intel quad core processors supporting GPU specification. For experimental

```

threshold = 0.8
anomalies = set()
for v in j:
    if(v[2] > threshold):
        rm.add(v[1])
print(anomalies)

```

```
{11, 19, 20, 22, 24, 27, 28, 31, 34, 35, 36, 37, 38, 39, 40, 41, 43, 45, 46, 48, 49, 50, 51, 54}
```

Fig. 4 Result of graph vertex outlier

evaluation, we use Python-3.6 utilizing machine learning libraries comprising scikit-learn version 0.19.1 and NumPy version 1.14.0 to perform mathematical operations.

We evaluated our algorithm on the Facebook dataset from network repository which has 620 nodes and 2102 edges. The fb_pages_food is an undirected graph with nodes representing the Facebook pages and edges having mutual likes among them. In addition to that we have also added 20 randomly generated edges or anomalies using Erdos_Renyi algorithm. To incorporate anomalous nodes, reindexing of the node pairs can be done in the initial dataset.

4.2 Generation of Anomalous Edges

To generate random anomalous edges in the existing graph, Erdos–Renyi Model is used. In this model, each edge distribution has a fixed probability of being available or missing, regardless of the number of connecting links in the network. The probability of edge distribution existing for our model has been set as 0.5 since it gives an equal distribution of the edges. Figure 5 gives the degree of distribution of the edges across the existing network.

4.3 Result and Experimental Setup

Figure 6 shows the anomalous nodes that are obtained from the proposed algorithm. The node information helps to prevent malicious activity. Nodes are crawled such that outlier nodes are extracted on configuring the threshold value set to 0.8. A node exceeding the limit is outliers and monitoring the network activity of the node is encouraged.

To evaluate the performance metric of the detection model, a confusion matrix is employed. The confusion matrix is used to solve this binary classification as either genuine/ normal node or outlier node. The confusion matrix gives values such as

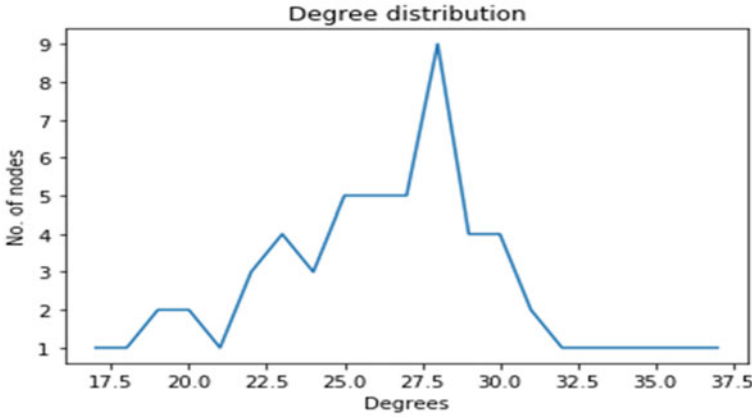


Fig. 5 Degree distribution of Facebook network

```
In [39]: threshold = 0.8
anomolies = set()
for v in j:
    if(v[2] > threshold):
        anomolies.add(v[1])
print(anomolies)

{'78', '221', '422', '87', '541', '603', '260', '34', '35', '36', '37', '38', '39', '40', '41', '487', '42', '44', '390', '252', '43', '25', '207',
'80', '415', '214', '400', '63', '305', '128', '131', '292', '13', '564', '47', '475', '202', '82', '498', '583', '245', '212',
'482', '552', '613', '443', '577', '265', '22', '211', '502', '175', '243', '342', '85', '126', '533', '277', '525', '236', '39
2', '614', '46', '266', '478', '417', '24', '2', '557', '143', '587', '198', '62', '454', '370', '306', '165', '573', '174', '2
59', '41', '517', '602', '489', '593', '274', '604', '432', '431', '465', '234', '537', '594', '23', '499', '35', '258', '444',
'118', '17', '241', '413', '495', '513', '437', '232', '100', '514', '435', '56', '397', '562', '330', '149', '303', '148', '38
6', '116', '582', '504', '122', '96', '317', '321', '442', '610', '136', '509', '607', '21', '104', '217', '179', '322', '430',
'166', '3', '521', '267', '137', '57', '278', '79', '275', '90', '239', '182', '146', '218', '11', '548', '204', '244', '485',
'7', '520', '298', '561', '132', '16', '184', '337', '459', '500', '452', '436', '450', '185', '584', '530', '36', '488', '45
5', '10', '169', '449', '297', '284', '294', '77', '427', '597', '135', '154', '592', '200', '42', '72', '155', '578', '534',
'532', '575', '357', '168', '172', '208', '263', '296', '484', '348', '327', '355', '242', '147', '473', '511', '203', '418',
'240', '343', '28', '124', '591', '608', '356', '167', '27', '50', '12', '29', '605', '350', '152', '406', '279', '468', '428',
'173', '164', '219', '213', '183', '205', '565', '289', '503', '201', '139', '15', '255', '471', '253', '191', '324', '316', '2
15', '151', '14', '94', '251', '412', '524', '501', '51', '171', '66', '368', '223', '611', '569', '97', '469', '540', '103',
'318', '612', '414', '387', '375', '606', '318', '74', '325', '59', '302', '249', '429', '206', '510', '115', '123', '426', '40
5', '119', '407', '345', '367', '419', '409', '398', '464', '70', '364', '556', '369', '560', '84', '550', '445', '568', '385',
'33', '237', '192', '272', '281', '416', '156', '162', '347', '536', '381', '256', '567', '193', '199', '109', '299', '220', '5
38', '111', '53', '522', '323', '434', '523', '93', '588', '404', '194', '566', '270', '366', '187', '421', '470', '161', '15
3', '225', '539', '83', '40', '142', '336', '423', '0', '328', '269', '60', '140', '112', '433', '238', '98', '209', '181', '35
8', '480', '526', '378', '481', '64', '331', '102', '360', '529', '493', '332', '492', '75', '76', '590', '230', '61', '291',
'6', '372', '486', '404', '320', '402', '558', '554', '453', '441', '610', '476', '424', '307', '396', '334', '447', '160', '55
9', '446', '333', '226', '117', '235', '617', '233', '101', '354', '222', '505', '55', '599', '616', '120', '106', '479', '15
7', '376', '37', '65', '580', '353', '544', '45', '535', '527', '384', '531', '425', '600', '89', '108', '344', '362', '438',
'121', '363', '19', '9', '448', '91', '401', '105', '349', '340', '553', '276', '133', '189', '196', '497', '39', '490', '373',
'114', '134', '262', '250', '467', '329', '601', '44', '618', '477', '300', '551', '463', '312', '462', '563', '248', '374', '4
61', '457', '371', '516', '383', '163', '125', '315', '280', '420', '403', '288', '271', '261', '576', '586', '451', '361', '21
0', '43', '52', '290', '287', '69', '67', '518', '138', '31', '319', '144', '346', '190', '382', '178', '127', '257', '496', '3
4', '393', '335', '301', '555', '410', '186', '38', '216', '460', '145', '380', '311', '92', '508', '395', '440', '472', '585',
'341', '543', '379', '8', '95', '466', '99', '58', '246', '572', '30', '113', '188', '326', '515', '231', '574', '570', '377',
'596', '73', '313', '456', '528', '458', '54', '408', '293', '512', '224', '268', '150', '394', '68', '228', '519', '389', '2
6', '229', '388', '282', '304', '338', '71', '118', '399', '180', '273', '129', '391', '295', '507', '571', '474', '130', '28
6', '187', '247', '81', '579', '20', '609', '615', '545', '595', '159', '32', '285', '264', '546', '581', '141', '491', '176',
'170', '197', '359', '439', '506', '177', '314', '195', '542', '339', '254', '483', '283', '589', '411', '48', '158', '351', '3
08', '547', '352', '88', '227', '5', '309', '1', '549', '86', '598', '4', '365'}
```

Fig. 6 Outlier nodes detected

True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The mathematical formulation of the performance parameters are represented in Eqns. 2–5, where TP is nodes correctly classified as outlier, TN is node correctly classified as normal, FP is nodes misclassified as outlier from normal, and FN is nodes misclassified as normal from outlier.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total Nodes} \tag{2}$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \tag{3}$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \tag{4}$$

$$f1 - \text{score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{5}$$

Figure 7 tabulates the *f1*-score, support, precision, and recall for the trained model. The KNN uses the Minkowski metric for this step. We have also measured the error rate for different *k* values (*k* lies between 1 and 40).

Figure 8 depicts that the highest error rate is 11% for *k* = 0 and the lowest error rate of 6% is for *k* = 6. The error rate remains constant (i.e.) 6% for the value of *k* greater than 6. The algorithm has been trained using KNN model, and its accuracy is found to be 93% with the maximum error rate of 0.14. Table 1 summarizes the comparison of the proposed model outperforming other methods [23, 26].

```
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

[[5726	4]				
[416	5]]				
		precision	recall	f1-score	support
	0	0.93	1.00	0.96	5730
	1	0.56	0.01	0.02	421
	accuracy			0.93	6151
	macro avg	0.74	0.51	0.49	6151
	weighted avg	0.91	0.93	0.90	6151

Fig. 7 Classification report

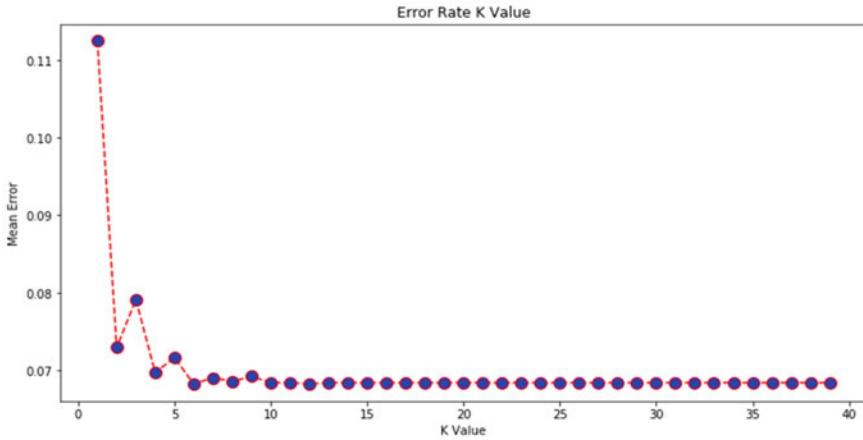


Fig. 8 Error rate calculation

Table 1 Comparison table

S. No	Model Used	Accuracy(%)	Time (ms)
1	<i>K</i> -nearest neighbor (KNN)	89	0.89
2	Support vector machine (SVM)	84	0.78
3	Node similarity communication matching	93.14	0.45
4	Naive bayes	83	0.64
5	Proposed model (Random forest + KNN)	93.495	0.40

5 Conclusion

To understand the behavior patterns of large networks, detection of anomalies are very important. We propose a novel strategy that discovers anomalous nodes on the basis of characteristic traits of the underlying network structure. The method adopts current techniques in machine learning to generate useful communities. We experimented this method in the Facebook social network dataset, but there is considerable scope to be tested with a diverse application. The proposed method is trained using KNN model and its accuracy is found to be 93% with the maximum error rate of 0.14. As for future work, the algorithm can be prolonged to be applied on undirected and dynamic real-world graph datasets. We plan to extend the work to incorporate a neural network learning model so as to increase the performance of the proposed model.

References

1. Hodge VJ, Austin J (2004) A survey of outlier detection methodologies. *Artif Intell Rev* 22(2):85–126
2. Saranya S, Rajalakshmi M (2022) Certain St rategic study on machine learning-based graph anomaly detection. In: Shakya S, Bestak R, Palanisamy R, Kamel KA (eds) *Mobile computing and sustainable informatics. Lecture notes on data engineering and communications technologies*, vol 68. Springer, Singapore. https://doi.org/10.1007/978-981-16-1866-6_5
3. Noble CC, Cook DJ (2003) Graph-based anomaly detection”, *ACM SIGKDD '03*, August 24–27, 2003
4. Roy PK, Chahar S (Dec.2020) Fake profile detection on social networking websites: a comprehensive review. *IEEE Trans Artif Intell* 1(3):271–285. <https://doi.org/10.1109/TAI.2021.3064901>
5. Moonesinghe HDK, Tan PN (2008) OutRank: a graph-based outlier detection framework using random walk. *Int J Artif Intell Tools* 17(1)
6. Altschuler Y et al (2013) Detecting Anomalous behaviors using structural properties of social networks. In: Greenberg AM, Kennedy WG, Bos ND (eds) *Social computing, behavioral-cultural modeling and prediction*, SBP 2013. *Lecture Notes in Computer Science*, vol 7812. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-37210-0_47
7. Fire M, Katz G, Elovici Y (2012) Strangers Intrusion detection—detecting spammers and fake profiles in social networks based on topology anomalies. In: *ASE Human J*
8. Mohammadrezaei M, Shiri ME, Rahmani AM (2018) Identifying fake accounts on social networks based on graph analysis and classification algorithms. *Secur Commun Netw* 1–8. <https://doi.org/10.1155/2018/5923156>
9. Sun J, Qu H, Chakrabarti D, Faloutsos C (200) Neighborhood formation and anomaly detection in bipartite graphs. In: *Fifth IEEE international conference on data mining (ICDM'05)*, p 8. <https://doi.org/10.1109/ICDM.2005.103>
10. Fire M, Kagan D, Elyashar A et al (2014) Friend or foe? Fake profile identification in online social networks. *Soc Netw Anal Min* 4:194. <https://doi.org/10.1007/s13278-014-0194-4>
11. Campbell WM, Dagli CK, Weinstein CJ (2013) Social network analysis with content and graphs
12. Nandhini DM, Das BB (2016) Profile similarity technique for detection of duplicate profiles in online social network
13. Vempala S (2005) Geometric random walks: a survey. *Comb Comput Geom MSRI Publ* 52:573–612
14. Perozzi B, Al-Rfou R, Skiena S (2019) DeepWalk: online learning of social representations. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*. <https://doi.org/10.1145/2623330.2623732>
15. Tan Q, Liu N, Hu X (2014) Deep representation learning for social network analysis. *Front Big Data* 2:2
16. Yu W, Cheng W, Aggarwal CC, Zhang K, Chen H, Wang W (2018) NetWalk: a flexible deep embedding approach for anomaly detection in dynamic networks. pp 2672–2681. <https://doi.org/10.1145/3219819.3220024>.
17. Maya S, Ueno K, Nishikawa T (2019) dLSTM: a new approach for anomaly detection using deep learning with delayed prediction. *Int J Data Sci Anal*
18. Kagan D, Elovichi Y, Fire M (2018) Generic anomalous vertices detection utilizing a link prediction algorithm. *Soc Netw Anal Mining*
19. Fadaee SA, Haeri MA (2019) Classification using link prediction. *Neurocomputing* 359:395–407. <https://doi.org/10.1016/j.neucom.2019.06.026>
20. Fire M, Tenenboim L, Lesser O, Puzis R, Rokach L, Elovici Y (2011) Link prediction in social networks using computationally efficient topological features. In: *IEEE Third Int'l conference on privacy, security, risk and trust and IEEE third Int'l conference on social computing*. <https://doi.org/10.1109/passat/socialcom.2011.20>

21. Lichtenwalter RN, Lussier JT, Chawla NV (2010) New perspectives and methods in link prediction. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining—KDD '10,2010. <https://doi.org/10.1145/1835804.1835837>
22. Bhatia V, Saneja B, Rani (2017) INGC: graph clustering & outlier detection algorithm using label propagation. In: International conference on machine learning and data science 2017
23. Homsy A, Al Nemri J, Naimat N, Kareem HA, Al-Fayoumi M, Snober MA (2021) Detecting twitter fake accounts using machine learning and data reduction techniques. DATA
24. Primartha R, Tama BA (2017) Anomaly detection using random forest: a performance revisited. <https://doi.org/10.1109/ICODSE.2017.8285847>
25. Rahman O, Quraishi MA (2019) Experimental analysis of random forest, K-nearest neighbor and support vector machine anomaly detection. <https://doi.org/10.13140/RG.2.2.19998.18245>
26. Revathi S, Suriakala M (2018) Profile similarity communication matching approaches for detection of duplicate profiles in online social network. In: 2018 3rd International conference on computational systems and information technology for sustainable solutions, pp 174–182

Comparison of Various CNN Models for Image Classification



S. Sony Priya and R. I. Minu

Abstract Classification is one of the core tasks of computer vision. In image classification, the system trained using one of the classification algorithms receives the input image, analyzes the image, and assigns the label based on its training. Image classification provides a vital role in various fields like medical imaging, automated vehicle, social media and many more. This paper compares convolutional neural network (CNN) models like AlexNet, VGG16, InceptionV3, Deep residual networks (ResNet), and DenseNet using the accuracy. Here, the CIFAR-10 dataset is used for comparison.

Keywords Image classification · Convolutional neural network · Computer vision · AlexNet · VGG16 · InceptionV3 · ResNet · DenseNet

1 Introduction

Classifying images is not an uphill battle for humans, but it is strenuous for machines to correctly identify and categorize pictures because of different backgrounds, poses, sizes, and others. Deep learning inspired by the organic nervous system can tackle the problem. Deep learning is based on traditional neural network architecture. Traditional neural networks have fewer layers, but we have more in the deep neural network. The advent of deep learning brought great successes and achieved above human-level performance in image classification. McCulloch et al. [1] took the first step toward artificial neural networks in 1943. The authors used threshold logic to model a primary neural network.

Hebb [2] developed a learning rule in 1949 that outlines how neural activity impacts the connection between neurons, a phenomenon known as neuroplasticity. Hebb's rule is another name for this rule. Within a neural network, it offers a method

S. Sony Priya (✉) · R. I. Minu
SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India
e-mail: sp3454@srmist.edu.in

R. I. Minu
e-mail: minur@srmist.edu.in

for updating the weight of neuronal connections. In 1958, Rosenblatt [3] invented the perceptron algorithm for pattern recognition.

In 1975, the Werbos [4] backpropagation method was used to train the network, a significant catalyst for increased interest in neural networks. Backpropagation reduces the error rate by adjusting the weights at each node backward.

Fukushima and Miyake [5] invented the neocognitron, a hierarchical, multilayered artificial neural network, in 1979. It inspired convolutional neural networks and utilized Japanese handwritten character identification and other pattern recognition applications. LeCun et al. [6] released a paper in 1989 demonstrating how backpropagation is used to train the neural network algorithms. Using a neural network in this study successfully recognized handwritten postal code digits given by the US Postal Service. By this, deep learning has become popular. CNN models are no longer the same as LeNet, despite being all based on it. Because of its accuracy in demanding categorization tasks, CNN has gained popularity. In addition, typical manual image processing approaches for feature extraction are no longer required [7, 8].

CNN is utilized in a wide range of applications. In the paper [9], the authors proposed a CNN model for automatically recognizing license plates in the nighttime, which is helpful for the intelligent transportation system. Ge et al. [10] proposed the AlexNet-Emotion model to find the facial expression in a static and dynamic sequence. Feroz et al. [11] proposed Single Shot Detector (SSD) and You Only Look Once (YOLO) models for object detection and recognition in low-quality videos, which are used for a variety of applications like classification, surveillance is name few. CNN is also applied in multiple agricultural domains like detecting diseases, identification of various plants, Land Cover and Land use Classification (LCCS), automatic fruit or vegetable counting, etc. Roy et al. [12] proposed an advanced version of the You Only Look Once V4 algorithm for detecting four different types of disease: early blight fungal disease, late blight fungal disease, *Septoria leaf* spot, and *Passiflora fulva* in real time. CNN models are used in medical decision support systems. Maweu et al. [13] proposed CNN Explainability Framework for ECG signals (CEFEs) that uses highly structured ECG signals to provide Interpretable explanations. Rehman et al. [14] proposed 3D CNN-based architecture for brain tumor extraction and used VGG19 to classify the tumor type [15–17]. The authors used BraTS 2015, 2017, and 2018 datasets for tumor type classification and achieved an accuracy of 98, 97, and 93%. CNNs are also used in drug discovery. Peng et al. [18] proposed DTI-CNN to forecast how various chemicals will interact with various protein targets. First, they used heterogeneous networks to identify the pertinent characteristics of medications and targets. The DAE model is then used to reduce dimensions and identify the fundamental components. Finally, by using CNN, predictions of DTIs are made. Similarly [18, 19] developed KenDTI, which utilizes CNN for drug-target interaction prediction.

The first portion of this study discusses CNNs history and architecture. The architecture of CNN models is detailed in the next section. The experimental setup and findings are presented in Sect. 3. The conclusion section discusses which model is best for image classification and how accuracy may be improved.

2 Classification Process

For classifying images, we must consider the factors like the type of classification model type of samples by which we will train the model. Then we perform data preprocessing, ensuring the training samples/images have the same size as the classification model. In the preprocessing step, we perform normalization to rescale the pixel values to lie within a restricted range. Often, the amount of data we have is insufficient to complete the classification assignment. We use data augmentation, which involves flipping, rotating, cropping, translating, lighting, scaling, adding noise, etc., to get around this issue. We fed the data into the CNN model when the preprocessing stage was finished. Figure 1 depicts the CNN architecture.

The CNN receives input images, extracts, and learns the features to classify the images. It contains convolution, pooling, and fully connected layers. In the convolution layer, the filters convolved with the input image by taking one or more steps along the x or y -axis of the image to produce a feature map. This step is called stride. This feature map extracts information like edges and corners in the image. Additionally, the convolution technique minimizes the resulting image's size. Consider the situation in which we do not want to reduce the output image size and corner information loss; then, we use padding. Adding extra rows and columns in the input matrix is called padding. Consequently, the formula below determines the size of the output feature map.

$$h_{\text{new}} = \frac{h_{\text{old}} - f_s + s_l + p_s}{s_l} \tag{1}$$

$$w_{\text{new}} = \frac{w_{\text{old}} - f_s + s_l + p_s}{s_l} \tag{2}$$

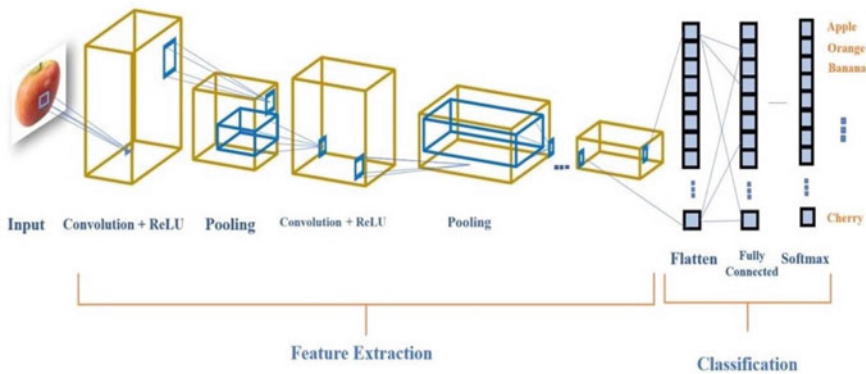







Fig. 1 CNN architecture

Table 1 Activation functions

Activation function	Formula	Plot	Range
Linear	$f(x) = x$		$(-\infty, \infty)$
Sigmoid	$f(x) = \frac{1}{1+e^x}$		(0 to 1)
Hyperbolic tangent	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$		$(-1 \text{ to } 1)$
Rectified linear unit (ReLU)	$f(x) = \text{Max}(0, x)$		$(0, \infty)$
Softmax	$f(x) = \ln(1 + e^x)$		$(0, \infty)$

where, $h_{\text{new}}, w_{\text{new}}$ is new height and width of the feature map, $h_{\text{old}}, w_{\text{old}}$ is the previous height and width of the feature map, f_s is filter size, s_l is stride length, p_s is padding size. The convolution procedure's output is processed via an activation function. To study the intricate interaction between network variables, activation functions are applied. Furthermore, it determines whether or not a neuron will fire. Table 1 provides some key activation functions as well as their formulae.

The convolution layer is followed by the pooling layer. The feature map from the convolution layer is down sampled using the pooling layer. Additionally, the position of the features in the feature maps produced by the convolution layer affects them greatly. As a result, any adjustments to the feature map could result in an erroneous forecast overall. This problem is addressed by the pooling layer. Max pooling and average pooling are the pooling layers that are used the most frequently. The equation below provides the pooling layer's output.

$$\text{output}_{\text{pooling}} = \frac{f_h - F_s + 1}{s * (f_w - F_s + 1) * C} \quad (3)$$

Here, f_h, f_w signifies the feature maps height and width, F_s signifies filter size, C denotes number of channels, s represents stride size.

The pooling layer's output is then flattened into a one-dimensional array by flattening procedure, i.e., a long single feature vector. Next, this vector is sent into the fully connected layer for classification. In a fully connected network, all the neurons are densely connected; also, it is placed at the end of the architecture. The probabilities for each class are calculated using the SoftMax activation function in the last layer of CNN. It also assigns input to the class with the highest likelihood. It is computed by the following formulae

$$\text{softmax}(Z_i) = e^z / \sum e^z \quad j \rightarrow 0 \text{ to } n \tag{4}$$

where, Z_i denotes i th vector value, n denotes number of classes. After finding the class of an object, the loss of the model is calculated. The discrepancy between the actual and anticipated output is known as loss. A good classifier should have the minimum loss. There are various loss functions used to calculate the loss. Binary cross-entropy, soft margin classifier, margin classifier, and negative log likelihood are name few. To reduce the loss, we use optimizers that change the neural network’s weights or learning rate. The most commonly used optimizers are gradient descent, stochastic gradient descent, mini-batch gradient descent, SGD with momentum, AdaGrad, Adadelta, Adam, etc., After implementing the model, we finally measure the model’s performance.

Several methods measure the classification model’s performance. The most popular methods are confusion matrix, accuracy, precision, and recall. The confusion matrix is represented as a table given below, and it is used to identify misclassification. The general format of confusion matrix is depicted in Fig. 2.

Here, True Positive (TP): Predicted values and actual values are True Negative (TN): Predicted values and actual values are FP (Type1 error): Predicted value is true, but actually, it is false FN (Type2 error): Predicted value is negative, but actually, it is true Accuracy is another performance measure used to find how many you got right. It is described as the proportion of accurate predictions to all predictions.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

How accurately the model predicts the positive class that is a member of the positive class is known as precision, i.e., the proportion of TP to every positive result

		Real Value	
		Positive (1)	Negative (0)
Predicted Value	Positive (1)	TP	FP (Type1 error)
	Negative (0)	FN (Type2 error)	TN

Fig. 2 Confusion matrix

the model anticipated.

$$\text{Precision} = \frac{\text{TP}}{\text{Predicted positive value}} \quad (6)$$

The ratio of all the dataset's positives to the TP is known as recall.

$$\text{Precision} = \frac{\text{TP}}{\text{Real positive value}} \quad (7)$$

3 CNN Architectures

The convolutional neural network (CNN or ConvNet) is excellent in image classification among various deep neural networks. Nowadays, we have various pretrained models for image classification. A pre-trained model means the model created by someone to solve a particular problem can be used by other researchers instead of re-inventing everything from scratch.

4 LeNet-5

LeNet was proposed by LeCun et al. [20] in 1998. This architecture was established for the recognition of handwritten digits. Figure 3 shows LeNet-5 architecture. The size of the input image is 32×32 . Besides the input layer, LeNet has seven layers. In this Figure, C_x denotes convolution layers, S_x denotes sub-sampling layers, F_x denotes fully connected layers. The first convolution layer comprises six 5×5 filters convolved with the input picture by taking 1 step to generate an output of 28×28 in size. The following layer is the pooling layer, which employs a sliding window of 2×2 and stride = 2 to provide an output of 14×14 pixels. The following convolution layer employs 16 filters of size 5×5 , padding 0, and stride 1 to provide the output size of 10×10 [21]. The output of the next pooling layer is 5×5 , and it employs stride 2. There are two fully connected layers with 120 and 84 neurons in each. Finally, the last layer has ten neurons that employ the SoftMax activation function to categorize the digits. The number of parameters in LeNet5 is 60,000.

5 AlexNet

AlexNet [22] was proposed by Krizhevsky et al. in 2012. This architecture is developed for ImageNet dataset to classify images. It includes 8 layers, including 3 fully

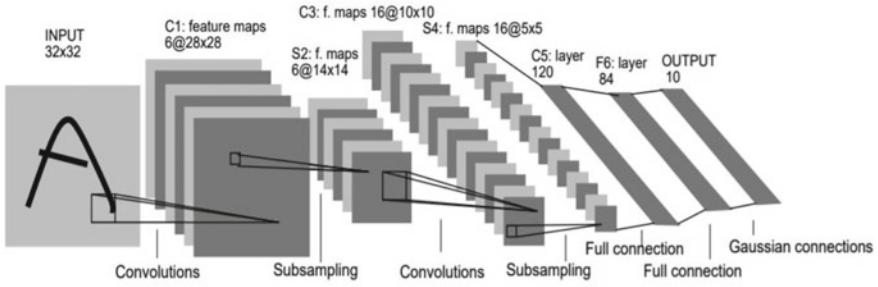


Fig. 3 LeNet-5 architecture [20]

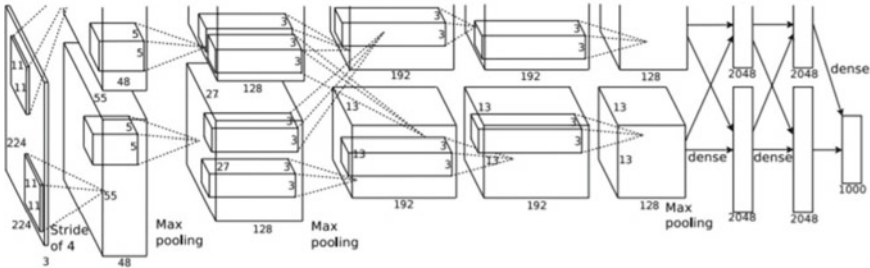


Fig. 4 AlexNet architecture [22]

connected layers and 5 convolutional layers. The SoftMax layer, the last layer of this architecture, has 1000 neurons. The authors also used ReLU to avoid vanishing gradient problem and dropout to reduce the overfitting. This architecture totally has sixty million parameters and 650,000 neurons. Also, it is trained on two GPUs. Figure 4 shows AlexNet architecture.

6 VGGNet

VGGNet [23] was proposed by Simonyan and Zisserman in 2014. It is a deeper network with smaller filters. The total number of layers in VGGNet is 16–19 layers with 3×3 kernels. It comprises 138 million parameters. It uses 3×3 CONV stride 1 with pad 1 and 2×2 max pool with stride 2. 4096 neurons make up the first two fully-connected layers, while the final layer, which does classification, contains 1000 channels. The last layer uses the SoftMax activation function to perform classification.

7 ResNet

ResNet [24] was proposed by He et al. in 2015 with 152 layers. It is an intense network using residual connections. The model performs worse in a plain convolutional neural network due to optimization problems when we stack more layers. ResNet overcomes this. In ResNet, the authors used the concept of skip connection, which means skipping some of the connections and adding the output from the previous layer to the next layer.

8 Experimental Setup and Results

Here, the CIFAR-10 [25] dataset compares classification accuracy between the CNN models. In the CIFAR10 dataset, there are 60,000, 32×32 RGB pictures. The combination of train and test images is 50,000 and 10,000. Ten classes are represented in this dataset: truck, frog, horse, ship, vehicle, bird, cat, deer, and airplane. Figure 5 shows the sample images in the CIFAR10 dataset. The coding is done in Google Colab with GPU using Python 3.7, and the deep learning framework PyTorch is used. The mean squared error (MSE) loss function and Adam optimizer are used in this paper. Table 2 gives the performance of CNN models concerning precision and recall on the CIFAR-10 dataset. The confusion matrix depicted in Figs. 6, 7, 8, and 9 evaluates the performance of the CNN models. From Table 2, we can understand that VGG16 has the highest accuracy (92.39%) in classifying images. Moreover, the two-layer CNN model has a classification accuracy of 60.83%; also, it is lower than other models. We can understand that the classification accuracy will improve by increasing the number of layers. Comparing these four models, VGG16 provides good classification accuracy, precision, and recall value of 92.39, 92.1, and 92.6, respectively. Figure 10 provides a graphical representation of Table 2.

9 Conclusion

This paper performs the classification using CNN with two convolutional layers, VGG16, AlexNet, and ResNet. From the above results, we can observe the VGG16 model provides better accuracy for image classification. Therefore, we can use this image classification technique for object recognition, medical imaging, and face recognition for security purposes. Also, we can improve the accuracy of classification by increasing the number of epochs, training, and testing using more extensive datasets.

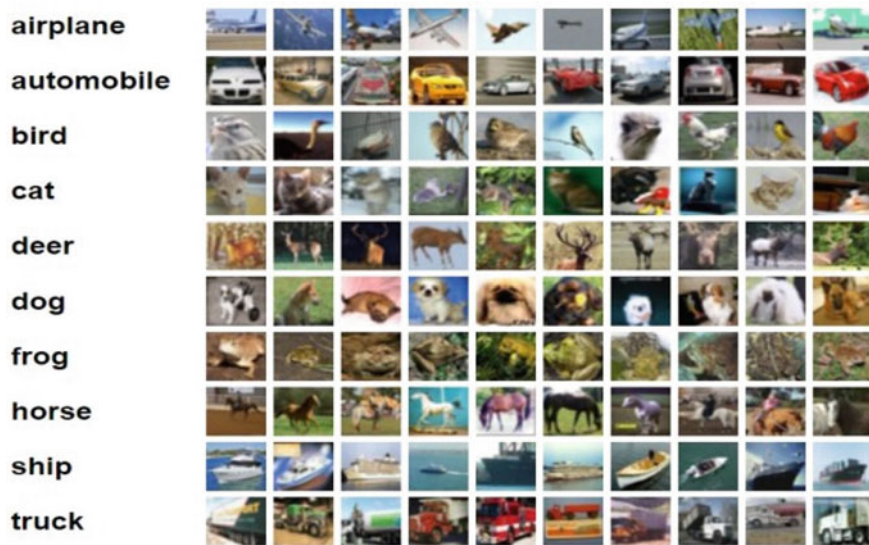


Fig. 5 CIFAR10 dataset [25]

Table 2 Performance measure

CNN model	Accuracy (%)	Precision (%)	Recall (%)
CNN (2 layers)	60.83	60.9	62.7
AlexNet	74.94	74	74.3
ResNet50	79.48	79.6	79.7
VGG16	92.39	92.1	92.6

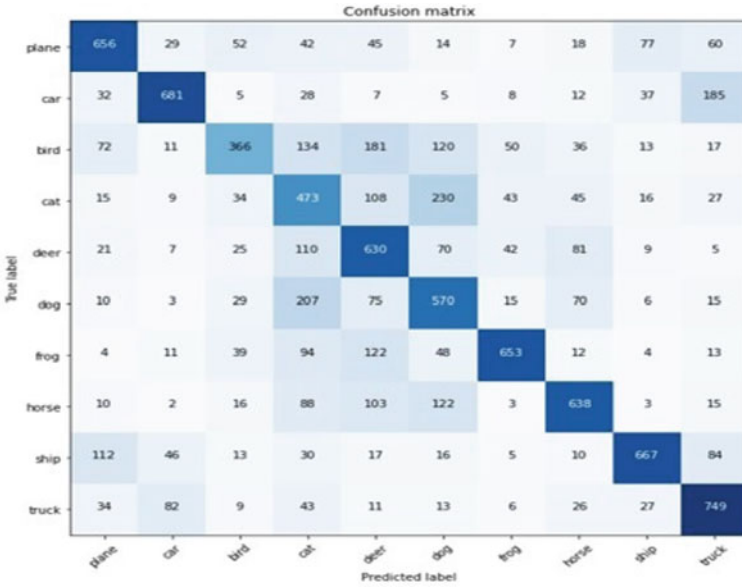


Fig. 6 CNN (2 layers) confusion matrix

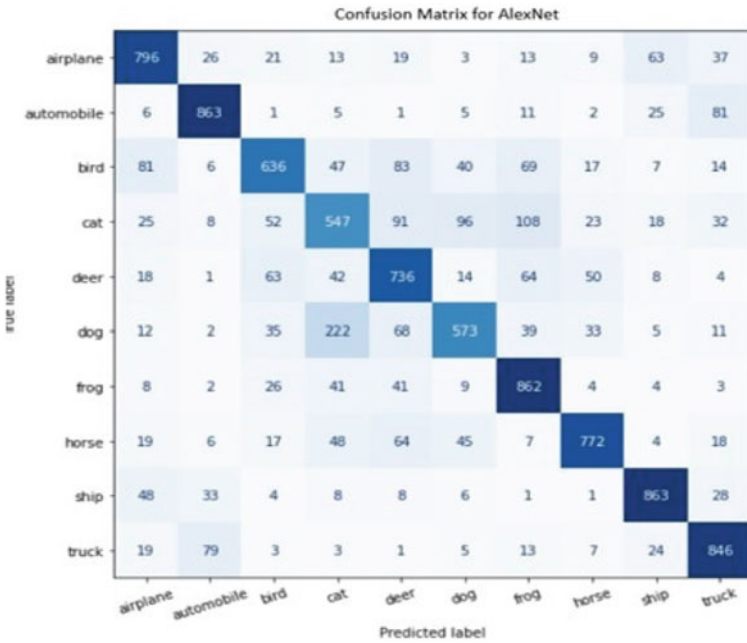


Fig. 7 AlexNet confusion matrix

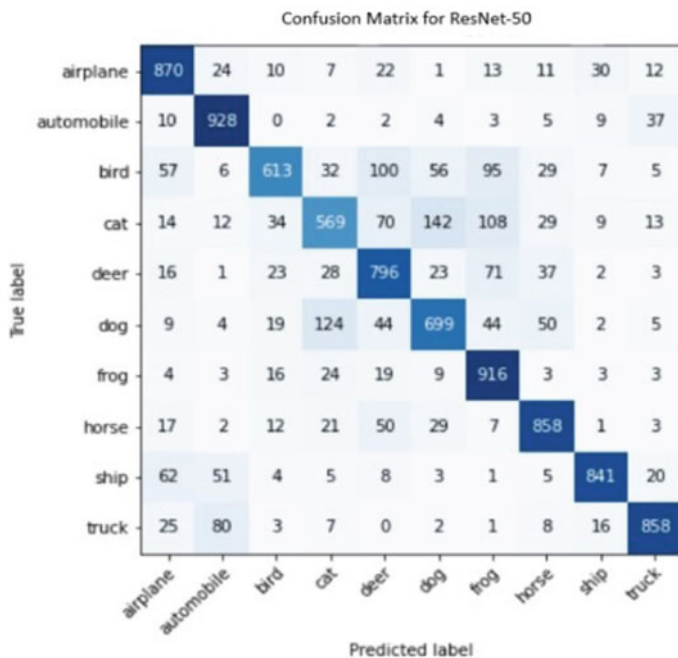


Fig. 8 ResNet confusion matrix

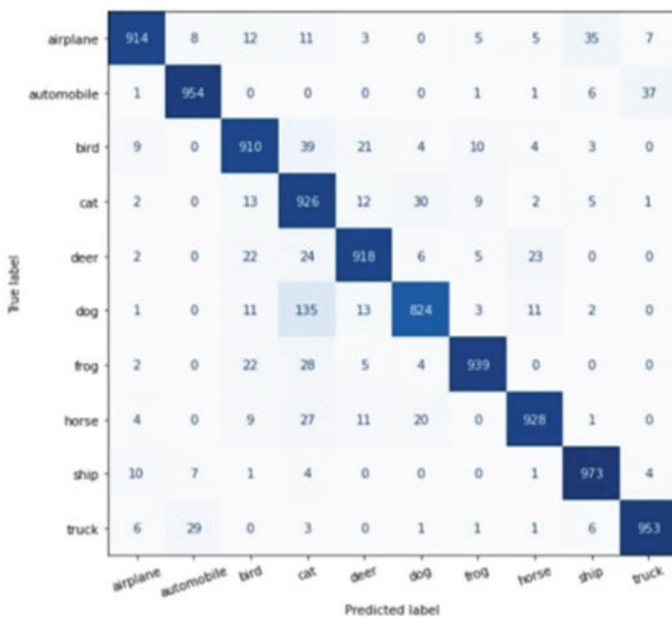


Fig. 9 VGG16 confusion matrix

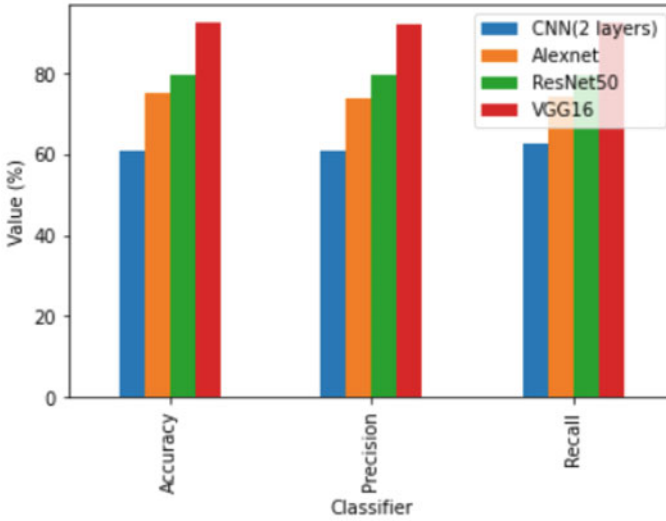


Fig. 10 Performance graph

References

1. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5(4):115–133
2. Hebb DO (2005) *The organization of behavior: a neuropsychological theory*. Psychol Press
3. Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65(6):386
4. Werbos P (1974) *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. PhD dissertation, Harvard University
5. Fukushima K, Miyake S (1982) Neocognitron: a self-organizing neural network model for a mechanism of visual pattern recognition. In: *Competition and cooperation in neural nets*. Springer, Berlin, Heidelberg, pp 267–285
6. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1(4):541–551
7. Weng JJ, Ahuja N, Huang TS (1993) Learning recognition and segmentation of 3-D objects from 2-D images. In: *1993 (4th) international conference on computer vision*. IEEE, pp 121–128
8. Hussain M, Bird JJ, Faria DR (2018) A study on cnn transfer learning for im- age classification. In: *UK Workshop on computational intelligence*. Springer, Cham, pp 191–202
9. Kaur P, Kumar Y, Ahmed S, Alhumam A, Singla R, Ijaz MF (2022) Automatic license plate recognition system for vehicles using a CNN. *CMC—Comput Mater Continua* 71(1):35–50
10. Ge H, Zhu Z, Dai Y, Wang B, Wu X (2022) Facial expression recognition based on deep learning. *Comput Methods Programs Biomed* 106621
11. Feroz M, Sultana M, Hasan M, Sarker A, Chakraborty P, Choudhury T (2022) Object detection and classification from a real-time video using SSD and YOLO models. In: *Computational intelligence in pattern recognition*. Springer, Singapore, pp 37–47
12. Roy AM, Bose R, Bhaduri J (2022) A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. *Neural Comput Appl* 1–27
13. Maweu BM, Dakshit S, Shamsuddin R, Prabhakaran B (2021) CEFES: a CNN explainable framework for ECG signals. *Artif Intell Med* 115:102059

14. Rehman A, Khan MA, Saba T, Mehmood Z, Tariq U, Ayesha N (2021) Micro-scopic brain tumor detection and classification using 3D CNN and feature selection architecture. *Microsc Res Tech* 84(1):133–149
15. Abudhagir US, Anuja K, Patel J (2022) Faster RCNN for face detection on a FACENET model. In: *Advances in mechanical and materials technology*. Springer, Singapore, pp. 283–293
16. Andriyanov NA, Dementiev VE, Tashlinskii AG (2022) Detection of objects in the images: from likelihood relationships towards scalable and efficient neural networks. *Comput Opt* 46(1):139–159
17. Abudhagir US, Anuja K, Patel J (2022) Faster RCNN for face detection on a FACENET model. In: *Advances in mechanical and materials technology*. Springer, Singapore, pp 283–293
18. Peng J, Li J, Shang X (2020) A learning-based method for drug-target interaction prediction based on feature representation learning and deep neural network. *BMC Bioinform* 21(13):1–13
19. Yu Z, Lu J, Jin Y, Yang Y (2021) KenDTI: an ensemble model based on network integration and CNN for drug-target interaction prediction. *IEEE/ACM Trans Comput Biol Bioinform*
20. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
21. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1–9
22. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 25:1097–1105
23. 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)
24. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
25. Alex Krizhevsky's home page. <https://www.cs.toronto.edu/kriz/cifar.html>.

Enhancing Customer Prediction Using Machine Learning with Feature Selection Approaches



R. Siva Subramanian, B. Maheswari, S. Nikkath Bushra, G. Nirmala, and M. Anita

Abstract In each enterprise, customer data is collected in large quantities and the analysis of such data aids in the development of more accurate insights into customer patterns and the identification of high-risk customers in the enterprise. In today's world, machine learning approaches are being used more and more in customer pattern analysis and prediction, as well as in other areas. But in real-time, the consumer dataset often holds a poor quality of data such as irrelevant, corrected, and noisy variables. With these poor quality data, better prediction using the machine learning approaches cannot be witnessed. The removal of correlated variables and the selection of variables with high information content helps to improve the accuracy of customer predictions. With the help of feature selection mechanisms and machine learning techniques, we devise a framework for conducting more accurate customer pattern analysis in this study. The research investigation makes use of UCI's customer dataset and compares machine learning results obtained with and without feature selection. Performing customer analysis helps businesses to improve their decision-making and marketing strategies, which in turn improves customer retention and satisfaction, as well as the enterprise's overall business processes. The exploratory process is done with two distinct feature selection approaches, and feature subset captured is modeled using two different machine learning mechanisms. It is clear from the exploratory results that the use of feature selection allows for the removal

R. Siva Subramanian (✉) · M. Anita
Department of CSE, S. A. Engineering College, Chennai, Tamil Nadu, India
e-mail: sivasubramanian12@yahoo.com

B. Maheswari
Department of CSE, Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India

S. Nikkath Bushra
Department of IT, St. Joseph's Institute of Technology, Chennai, Tamil Nadu, India

G. Nirmala
R.M.D. Engineering College, Kavaraipeetai, Tamil Nadu, India

of correlated variables and the selection of relevant variables, which in turn aids in the improvement of the classifier's performance.

Keywords Customer prediction · Naive Bayes · K-NN · Machine learning · Artificial intelligence

1 Introduction

In each enterprises, the raw information about customer data is collected and stored in enormous volume using different mediums. The accumulated data can be wisely applied to discover interesting patterns about customers, which will be useful in developing better strategies to increase the customer base and increase the profitability of the business. Customer data analysis can be used for a variety of purposes, including: (1) customer segmentation, (2) customer prediction, (3) marketing efficiency, (4) customer retention, (5) profit margin, (6) customer lifetime, and so on [1]. To conduct effective analysis about the customers, machine learning approaches are widely applied. ML techniques helps to find the interesting patterns in data, which in turn further applied to improve the enterprises business. Based upon the nature of the dataset, the use of ML approaches differs. The ML approaches can be widely classified into three types: one is reinforcement learning, second one is unsupervised learning, and the third one is supervised leaning techniques. Supervised learning techniques are applied for labeled datasets and unsupervised learning techniques are applied for unlabeled datasets. The learning algorithms that are applied in supervised learning are: SVM, Naive Bayes, decision tree, neural networks, etc. [2]. The learning algorithms that are applied in unsupervised learning are: K-means, association rules, etc. But in practical, the real-time data generated consists of uncertainties variables and also in high dimensional, and with these data, effective analysis about the customer cannot be captured. To eliminate the uncertainties variables and to minimize the high dimensional, the use of attribute selection (AS) is conducted. AS enables to choose the right set of variables which are relevant to class variables and improve the accuracy of the classifier. AS can be classified into three types: one is filter, embedded, and third one is wrapper approaches [3]. Filter attribute selection uses statistical evaluation approach to assess the influential of the features individually. Based upon the evaluation, the attributes are sorted, and then by using suitable cut-off values, the attribute set is generated. Wrapper AS uses search techniques and induction model to find the best attributes from the whole subset. Best attribute sets are generated in the wrapper approach, but the model is slow with respect to the filter approach. In this study, the filter AS method is used to choose the right set of attributes from the whole set, since this approach is fast. Further, the attribute sets captured are feed into ML model to get better customer analysis. The experimental procedure is carried in two different perspectives: one is ML with attribute selection and the another is ML without attribute selection. The empirical outcome received is projected and compared using different performance metrics parameters. The results

analysis implies the ML with AS approaches performs wisely compared to the ML without AS approach. Since this is to use of relevant variables aid to enhance classifier efficiency. The remaining of the work is followed by literature survey, methodology, experimental results, and conclusion.

2 Literature Survey

In [4], the author addresses the importance of customer retention in the enterprises (financial) and explores the customer analysis which is mostly carried out using the structured data, and the unstructured data analysis is not performed. In this work, the authors apply different ML approaches to apply on unstructured data to conduct churn prediction. In [5], the author explores a study on development WSN and addresses the different attacks in the WSN. Further, the author implies the use of FS aids to minimize the redundant variables and makes to create the accurate model and identify the solution. In this research, the author applies different FS models like CBFS, PCA, LDA, RFE, UFS, and feature set obtained are evaluated using different ML models. The results implies XGBOOST performs superior compare to others. In [6], the author explores the importance of life-threatening arrhythmia disease and addresses the issue in the arrhythmia dataset such as high dimensional, class imbalance, and correlation. To address the issue, the author applies RF-RFE FS technique to get the relevant attribute set. Further, the attribute set is applied with various ML model to conduct the arrhythmia disease analysis. The results imply random forest performs superior compare to others. In [7], the author explores the important study on heart disease prediction and analysis of disease in efficient and timely manner aid to help to save the life. Further, the author addresses the need of FS and applies GCSA FS method to get right set of attributes. In this work, the author applies deep CNN model as ML for classification. The empirical outcome shows GCSA approaches get 94% classification performance compare to other FS approaches. In [8], the author conducts a study on feature selection techniques and applies in facial expression recognition application. In this work, the author uses two different FS model: one is Chi-Square and the other one is ReliefF, and the attributes captured are modeled with four different ML models: SVM, KNN, DT, and RBF. The empirical results captured from both FS are projected and compared. From the result analysis, it shows K-NN achieves better performance compare to others. In [9], the author implies heart contraction prediction is challenging for medical practitioners. In this work, the author applies GWO attribute selection for minimizing the attributes and selecting the relevant attributes. Further, the selected relevant variables set are modeled using the ANN ML model. The empirical results obtained are projected and compared using SVM, DT, K-NN, NB, RF, LR. Results analysis implies proposed approach performs wisely compare to others. In [10], the author performs a study prediction of students performance in educational domain. Further, the author addresses the importance of the feature selection role in improving the performance of the classifier. In this work, the author proposes HVS approach to select the right set of variable

set to model with the ML model. The empirical results obtained are compared using exiting FS approaches. Results reveal the proposed FS model performs better than existing approaches. The proposed approach gets 90% accuracy compare to existing approach.

3 Methodology

This section describes detailed description about the proposed methodology to improve the customer prediction. The overall methodology is shown in Fig. 1.

3.1 Problem Statement

The objective of this research to selected most influential feature which aids to enhance the analysis about the customer prediction. Given the customer dataset which consists $D = \{v_1, v_2, v_n|c_n\}$. The aim is to find best influential variables and remove the inappropriate and correlated features.

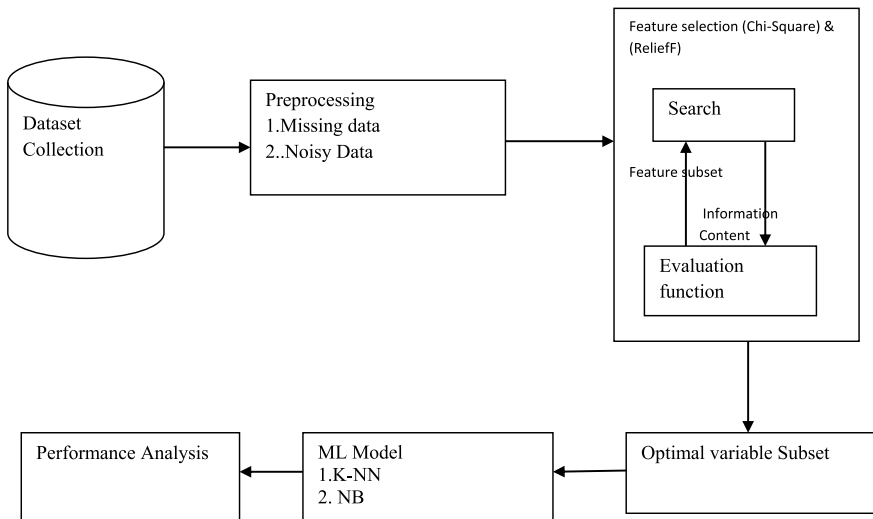


Fig. 1 Overall methodology

3.2 *Algorithm*

1. Given the customer dataset $D = \{v_1, v_2, v_n | c_n\}$.
2. Apply filter FS approach to sort the features
3. Use cut-off value to choose the influential features
4. Model the influential feature using ML classifier.
5. Project the experimental outcome using different metrics.
6. Evaluate the results

3.3 *Data Collection*

The first step of the methodology is collection of related customer for the problem. In today years, the customer data generated is from different customer interaction approach. Further, the customer dataset collected has high possible of holding the correlated and inappropriate variables.

3.4 *Data Preprocessing*

Data preprocessing is a necessary step in ML to convert the raw customer data collected into useful format. The raw data collected is further processed to remove the missing and noisy data. In case of missing data, the fields in the dataset are checked. In case of any missing tuples, the field is filled by using manually or using mean methods. In case of noisy data, ML model cannot be interpreted the data, so the noisy data is eliminated.

3.5 *Feature Selection*

In the third step, the influential features are selected using the FS approaches. Here, the most relevant and most contribution features are selected from the entire dataset to model with ML models. All the features in the customer dataset may not contribute to understand the underlying structure of the data. In such case, FS approaches are carried out to find out the interesting features from the dataset that are highly associated with target class and remove features which are less optimal from the dataset [11]. There are different types of FS approaches in practice, and in this work, filter FS mechanism is conducted. In filter FS approaches, two different FS approaches are considered: one is Chi-Square and another one is ReliefF. These both approaches are applied separately with the dataset and best influential features are selected.

Chi-Square Steps [12]

1. Define the hypothesis
2. Calculate the expected values
3. Calculate $\frac{(O-E)^2}{E}$ for each cell in the table
4. Calculate the test statistic X_2

ReliefF [12]

1. Randomly select an instance
2. Evaluate nearest miss and hits
3. Evaluate the weight
4. Sort the features

The preprocessed dataset is further applied with feature selection to remove correlated, inappropriate variables, and select the influential feature from the entire dataset. This filter FS approach will sort the variables based on their relevance to the class table. Further, to choose the required variables, the use of cut-off mechanism is encouraged. Since using the cut-off value, the range of ranked values are selected. So based upon the cut-off values, the number of variables selected differs. Here, in this step, 50% cut-off is applied to choose the required influential features. Further, the features selected are processed using the NB and K-NN model in the next step.

3.6 ML Classifier

The selected influential features from the above FS step is processed using modeled using ML classifier to perform efficient customer efficient. In this step, two different ML approaches are considered: one is Naive Bayes and the other one is K-NN.

3.6.1 Naive Bayes

NB is efficient and simple model which is based on the Bayes rule. To classify new instance using Naive Bayes [13, 14],

$$\hat{y} = \arg \max_{k \in \{1, \dots, K\}} p(B_k) \prod_{i=1}^n p(a_i | B_k) \quad (1)$$

The most significant presumptions made by NB is that the attributes in the dataset are conditionally independent and equal. The breach of these presumptions makes the model to perform suboptimal prediction about the customer analysis.

3.6.2 K-NN Algorithm

1. Specify the number of neighbors k .
2. Secondly, compute the Euclidean distance between a set of K number of neighbors.
3. Consider the K nearest neighbors as per the evaluated Euclidean distance.
4. Count the number of data points in each category among the k neighbors.
5. The new data points are assigned to the category for which the number of the neighbor is maximum [15].

Further, the exploratory results captured are projected and compared using different performance metrics.

3.7 Performance Metrics

The experimental outcome obtained from the proposed methodology is evaluated and compared using different performance metrics. The parameter considered is given below [16].

$$\text{Accuracy} = \frac{(Tn + Tp)}{(Tp + Tn + Fp + Fn)} \quad (2)$$

$$\text{Sensitivity} = \frac{(Tp)}{(Fn + Tn)} \quad (3)$$

$$\text{Specificity} = \frac{(Tn)}{(Fp + Tn)} \quad (4)$$

$$\text{Precision} = \frac{(Tp)}{(Fd + Tp)} \quad (5)$$

Further, the exploratory results captured from the proposed ML with FS are compared and validated using the approach ML without FS. Next section briefly addresses the experimental results obtained from the two different approaches.

4 Experimental Results

The experimental procedure is performed in two different perspectives. One is using FS mechanism to select the relevant subset to model with ML and another one is using ML without using FS mechanism. The study is conducted using the Australian credit dataset. The exploratory outcome got are compared using good number of parameters. The validated results are shown in Figs. 2, 3, 4, 5, 6 and 7.

Fig. 2 Accuracy comparison of NB and K-NN

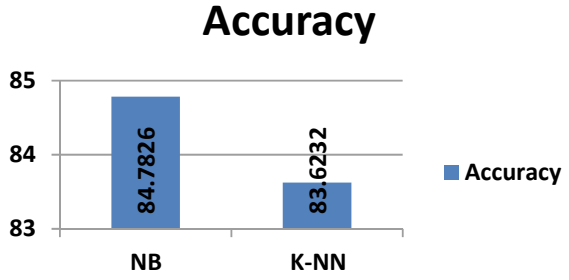


Fig. 3 Sensitivity, specificity, and precision comparison of NB and K-NN

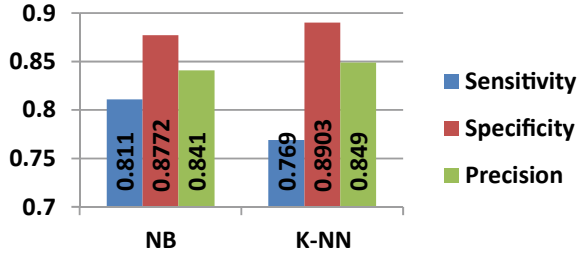


Fig. 4 Accuracy comparison of NB and K-NN

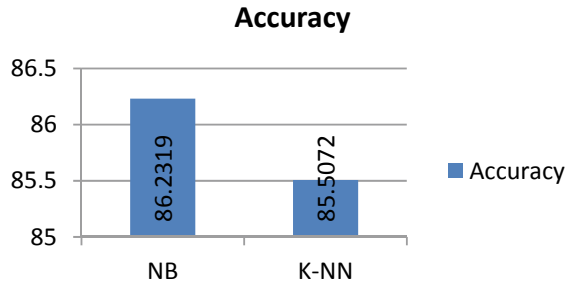


Fig. 5 Sensitivity, specificity, and precision comparison of NB and K-NN

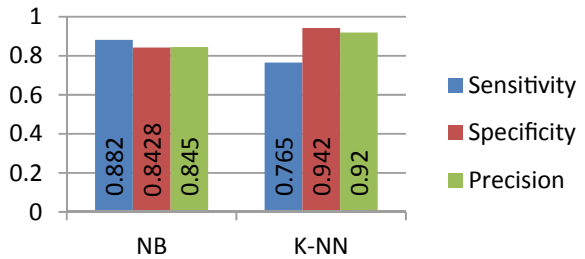


Fig. 6 Accuracy comparison of NB and K-NN

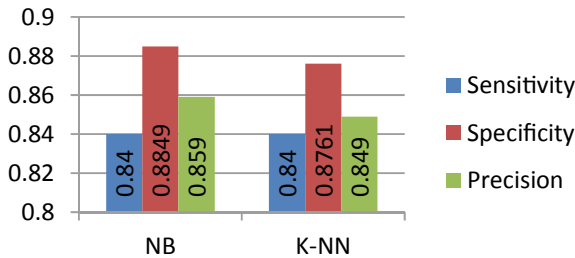
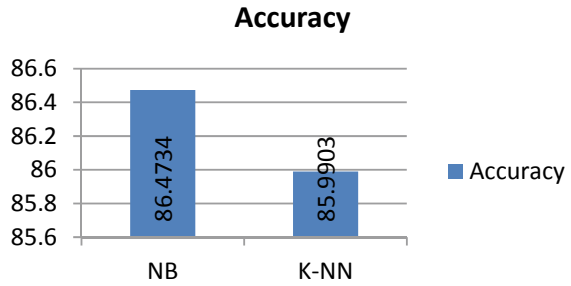


Fig. 7 Sensitivity, specificity, and precision comparison of NB and K-NN

4.1 Experimental Procedure

1. First using the filter FS mechanism, relevant attribute sets are captured. This filter FS approach will sort the variables based on their relevance to the class table. Further, to get high information variables, a cut-off value is considered. In this research, 50% cut-off rate is considered to get right set of variables, and chi-Square and ReliefF approaches are considered.
2. Next feature obtained are modeled with two different ML approaches: one is Naive Bayes and the other one is K-NN.
3. Then without using any FS approaches the dataset is directly modeled with two different ML approaches.
4. The results captured from the two different methodologies are projected and compared using different parameters.
5. From the exploratory result analysis, its shows use of FS approaches aids to get better customer prediction results. Results obtained are represented in Tables 1, 2 and 3.

Table 1 Results of ML without FS approach using four different parameters

S. No.	Model	Accuracy	Sensitivity	Specificity	Precision
1	NB	84.7826	0.811	0.8772	0.841
3	K-NN	83.6232	0.769	0.8903	0.849

Table 2 Result of ML using chi-square filter FS approach

S. No.	Model	Accuracy	Sensitivity	Specificity	Precision
1	NB	86.2319	0.882	0.8428	0.845
3	K-NN	85.5072	0.765	0.942	0.92

Table 3 Result of ML using ReliefF filter FS approach

S. No.	Model	Accuracy	Sensitivity	Specificity	Precision
1	NB	86.4734	0.84	0.8849	0.859
3	K-NN	85.9903	0.84	0.8761	0.849

Table 1 represents the results of ML without using any FS approaches. In this methodology, the customer dataset is directly applied to ML without using any FS techniques. Here, we can see that Naive Bayes model achieves superior results compare to K-NN in terms of accuracy.

Table 2 represents the results of ML using any Chi-Square FS approach. In this methodology, the customer dataset is processed using Chi-Square FS and features selected are applied with ML. Here, we can see that both ML models achieve better results compared to ML without FS approach. Further compared to both models, Naive Bayes model achieves superior results compare to K-NN in terms of accuracy. Removal of unnecessary variables helps to enhance classifier efficiency. Compared to Table 1 results, this methodology performs superior.

Table 3 represents the results of ML using any ReliefF FS approach. In this methodology, the customer dataset is processed using ReliefF FS and features selected are applied with ML. Here, we can see that both ML models achieve better results compare to ML without FS approach. Further compared to both models, Naive Bayes model achieves superior results compared to K-NN in terms of accuracy. Removal of unnecessary variables helps to enhance the classifier efficiency. Compared to Table 1 results, this methodology performs superior.

4.2 Result Discussion

The exploratory results captured from two different methodologies are presented in Tables 1, 2 and 3. Table 1 represents the results of ML without using any FS approaches. In Table 1, results when compared to K-NN, the NB has a higher accuracy of 84.7826. Correspondingly, NB gets high sensitivity of 0.811 compared to K-NN. But in case of precision, we can see K-NN gets high rate compared to NB. Table 2 represents the results of ML using FS approaches. In this methodology, Chi-Square FS is applied. In Table 2, results when compared to K-NN, the NB has a higher accuracy of 86.2319. Correspondingly, NB gets high sensitivity of 0.882 compared to K-NN. But in case of precision, we can see K-NN gets high rate compared to

NB. Table 3 represents the results of ML using FS approaches. In this methodology, ReliefF FS is applied. In Table 3, results when compared to K-NN, the NB has a higher accuracy of 86.4734. Correspondingly, NB and K-NN get same sensitivity values. Likewise, NB gets high precision compared to K-NN. From the exploratory results its clear, the use of feature selection makes to remove correlated variables and select the relevant variables and aids to increase the performance of the classifier.

4.3 Research Findings

1. The customer dataset captured has high possibility of consists of correlated and inappropriate variables.
2. The use of these datasets directly without FS makes to get less suboptimal outcome about the customers.
3. The use of FS aids to get relevant variables and removes the correlated and inappropriate variables from the entire dataset.
4. The use of pertinent variables influences the efficiency of the ML model and also minimizes the computational complexity of the classifier.
5. From the exploratory results its clear, the use of feature selection makes to remove correlated variables and select the relevant variables and aids to increase the performance of the classifier.

5 Conclusion

In recent years, the data about the customer generation is enormous in each enterprise. The use of the customer data in efficient ways helps the enterprises to figure the risk customers and customer patterns with an enterprise. Further, with the aid of customer analysis, more intelligent business plans can be developed to increase customer satisfaction and retention wisely. To analyze these customer data in productive way, the use of ML approaches is encouraged. ML approaches are efficient in handling these high- and large-dimensional datasets and aid in understanding the underlying structure of the data. But in real scenario, the customer dataset collected consists of correlated and inappropriate variables. The use of these attributes without any preprocessing techniques makes to get suboptimal prediction results. To overcome these consequences, the use of FS approach is encouraged. The intent of FS is to get rid of correlated and inappropriate features in the dataset and select the relevant variables to model with the ML model. But problem with filter approach is it only sort the variables accordingly to correlation (target class) and to choose the relevant features use of cut-off value is considered. The influential variables are chosen in accordance with the cut-off values. Under this study, two distinct filter FS aspects are investigated: one is the Chi-Square filter, and the other is the ReliefF filter. Further, the feature set captured by the 2 FS techniques is tested individually to

two different ML classifier: one is Naive Bayes and another one is K-NN. The results obtained from the from above approaches are evaluated and projected using various performance metrics. The findings are presented in Tables 1, 2 and 3. Table 1 presents the results of ML without any FS approaches. Table 2 represents the results of ML using any Chi-Square FS approach and Table 3 represents the results of ML using any ReliefF FS approach. From the exploratory result analysis, it clearly shows NB and K-NN achieve higher accuracy using feature selection compared to without feature selection. Results analysis imply the use of relevant variables aid to improve classifier efficiency and also minimize the computational complexity of the classifier. In future works, use of different variable selection is encouraged.

References

1. Siva Subramanian R, Prabha D (2017) A survey on customer relationship management. In: 4th International conference on advanced computing and communication systems (ICACCS). Coimbatore, pp 1–5. <https://doi.org/10.1109/ICACCS.2017.8014601>. Electronic ISBN: 978-1-5090-4559-4
2. Subramanian S, Prabha D (2018) Prediction of customer behaviour analysis using classification algorithms. AIP Conf Proc 1952:020098. <https://doi.org/10.1063/1.5032060>. ISBN: 978-0-7354-1647-5
3. Siva Subramanian R, Prabha D (2020) Optimizing Naive Bayes probability estimation in customer analysis using hybrid variable selection. In: 3rd International conference on computer networks and inventive communication technologies (ICCNCT). Springer Publication
4. Vo NNY, Liu S, Li X, Xu G (2021) Leveraging unstructured call log data for customer churn prediction. *Knowl-Based Syst* 212:106586
5. Panda S, Khamparia A, Gupta D (2021) Feature selection and comparison of classification algorithms for wireless sensor networks. *J Ambient Intell Human Comput*
6. Wang T, Chen P, Bao T, Li J, Yu X (2021) Arrhythmia classification algorithm based on SMOTE and feature selection. *Int J Perform Anal Sport* 17(3)
7. Senthil Murugesan N, Muthukumar V, Murusegan R, Bindu R, Meram M, Prathik (2021) Innovative feature selection and classification model for heart disease prediction. *J Reliable Intell Environ*
8. Mahmood MR (2021) Two feature selection methods comparison chi-square and Relief-F for facial expression recognition. *J Phys* 1804
9. Le MT, Vo MT, Pham NT, Dao SVT (2021) Predicting heart failure using a wrapper-based feature selection. *Indonesian J Electr Eng Comput Sci* 21(3):1530–1539
10. Zaffar M, Hashmani MA, Habib R, Quraishi KS, Irfan M, Alphtani S, Hamdi M (2021) A hybrid feature selection framework for predicting students performance. *Comput Mater Continua* 70(1):1893–1920
11. Balakrishnan S, Prabha D, Karpagam M, Siva Subramanian R (2019) Performance evaluation of Naive Bayes classifier with and without filter based feature selection. *Int J Innov Technol Explor Eng* 2278(3075):2154–2158
12. Siva Subramanian R, Prabha D, Aswini J, Maheswari B (2022) Evaluation of different variable selection approaches with naive Bayes to improve the customer behavior prediction. In: Smys S, Balas VE, Palanisamy R (eds) *Inventive computation and information technologies. Lecture notes in networks and systems*, vol 336. Springer, Singapore. https://doi.org/10.1007/978-981-16-6723-7_14
13. Siva Subramanian R, Prabha D (2022) Ensemble variable selection for naive Bayes to improve customer behaviour analysis. *Comput Syst Sci Eng* 41(1):339–355

14. Shobana G, Bushra SN (2021) Prediction of cardiovascular disease using multiple machine learning platforms. In: 2021 international conference on innovative computing, intelligent communication and smart electrical systems (ICSES), pp 1–7. <https://doi.org/10.1109/ICSES52305.2021.9633797>
15. Siva Subramanian R, Prabha D, Maheswari B, Aswini J (2022) Wrapper-Naive Bayes approach to perform efficient customer behavior prediction. In: Raj JS, Kamel K, Lafata P (eds) Innovative data communication technologies and application. Lecture notes on data engineering and communications technologies, vol 96. Springer, Singapore. https://doi.org/10.1007/978-981-16-7167-8_2
16. Shobana G, Bushra SN (2020) Classification of Myopia in children using machine learning models with tree based feature selection. In: 2020 4th international conference on electronics, communication and aerospace technology (ICECA), pp 1599–1605. <https://doi.org/10.1109/ICECA49313.2020.9297623>

Domain Name System Resolution System with Hyperledger Fabric Blockchain



Khoi Tuan Huynh Nguyen, Hy Khang Vuong Nguyen, Dung Ngoc Vo Lam, Tuan Anh Dao, Nhan Trong Van Pham, Huong Hoang Luong, and Hai Thanh Nguyen

Abstract There is a strong demand for a trusted domain name resolution mechanism because of endless cyber-attacks. However, the existing collaborative Domain Name System (DNS) security schemes have low credibility and an imperfect validating method. Therefore, we propose a multi-DNS resolution model, namely, HFDNS, which can improve the credibility of DNS resolution results by establishing a complete chain of trust by combining an automated DNS system with Hyperledger Fabric Blockchain. Our team developed a DNS recursive server cluster in which nodes jointly resolve domains. Therefore, hackers must compromise nodes simultaneously to poison our DNS system successfully. All verified records are then saved in a secure place, which, in our project, is a Hyperledger Fabric network. Our system can detect and discard malicious DNS packets from this validation scheme. Hyperledger Fabric Blockchain is a carrier of the peer-to-peer network to reduce the impact of illegal access and complicity tampering on the DNS credibility. Hyperledger Fabric Blockchain has four characteristics: permission network, confidential transaction, non-crypto currency, and programmable. Furthermore, the DNS records stored in the Hyperledger network are immutable, thus maintaining their validity. This system is expected to be used by enterprise or service provider networks. The experiment shows that our system can consistently resolve users' queries within 192 ms for uncached records and 14 ms for cached records. Furthermore, our validation algorithm successfully returned a valid response for 84% of the total queries.

Keywords Domain name service · Hyperledger fabric · Blockchain · DNS cache poisoning

K. T. H. Nguyen · H. K. V. Nguyen · D. N. V. Lam · T. A. Dao · N. T. V. Pham · H. H. Luong
FPT University, Can Tho, Viet Nam
e-mail: khoiht@ieee.org

D. N. V. Lam
e-mail: dunglvn299@gmail.com

N. T. V. Pham
e-mail: nhanpvt.ct@gmail.com

H. T. Nguyen (✉)
Can Tho University, Can Tho, Viet Nam
e-mail: nthai.cit@ctu.edu.vn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_5

1 Introduction

DNS has always been a critical component of the Internet. Due to its popularity, DNS is an attractive target for hackers. Therefore, the security of the DNS system is of significant importance. Any successful attempts at compromising the DNS servers will put customers, and enterprises' networks at devastating risks [1]. There have been various novel attacks against DNS, such as cache poisoning, Distributed Denial of Service (DDoS), and amplification. Unfortunately, the current defending techniques like The Domain Name System Security Extensions (DNSSEC) or DNS over Hypertext Transfer Protocol Secure/Transport Layer Security (DNS over HTTPS/TLS) have yet to be widely adopted [2]. DNSSEC still has many weaknesses, such as the availability of DNSSEC resolution and validation service is not efficient. An attacker can take advantage of this vulnerability to perform a Denial of Service (DoS) attack [3].

Furthermore, the DNSSEC additional encryption process could cause the performance of the DNS system to decrease because it included many extensive records containing signatures or cryptographic keys [4]. In addition, the implementation of DNSSEC will affect the DNS system that is usually working, resulting in loss of availability, so it is not possible to immediately deploy DNSSEC [5, 6]. Therefore, there is a growing demand for viable solutions that can enhance the security of the DNS infrastructure.

After researching, we developed a DNS architecture based on a permission-based Blockchain. Our system is divided into two major components: the DNS resolvers and the Hyperledger Fabric. In our system, we use the multi-DNS resolution mechanism. Several precursors will resolve a domain to select the most returned Internet Protocol (IP) addresses. Our implementation minimizes the need for complex encryption techniques like DNSSEC. We also use Hyperledger Fabric Blockchain as our cache storage to take advantage of a permission-based Blockchain's immutable and decentralized characteristic [7] to mitigate cache vulnerabilities. With its consensus algorithm ensuring conformity in ledgers and authentication certificates encrypting all the transmission, Hyperledger Fabric could create a truth-worthy system. On top of that, not only is Hyperledger Fabric a non-concurrency Blockchain to simplify and deduce the system's computing, but its ledger also has two states (world state and blockchain state) to improve the Blockchain performance¹ significantly. In addition, Fabric is more advanced and convenient than other traditional Blockchain in terms of storage.

Our work is expected to make the following contributions:

- We introduce a novel DNS resolution technique that thwarts DNS poisoning attempts. Specifically, we illustrate how a group of DNS can cooperate to resolve one single domain, thereby minimizing the chance of a prosperous poison (Multi-DNS). We also explain how our resolution technique differs from previous studies to decrease resolution latency.

¹ <https://hyperledger-fabric.readthedocs.io/en/release-2.2/>.

- We introduce a Blockchain system for DNS cache storage to improve monitoring capability.
- We attempt to reduce the latency of Blockchain by Hyperledger Fabric.

The rest of the paper is organized as follows (Sect. 1). Then, we will discuss several related studies on DNS attacks and Blockchain (Sect. 2). After that, in section three, we will elaborate on our architecture and algorithms (Sect. 3). Subsequently, our experiments for the system will be explained (Sect. 4). Finally, in conclusion, we will summarize our studies' key features and development plan (Sect. 5).

2 Related Work

There have been numerous studies on Blockchain and DNS cache poisoning detection/prevention. This section will introduce them and explain how our work will differ from similar implementations. We will first show some studies on Blockchain applications in security before moving on to Hyperledger Fabric. Finally, three works on using multi-DNS resolution in preventing DNS cache poisoning will be discussed in-depth, as they are the inspiration for this paper.

2.1 *Blockchain in Trusted Sharing Mode*

Many cyber breaches have made some heavy impacts on DNS security. Blockchain technology has been applied to DNS services to tackle these problems. It required registration permission to access data in blockchain [8]. The read and write permissions are restricted to the members, and their rights are defined by administrators according to the blockchain rules [9]. It means the data sharing between users or organizations is harder to get stolen, ensuring data quality and maintaining its privacy and integrity. With this feature, a DNS-blockchain can prevent threats such as phishing, a man-in-the-middle attack (MITM), or denial of service [10]. There are some studies about using Blockchain to secure data protection as a trusted server [11], infrastructure for smart cities [12] and SPDS: A Secure and Auditable Private Data Sharing Scheme for Smart Grid Based on Blockchain [13]. These show the possibility and efficiency of using Blockchain to prevent data breaches and ensure data consistency. For our research, we use Hyperledger Fabric Blockchain for its endorsement mechanism [14] in data protection. Also, its performance and scalability would help the operation run smoothly.

2.2 *Hyperledger Fabric—Permissioned Blockchain Technology*

A permissioned blockchain secures interactions between organizations with a shared aim but who may not completely trust one another. A permissioned blockchain can utilize more typical Crash Fault Tolerant (CFT) [15] or Practical Byzantine Fault Tolerance (PBFT) [16] consensus methods that do not require costly mining by depending on the identities of the participants. Fabric is the first truly extensible blockchain system for running distributed applications [17]. Fabric implements the permissioned model by utilizing a portable idea of membership that can be coupled with industry-standard identity management systems. Fabric proposes an altogether new blockchain design to accommodate such flexibility and a revamped approach to dealing with nondeterminism, resource exhaustion, and performance threats. Several research articles examining and evaluating the performance capabilities of Hyperledger Fabric have been published. Fabric was recently scaled to 20,000 transactions per second [18].

2.3 *Preventing DNS Cache Poisoning with Multi-DNS*

This subsection will specifically discuss the studies that implemented multi-DNS architecture like ours. In their research, Wang et al. [19] leverage a range of security solutions to build a DNS resolving system resistant to cache poisoning attacks without relying on complex cryptographic measures like DNSSEC. Multiple DNS recursive servers will handle the same request from a client (multi-DNS implementation). After getting the responses from the authoritative servers, they will select the best records based on the majority-based mechanism. DoX [20] is another notable research on cooperating multiple recursive servers to select the best answers for a single query. In their systems, similar servers will be grouped into one channel (verification channel). Members within the same channel will validate the others' new records. All validated records will be stored in Verification Cache (vCache). This cache is separated from the standard cache, and only stores validated records. Another critical characteristic of vCache is that the stored records do not have a Time-to-Live (TTL) value, which means they are only updated or deleted after validation or memory reaches its limit. The recursive server will perform a traditional lookup when a client requests a resolution. After that, it will compare with the records in the vCache. The server will request other peers to validate if the vCache does not have conflicting records. Should servers have the same results, they will update their vCache.

Yu et al. [10] proposed the architecture of DNSTSM to enhance the robustness of the DNS cache. DNSTSM is a security model which combines consortium blockchain with a stochastic distributed decentralized storage mechanism. DNSTSM system was compared to two secure DNS servers, DependDNS and HARD-DNS, in availability,

efficiency, and security. In the availability test, Zhong Yu launched many cache poisoning attacks and birthday attacks on these servers. DNSTSM provides several significant characteristics which can defend against cache poisoning and internal attack. With the security of blockchain databases and multi-recursive server parsing, the system ensures the credibility of the data. While using the Member Service System authenticated by Public Key Infrastructure (PKI) framework and Certificate Authority (CA) sign issuance, DNSTSM secures the legitimacy of certificates to manage members in the consortium. The credibility-based incentive servers stimulate the members to maintain their server security performance and the DNSTSM domain name resolution cycle.

Our work differs from the above studies in specific ways. First, as for the resolution process, the DNS gateway will not attempt to resolve the DNS query from the customers in our architecture. Instead, it will automatically forward the query to the Hyperledger Fabric or resolvers if no validated responses are available. This mechanism is expected to reduce the resolution latency. Another difference is how we handle DNS TTL, which will be discussed in the following section. Furthermore, we also implement Elastic to monitor and secure the network.

3 System Architecture

In this section, we will explain in detail the components of our system, along with the algorithms and workflows.

3.1 Overall

There are two prominent roles in our resolving infrastructure: the DNS gateways and the DNS resolvers. The DNS gateways will listen for DNS requests from clients and forward them to Blockchain, or, should the Blockchain fails to resolve, resolvers. The DNS Gateways must execute a running program to contact the controller's resolvers. The controller program is a crucial component of the infrastructure. It listens for requests from the DNS service and prompts the resolvers to resolve. After receiving responses from the resolvers, it will select the best record based on a majority-based mechanism and send it back to the DNS service. The DNS resolvers are conventional DNS recursive servers. The resolvers' only function is to resolve records sent from the controller. After selecting the records to return to users, the controller program will update the blockchain. Our Blockchain is based on Hyperledger Fabric. It is used to store validated DNS records. Due to its complexity, there will be a flowing section dedicated solely to Blockchain. The architecture is illustrated in Fig. 1.

Our mechanism reduces the chance of successful DNS spoofing in several aspects. First, hackers must spoof multiple resolvers to poison a DNS resolution simultaneously. It can be challenging in practice, as spoofing one resolver, hackers' packets

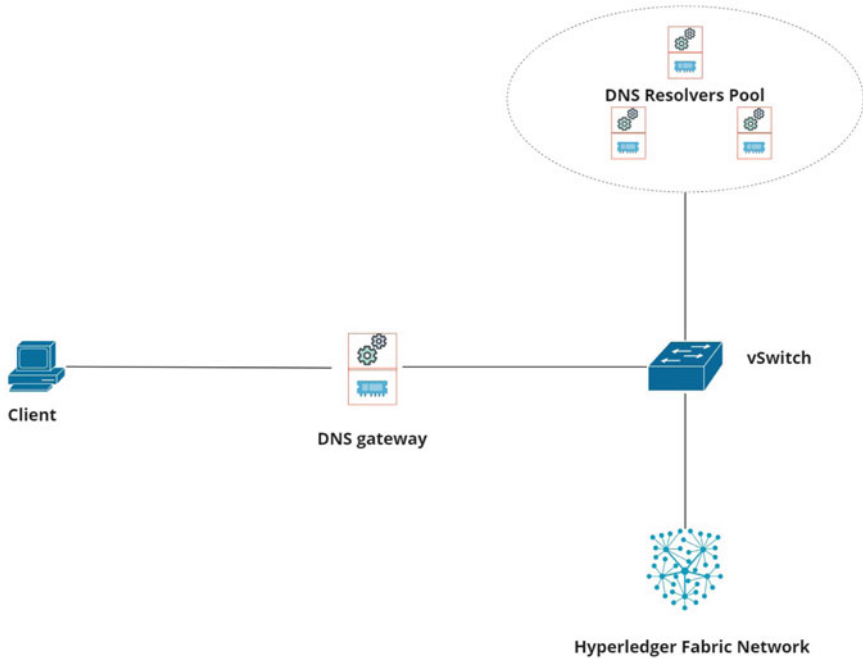


Fig. 1 Overall architecture

must have the same transaction ID and the source port of the original DNS request. Furthermore, they must guess which resolvers will be responsible for resolving their targeted domain.

3.2 DNS

This section will further explain the principles behind our resolution mechanism.

Request Resolution The process in which DNS gateway will request data from Blockchain or even the resolvers can be illustrated in Fig. 2. To be specific, the DNS gateway will first check if Blockchain has valid records for a domain. If yes, the gateway will forward the response to the client. Otherwise, the gateway will ask the controller program to ask the resolvers. After receiving responses from the resolvers, the controller will select the best one to return to the DNS service. The client will receive a DNS resolution error if the controller cannot choose a suitable record. The algorithm behind this selection will be discussed later.

Resolver Scheduling The controller will only send queries to a subset of the resolver pool. The controller application will randomly choose an odd number of resolvers to choose which resolvers to connect every 30 s. This randomization hides the actual

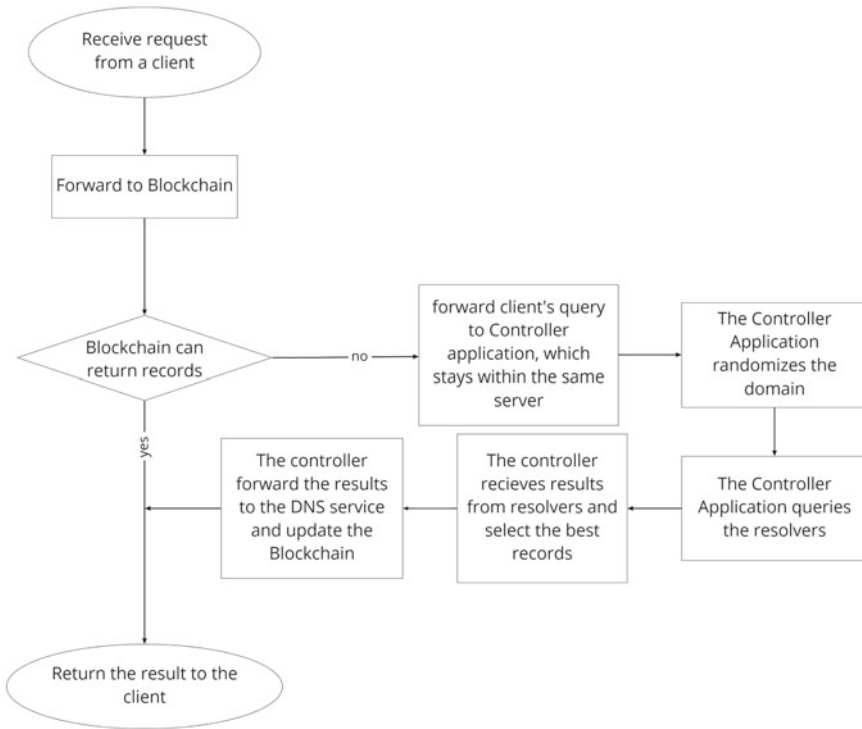


Fig. 2 Resolution workflow

servers that will connect to the Authoritative Servers. Therefore, to successfully spoof our resolution system, hackers either have to choose the scheduled resolvers or spoof all of the resolvers simultaneously, which is resource-consuming.

Record Selection The selection mechanism is majority-based, meaning the record returned by most resolvers will be selected. First, our program will calculate a majority number. Then, after prompting resolvers to resolve a domain, the controller will receive a set of IP addresses from each resolver. After that, all of the sets will be merged into one list. After that, the controller will pick the items whose frequency is higher than or equal to the majority number. The majority number can be calculated as $\frac{\text{Number Of Resolvers} + 1}{2}$.

Blockchain Update If the controller selects the best record, it will send the record to Blockchain after responding to the DNS service. The Blockchain system will receive and store the record for later retrieval.

Time to Live Calculation When developing this system, one major problem was calculating a TTL for a record. As multiple DNS resolvers attempt to resolve a single domain, there might be differences in TTL value. For the time being, our TTL selection is as follows. First, all TTL values from the scheduled resolvers that returned the

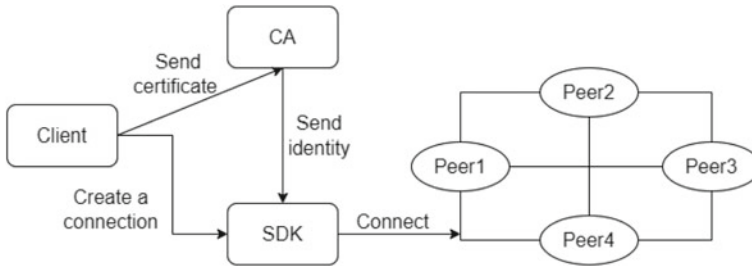


Fig. 3 Hyperledger fabric CA

chosen record will be gathered. Then the lowest value will be selected. The validation time and the TTL will be saved when updating Blockchain with the validated DNS record. When the DNS Gateway receives a record from the Blockchain, it will check the current time and determine if the TTL has been exceeded.

3.3 Blockchain Network

This section details the architecture of HFDNS in the Blockchain network aspect. Overall, the Blockchain network has five major components that are illustrated in Fig. 3. The primary responsibilities are storing and giving back domain name data when Recursive Servers make requests.

Authentication in Blockchain The Certificate Authority service is implemented on the Blockchain network due to the security enhancements of HFDNS. In the initial period, the root administrator creates accounts for Recursive Servers, each having its profile that determines the privileges. When a client tries to establish communication, the Certificate Authority service identifies the peer by username and password. The server responds to the connecting peer with an accepted signal with an attached certificate which initializes encrypted communication between the Recursive Server and the Blockchain network. The server can now call Blockchain Application Programming Interface (API) through chain code. The Recursive Server’s commands pass the verifying policies, and chain code will initiate new transactions, or commands are ignored.

3.4 Transaction in Blockchain

The Hyperledger Fabric Transaction Flow is shown in Fig. 4.

1. Transaction creation: The peer creates a connection within a client application connected to SDK.

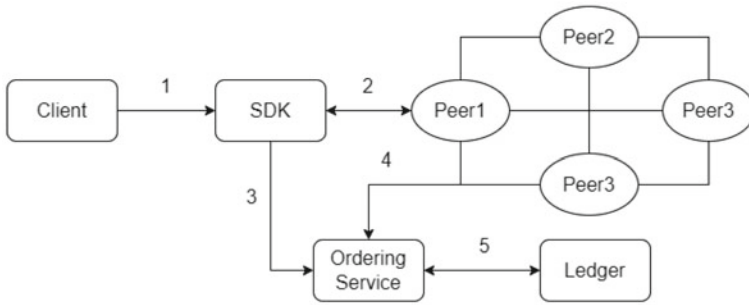


Fig. 4 Hyperledger fabric transaction

2. Transaction Endorsement: The SDK submits the transaction to the endorsing peers. Each endorsing peer executes the transaction and returns the response to the client.
3. Ordering peers submission: The transaction is subsequently sent to the ordering service, which utilizes consensus to order the transaction into a block inside the ledger once it has been endorsed and returned.
4. Transaction commitment: The saved transaction block is transmitted to all peers as part of the channel after requested.
5. Ledger Submission: The block is recorded into the ledger after the transaction(s) sent back to the peers have been confirmed and finished.

Storing and Querying Domain Name Data After Endorsers validate a transaction, it will be sent to the Ordering Service and put into a block. This block is recorded into the ledger. All domain name data will be stored. A ledger in Hyperledger Fabric contains two separated but related parts — a world state and a blockchain (Fig. 5). A world state is a database that keeps track of the current values of a set of ledger states. Instead of calculating the current values by traversing the complete transaction log, the world state allows a program to access it directly. By default, ledger states are expressed as key-value pairs, and Hyperledger Fabric offers flexibility. The world state can frequently change, as states can be created, updated, and deleted. Meanwhile, a blockchain is a transaction log that records all the changes in the current world state. Once the data is written in Blockchain, it is immutable.

4 Evaluation

This section will include our experiment with the system. Specifically, this section will describe how the system was deployed for testing.

Linode was the cloud provider for this system. One node was deployed as a DNS gateway and four as the resolvers. Moreover, the hardware specification was 1GB

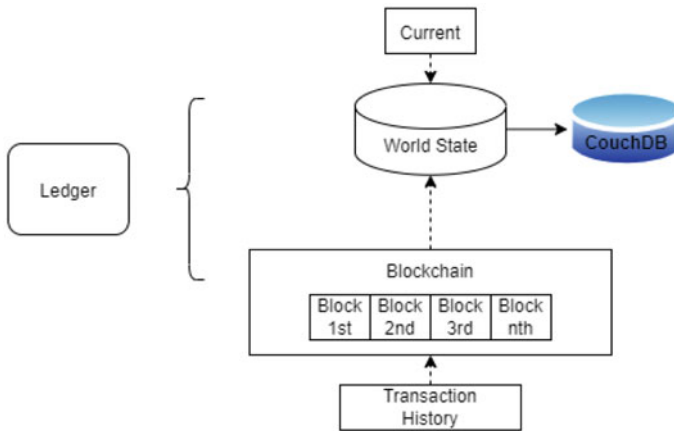


Fig. 5 Ledger components

RAM, 1 CPU, and 25GB of storage. In addition, one node with 4GB of RAM and 2 CPUs for Blockchain was also created. Besides the core servers, Elastic services were implemented to track the performance and secure the network.

Two tests were made for experimental purposes, each lasting for 30 min. First, all resolution logs were pushed to Elastic for visualization. We used a custom-made Python script to flood the DNS gateway with queries in each test. The first test was designated for monitoring the performance of uncached resolution (The Blockchain system had no validated records). The remaining one was used to track the latency of the DNS gateway in resolving cached records.

4.1 Uncached Records

In this test, 2318 queries were sent, and above 99% of them were uncached records. The specific proportion is shown in Fig. 6.

About the uncached latency, the figure fluctuated between 127 and 253 ms, with an average of 193 ms (Fig. 7). Most of the time, the latency stabilized at around 160–190 ms. There were only a few exceptions when the number topped at over 250 ms. However, all of the surges were followed by a profound decrease. We can see that the resolution speed is relatively consistent in the chart.

The latency of the validation algorithm was also tracked. The average latency was 136 ms (Fig. 8). The data pattern in this chart was similar to Fig. 7. The only difference was that dramatic changes in the numbers were rare. In other words, the performance of the validation algorithm was predictable at around 136 ms, as previously mentioned. From this analysis, it can be concluded that the sudden increase in uncached resolution latency may not have been because of the validation process.

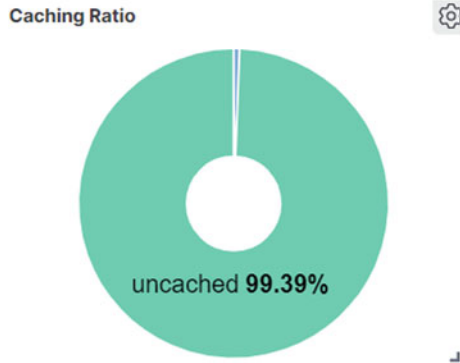


Fig. 6 Uncached resolution ratio

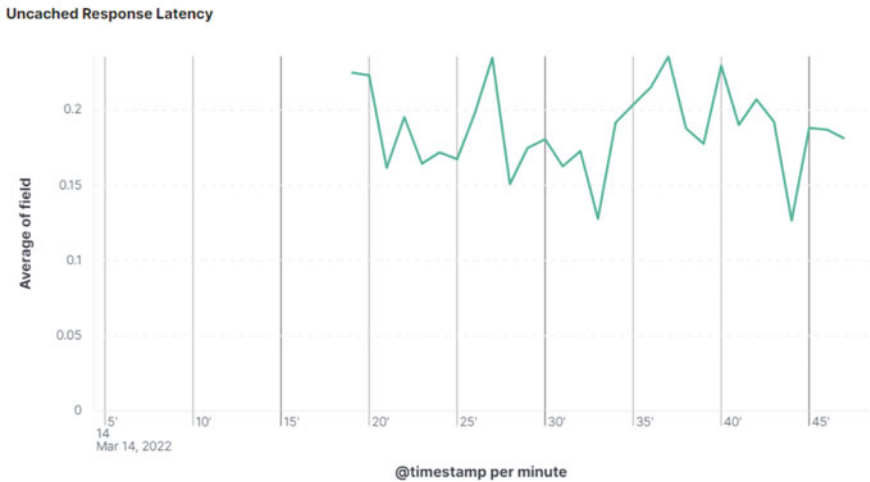


Fig. 7 Uncached resolution latency

Occasionally, the controller cannot determine the best record for the user. This may be attributed to conflicts in responses from resolvers or some errors. To detect any problems, we also monitored the result of the validation. 84.05% of all validation attempts were successful (Fig. 9).

4.2 *Cached Records*

After testing uncached resolution latency, we started testing with cached records. Within 30 min, we sent 44,274 queries. The average latency of this test was around 14 ms (Fig. 10). Contrary to the uncached tests, this test showed better stabilization.

DNS Validation Speed

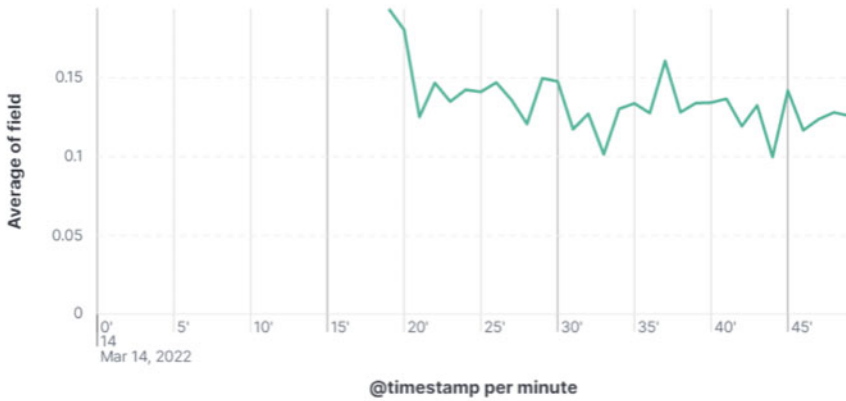
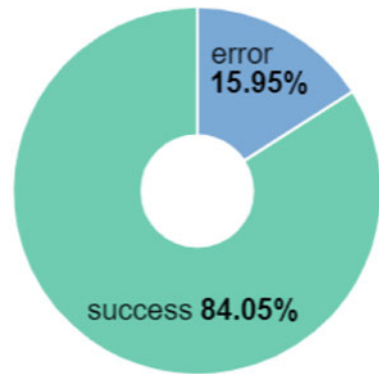


Fig. 8 DNS validation latency

Fig. 9 Validation result ratio

Validation Result



Based on the visualized figure, we can conclude that our system can log records, and the average latency suffered by clients will reduce considerably. However, this also indicates that the validation process still has room for improvement, as it is ten times more time-consuming than the resolution process with cached records.

From the above experiments, we can see that our validation algorithm consumed the most latency. However, the average latency is relatively acceptable for typical use cases without having to validate. It indicates that the Blockchain network has been set up efficiently. As for the validation, the latency is relatively consistent at around 192 ms. Therefore, we are actively modifying our codes for better speed.

Cached Response Latency

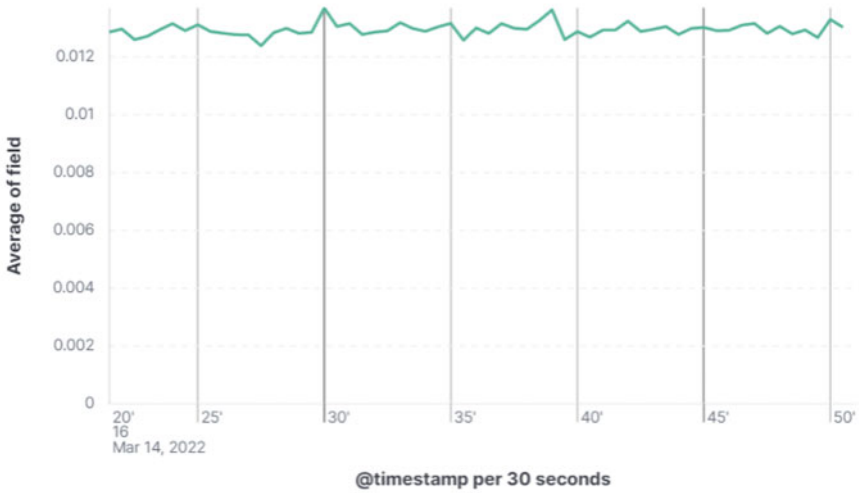


Fig. 10 Cached resolution latency

5 Conclusion

This study presented HFDNS, a DNS resolution system based on Hyperledger Fabric. It is expected that Hackers/crackers face more challenges in poisoning the cache, thanks to our DNS resolution mechanism. Furthermore, with Blockchain technology, we can monitor the storage of validated records, thereby detecting any sophisticated intrusions. Another notable benefit of our architecture is that there is no need for a complex cryptographic solution like DNSSEC. Therefore, we hope enterprises can use our model and deploy a safe DNS resolution system. In the future, we will continue minimizing the speed of our resolution algorithm. Also, we further search for any vulnerabilities in the validation process and implement security patches.

References

1. Lioy A, Maino F, Marian M, Mazzocchi D (2000) DNS security. In: Proceedings of the TER-ENA networking conference
2. Lambert KT, Souleymane O, Tiemoman K, Brice A, Pierre T (2008) Deployment of DNSSEC: problems and outlines. In: 2008 second international conference on research challenges in information science. IEEE, pp 343–348
3. Wang Y, Yun X, Xiong G, Li Z, Yao Y (2012) Availability analysis of DNSSEC resolution and validation service. In: 7th International conference on communications and networking in China. IEEE, pp 35–42

4. Migault D, Girard C, Laurent M (2010) A performance view on DNSSEC migration. In: 2010 International conference on network and service management. IEEE, pp 469–474
5. Atkins D, Austein R (2004) Threat analysis of the domain name system (DNS) RFC 3833. Tech Rep
6. Herzberg A, Shulman H (2013) DNSSEC: security and availability challenges. In: 2013 IEEE conference on communications and network security (CNS). IEEE, pp 365–366
7. Zhang R, Xue R, Liu L (2019) Security and privacy on blockchain. *ACM Comput Surv (CSUR)* 52(3):1–34
8. Feng Q, He D, Zeadally S, Khan MK, Kumar N (2019) A survey on privacy protection in blockchain system. *J Network Comput Appl* 126:45–58
9. Puthal D, Malik N, Mohanty SP, Kougianos E, Das G (2018) Everything you wanted to know about the blockchain: its promise, components, processes, and problems. *IEEE Consum Electron Mag* 7(4):6–14
10. Yu Z, Xue D, Fan J, Guo C (2020) DNSTSM: DNS cache resources trusted sharing model based on consortium blockchain. *IEEE Access* 8:13640–13650
11. Lin J, Shen Z, Miao C, Liu S (2017) Using blockchain to build trusted Lorawan sharing server. *Int J Crowd Sci*
12. Fu Y, Zhu J (2021) Trusted data infrastructure for smart cities: a blockchain perspective. *Build Res Inf* 49(1):21–37
13. Wang Y, Su Z, Zhang N, Chen J, Sun X, Ye Z, Zhou Z (2020) SPDS: a secure and auditable private data sharing scheme for smart grid based on blockchain. *IEEE Trans Ind Inform* 17(11):7688–7699
14. Androulaki E, De Caro A, Neugschwandtner M, Sorniotti A (2019) Endorsement in hyperledger fabric. In: 2019 IEEE international conference on blockchain (blockchain). IEEE, pp 510–519
15. Li W, Meese C, Nejad M, Guo H (2021) P-CFT: a privacy-preserving and crash fault tolerant consensus algorithm for permissioned blockchains, pp 26–31
16. Castro M, Liskov B (2002) Practical byzantine fault tolerance and proactive recovery. *ACM Trans Comput Syst (TOCS)* 20(4):398–461
17. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, Enyeart D, Ferris C, Laventman G, Manevich Y et al (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains, pp 1–15
18. Gorenflo C, Lee S, Golab L, Keshav S (2020) Fastfabric: scaling hyperledger fabric to 20,000 transactions per second. *Int J Network Manage* 30(5):e2099
19. Wang Z, Hu H, Cheng G (2019) Design and implementation of an SDN-enabled DNS security framework. *China Commun* 16(2):233–245
20. Yuan L, Kant K, Mohapatra P, Chuah CN (2006) Dox: a peer-to-peer antidote for DNS cache poisoning attacks. In: 2006 IEEE international conference on communications, vol 5. IEEE, pp 2345–2350

Parametric Analysis of Resource-Constrained Devices for Internet of Things Using FreeRTOS



Partha Pratim Ray 

Abstract Multithreading is a serious issue that is continuously repelling the growth of Internet of Things supported hardware platforms to get accommodated with the current real-time scenario. Further, power consumption and sleep period determining approaches are not feasible in such platforms. In this paper, a novel RiotModel framework is proposed to validate the appropriateness of resource-constrained IoT-based devices to cope up with the concurrency related issues in real time for an e-health care use case. The proposed model is implemented incorporating multiple real-time algorithms with help from FreeRTOS kernel, API and libraries. Results show promising approach to attain sustainability, concurrency and real-time affects into the existing IoT-based device pool.

Keywords IoT · Power consumption · Sleep period · Multithread · e-health care · Real-time OS

1 Introduction

Sustainability aspect of the existing resource-constrained embedded processing devices has become a major challenge for effective implementation of Internet of Things (IoT) [1–5]. Mainly, the understanding and analysis of the optimal requirement of minimal sleep period and power consumption of IoT-supported processors is the need of time. IoT-based applications such as sustainable e-health care could be benefited from these two perspectives. Moreover, similar other objectives could be facilitated by indulging these factors into the deployed system model. Thus, the issue related to power consumption and sustainability could be fulfilled in real life.

Run-time power management is directly related to the optimal selection of actual sleep duration [6–10]. These tasks are easily solved by any standalone computer system such as pc, laptop, workstation and server. However, such jobs become very difficult for the resource-constrained IoT-based devices (e.g., Arduino Uno, Mega

P. P. Ray (✉)

Department of Computer Applications, Sikkim University, Gangtok, Sikkim 737102, India
e-mail: parthapratimray1986@gmail.com; ppray@cus.ac.in

2560, Duemilanove, etc.). Real-time operating system (OS) is specialized and dedicated for solving these issues. But, due to aforesaid reasons, incorporation of real-time OS into such hardware become practically infeasible. Multithreaded task scheduling is also hard to achieve and implement in current sequential structure of such hardware. Thus, existing literature seems to fail to provide a concise model or methodology to counter these problems [11, 12]. In this paper, we propose novel RiotModel framework to facilitate the earlier issues in IoT-based scenario. A novel system model is thus prescribed to harness the capability of the RiotModel into a real-life experiment situation. We include FreeRTOS kernel supported API and header files to design a proof-of-concept for an e-healthcare application. We used a pulse sensor along with the RiotModel framework to demonstrate the IoT-based e-healthcare application while leveraging the multithreaded task scheduling in parallel.

The main contributions of this work are as follows.

- To propose and implement real-time RiotModel framework in conjunction to other IoT-supported tools;
- To propose the implied multithreaded procedures and algorithms to validate the applicability of the deployed RiotModel; and
- To compare and analyze the minimal sleep period and power consumption suitable for design of such application in run-time.

The rest of the letter is organized as follows. Section 2 presents the system design aspects of the proposed study. Section 3 analyzes and discusses the results obtained from the experiments done, leaving Sect. 4 to conclude this paper.

2 System Design

2.1 *RiotModel Framework*

We propose RiotModel framework to analyze the power consumption and sleep period in run-time for the targeted IoT-supported embedded devices, e.g., Arduino Uno, Arduino Mega 2560 and Arduino Duemilanove. The selected microcontroller boards have difference in their specifications that includes processor clock frequency, processor architecture, programmable memory capacity and pin-wise current propagation. We illustrate a novel e-healthcare scenario as a proof-of-concept along with the RiotModel framework to validate the effectiveness and related issues in real-life IoT-based deployments. The RiotModel consists of seven layers of intra-dependent segments that works in bottom-top manner. The aforesaid IoT devices act as host and core processing engine of the implied system. A pulse sensor is used to provide pulse rate into the framework. FreeRTOS kernel is positioned in middle layer of the framework to accommodate the system call mechanism via the API and header libraries. Mutex creation, power mode setup, sleep period fixing is done in conjunction to task generation segments in the said framework. An e-healthcare scenario is perceived

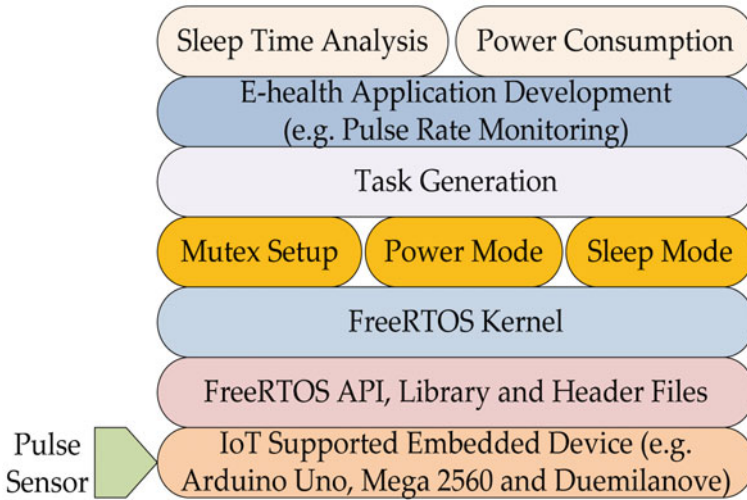


Fig. 1 Proposed RiotModel framework

on top of the RiotModel that in turn supports to analyze the sleep period and power consumption analysis that would be useful for betterment of a better heart and healthy life motto, when applied in reality. Figure 1 presents the concise architecture of the proposed framework.

2.2 Model Implementation

The RiotModel is applied by using three different IoT-supported devices as discussed earlier. Related programming, hardware setup and integration were done in support from a Dell Inspiron 5000 series machine (specifications Intel Core i7 7th generation processor, 8 GB DDR4 RAM, 1 TB HDD). Arduino IDE 1.8.2 was used for scripting, compilation and execution of the underlying programs. A pulse sensor manufactured by the World Famous Electronics LLC was attached to analog pin (A0) of the specific microcontroller board that receives the real-time pulse signal from a human’s finger. We incorporated the FreeRTOS kernel, API and header libraries to assist in the overall analysis of power consumption and sleep period while processing pulse signal in run-time. The FreeRTOS included herein supports heap generation, process queuing, message buffering, mutex semaphore, multithreaded task generation, timer optimization and port management. We incorporated one of the five algorithms at a time to analyze the power and sleep period calculations. Each one of the seven selected baud rates, e.g., 2400, 9600, 19,200, 57,600, 115,200, 500,000, 2,000,000 was preset into the algorithm to infer the variation in the output. The proposed system model data flow is illustrated in Fig. 2. In Fig. 2, RiotModel framework is fed with raw

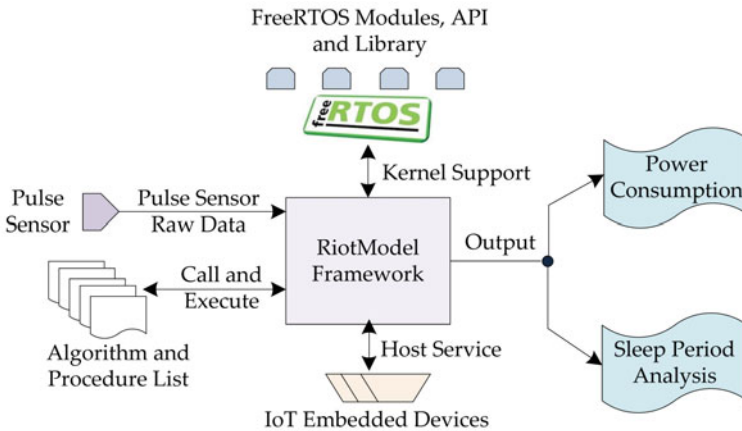


Fig. 2 Proposed system model data flow

pulse sensor data from user's body. FreeRTOS modules, application programming interfaces (API) and libraries are supported from the FreeRTOS kernel. Various algorithm and procedures are called and executed by RiotModel framework. IoT devices host various services. Final analysis is done in form of the power consumption and sleep period analysis.

2.3 Procedures and Algorithms

In this study, we have developed five algorithms, namely (i) empty loop, (ii) blink analog read, (iii) analog read digital write, (iv) blink analog read sleep and (v) blink analog read power. All these algorithms were written in wiring language and compiled using Arduino IDE 1.8.2. In all the algorithms, `Arduino_FreeRTOS.h` library was used. We also incorporated four different procedures such as, (i) multithreaded task creation, (ii) mutex creation, (iii) sleep mode selection and (iv) power consumption mode selection. The aim of the algorithm and associated procedures is to find the runtime (i) memory utilization, (ii) sleep duration analysis and (iii) power consumption.

Procedures. Usually, parallelization is never achieved in single core embedded processors as used in this study. Thus, Procedure 1 is used to create multithreaded tasks on top of the FreeRTOS kernel to run multiple tasks by the Arduino variants parallelly. Here, task names are explicitly mentioned with stack size of 128 bytes. Priority to the created task is allotted by 2 in first task "Blink" and 1 in "AnalogRead." More the value, more priority is achieved.

In Procedure 2, key part of mutex creation process is shown. While implementing this part, we included `semphr.h` header file to in this case to check whether the `serial_semaphore` is already created. Upon successful creation, a mutex is created to

access the serial port in run-time. It leaves the serial port when the job is done by calling `xSemaphoreGive()`.

In Procedure 3, sleep mode and duration is done in the `setup()` of the specific algorithm with conjunction to `avr/sleep.h`. Six different sleep modes were used to evaluate the outcomes that includes IDLE, ADC, PWR_DOWN, STANDBY and Ext_STANDBY while keeping SLEEP_MODE as common prefix. The mode selection was done by `set_sleep_mode()`. Locking on competing mutexting is controlled by keeping the flag of entering the critical section by `portENTER_CRITICAL()`. We used a set of various sleep functions to either make the CPU enable, disable, reset or push into the deep sleep.

Procedure 4 selects the power consumption varieties while utilizing the `avr/power.h` in context to mutually on/off the digital or analog pins. If an analog pin is being written via sensor data, digital pin buffer disabled and vice versa. Thus, power consumption is reduced. Pin sets such as DIR0 and 2 are also pre-initialized per the targeted architecture of processor. Similar analog comparator switch register (ACSR) could be switched on/off by looking into the flag value of analog comparator interrupt enable (ACIE) or analog comparator disable (ACD). A set functions are utilized to further reduce the power consumption as per the need of the application. Six types of power reduction functions are incorporated to disable the following modules (i) analog-to-digital converter, (ii) serial peripheral interface, (iii) I2C, (iv) timer 0, (v) timer 1 and (vi) timer 6.

Procedure 1. Multithreaded task creation

```
xTaskCreate (TaskBlink ("Blink", 128, NULL, 2, NULL));
xTaskCreate (TaskAnalogRead ("AnalogRead", 128, NULL, 1, NULL);
```

Procedure 2. Mutex creation

```
if (xSerialSemaphore == NULL) {
    xSerialSemaphore = xSemaphoreCreateMutex();
    if ((xSerialSemaphore) == NULL)
        xSemaphoreGive((xSerialSemaphore));}
```

Procedure 3. Sleep mode selection

```
set_sleep_mode(X) where X ∈ SLEEP_MODE_IDLE OR ADC OR PWR_DOWN OR
PWR_SAVE OR STANDBY OR EXT_STANDBY
portENTER_CRITICAL();
Select sleep_enable() OR sleep_cpu() OR sleep_mode() OR sleep_disable() OR sleep_reset()
```

Procedure 4. Power consumption mode selection

```
DIDR0 ← 0xFF OR 0x3F OR 0xF3
DIDR2 ← 0xFF OR 0x3F
Select ACSR & = ~_BV(ACIE) OR != _BV(ACD);
Select power_adc_disable() OR power_spi_disable() OR power_twi_disable() OR
power_timer0_disable() OR power_timer1_disable() OR power_timer2_disable();
```

Algorithms Algorithm 1 presents an empty `setup()` and `loop()` to keep itself as the base line in respect to other algorithms. Algorithm 2 calls Procedure 1 and utilizes two created tasks `TaskBlink()` and `TaskAnalogRead()` for performing two tasks (i) blinking on-board default D13 pin as per ratio to prefixed `portTICK_PERIOD_MS` and pulse sensor reading on the A0 pin.

Algorithm 3 first calls Procedure 1() and Procedure 2(). Then, implements two created task in parallel `TaskDigitalWrite()` and `TaskAnlogRead()` to read pulse sensor value while balancing with the push button to cope up with such task.

Algorithm 4 and 5 are revised version of Algorithm 2 to simulate the outcome behavior in the specified ways.

Though, Procedure 2 is used in both the algorithms. Procedure 3 and 4 are utilized in very specific manner in the respective algorithms. Here, algorithm 4 is executed multiple time to get variations in the IDLE sleep mode due to its coherent applicability into the three Arduino variants. Similarly, algorithm 5 is used to calculate power consumption in run-time while setting the `POWER_SAVE` mode on due to its normalized minimal current consumption rate in all the targeted Arduino boards.

Algorithm 1. Empty loop

```
void setup () {};  
void loop () {};
```

Algorithm 2. Blink analog read

```
Call Procedure 1 ();  
void TaskBlink ()  
    {for (;;) {vTaskDelay(1000/portTICK_PERIOD_MS)}}  
void TaskAnalogRead ()  
    {for (;;) {pulseValue ← analogRead(Pulse_Sensor);}}
```

Algorithm 3. Analog read digital write

```
Call Procedure 1 ();  
Call Procedure 2 ();  
void TaskDigitalWrite()  
    {for (;;) {pulse ← digitalRead(PushPulseButton);}}  
void TaskAnalogRead()  
    {for (;;) {pulseValue ← analogRead(Pulse_Sensor);}}
```

Algorithm 4. Blink analog read sleep

```
Call Procedure 2 ();  
Call Procedure 3 ();  
Call TaskBlink () and TaskAnalogRead of Algorithm 2
```

Algorithm 5. Blink analog read power

```
Call Procedure 2 ();  
Call Procedure 4 ();  
Call TaskBlink () and TaskAnalogRead of Algorithm 2
```

3 Results and Evaluation

3.1 Results

Obtained results may be partitioned into three ways such as, (i) run-time memory utilization, (ii) sleep duration comparison and (iii) power consumption comparison of the three different IoT-supported hardware platforms, namely Arduino Uno, Mega 2560 and Duemilanove. In storage computation phase, Algorithm 1–5 were separately implemented and results were obtained. Table 1 presents the comparative memory utilization (e.g., program storage, dynamic variable storage and local variable storage) of such algorithms in three different hardware platforms.

All values are represented in Bytes. (%) is integer approximation of total available program memory as follows, Arduino Uno: 32256 Bytes, Arduino Mega 2560: 253,952 Bytes, Arduino Duemilanove: 30,720 Bytes.

All current consumptions per modes are measured while executing the Blink Analog Read Sleep procedure in the Arduino Uno, Mega 2560 and Duemilanove.

Arduino Mega consumes least memory for algorithm 1, i.e., 2%. In other algorithms 2–5, we found that Arduino Mega utilizes minimal amount of memory than other two microcontroller platforms. Arduino Uno and Duemilanove have almost similar amount of storage capability thus a small change in respective algorithms. However, algorithm 3 has most memory consumable elements causing all the devices increase into respective storage limit, e.g., program storage 30%, 4% and 31% for Uno, Mega and Duemilanove, respectively. As seen, dynamic variables tend to

Table 1 Comparisons of storage parameters of IoT devices

Algorithm	Parameters	Uno	Mega	Duemilanove
1	Program storage	7018 (21%)	7086 (2%)	7018 (22%)
	Dynamic variables	163 (7%)	163 (1%)	163 (7%)
	Local variables	1885	8029	1885
2	Program storage	8690 (26%)	9184 (3%)	8690 (28%)
	Dynamic variables	356 (17%)	356 (4%)	356 (17%)
	Local variables	1692	7836	1692
3	Program storage	9718 (30%)	10,234 (4%)	9718 (31%)
	Dynamic variables	364 (17%)	364 (4%)	3624 (17%)
	Local variables	1684	7828	1684
4	Program storage	8738 (27%)	9220 (3%)	8738 (28%)
	Dynamic variables	365 (17%)	356 (4%)	356 (17%)
	Local variables	1692	7836	1692
5	Program storage	8670 (26%)	9160 (3%)	8670 (28%)
	Dynamic variables	356 (17%)	356 (4%)	356 (17%)
	Local variables	1692	7836	1692

Table 2 Comparisons of current consumption of IoT devices

Modes	Uno	Mega	Duemilanove
Active Mode	10 mA	20 mA	40 mA
SLEEP_MODE_PWR_DOWN	4 μ A	8.4 μ A	11.6 μ A
SLEEP_MODE_IDLE	2.4 mA	4.8 mA	6.5 mA
SLEEP_MODE_PWR_SAVE	6.5 μ A	8.8 μ A	9.4 μ A
SLEEP_MODE_STANDBY	0.2 mA	0.4 mA	0.6 mA
SLEEP_MODE_EXT_STANDBY	0.2 mA	0.6 mA	0.9 mA

increase from algorithm 2–5 subsequently in Arduino Uno, i.e., 7%–17%. Other two platforms do not respond drastically to such change, e.g., 7% and 15% for Mega and Duemilanove, respectively. Comparative results of the sleep mode selection and duration analysis are presented in Table 2 and Fig. 4. Table 2 compares the pin-wise current consumption for each of the earlier mentioned sleep modes. It is found that PWR_SAVE sleep mode consumes minimum current consumption in the pins of the hardware boards. Uno consumes the lowest of all, i.e., 6.5 μ A. We utilized MODE_IDLE to implement the algorithm 4 for the study due to its homogeneity and incorporation aspect in all the three boards. It is also observed that STANDBY or EXT_STANDBY sleep modes consume higher current, thus higher power. The three hardware were tested against the PWR_DOWN sleep mode that shows a linear increase of current consumption, i.e., 4 μ A, 8.4 μ A and 11.6 μ A. Once activated, only external hardware interrupts are able to awake the devices from deep sleep. Figure 4 illustrates the active mode and sleep mode duration per 1000 ms for each of the sleep modes as mentioned. Lower the duration of sleep, higher the power consumption option due to the occurrence of repetitive active mode that draws the maximum current, e.g., 10 mA, 20 mA and 40 mA for Uno, Mega and Duemilanove, respectively (Fig. 2).

Figure 3 presents the comparison of the power consumption occurred in all the sleep modes by the three devices. As expected, the power consumption is reduced as the sleep time increases. The maximum power consumption is found at 15 ms sleep mode causing Duemilanove to consume maximum amount of power. However, the Uno showed relatively more stable power consuming behavior throughout the sleep modes (Fig. 4).

3.2 Evaluation and Discussion

While evaluating the mapping and peak sensitivity of the deployed pulse sensor in this study, we found the following. Figure 5 shows the comparison of pulse sensor-based peak reachability in all the three hardware. Here, we applied different baud rate ranging from 2.4 to 2000 kbps to each of the device while evaluating the peak reachability accuracy mapping. We mapped the peak reachability accuracy within

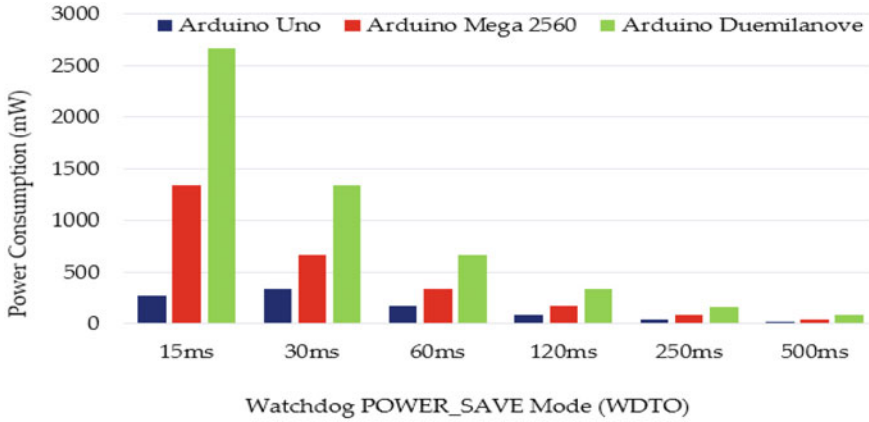


Fig. 3 Comparison power consumption versus watchdog power save mode

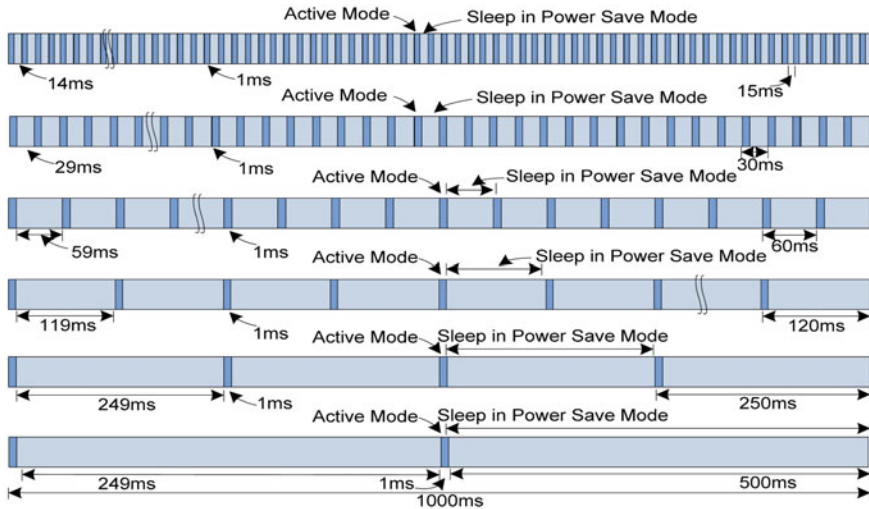


Fig. 4 Various sleep mode versus active mode slotting in 1000 ms duration

0–1 range to make it more convenient and comprehensive to understand. Uno reaches demonstrates the pulse peak reachability to the top at the 115.2 kbps serial clock rate. Same happens for the Mega at the 9.6 kbps clock rate. However, Duemilanove seems to have more peak reachability mapping points at the higher clock rates, e.g., 115.2, 500 and 2000 kbps. In other aspect, we can evaluate the pulse width wise sensitivity of all the hardware in different serial clock rates. We evaluated such sensitivity mapping within 0–1 range to express into more realistic way. We found that the lesser the pulse width gap, more the pulse rate sensitivity. Here, pulse width gap means the signal time period after the pulse sensor stop sending data to the hardware till start of

sending of pulse sensor data. It was observed that Duemilanove more sensitivity to such pulse width gaps. It outperforms other devices in 9.6, 19.2, 57.6 kbps serial clock rates. Uno shows least sensitivity at 57.6 kbps clock rate, while its sensitivity increases when clock rate increases. Mega also shows a promising sensitivity to width gaps at 9.6, 500 and 2000 kbps. Thus, we may infer that higher clock rate provides higher sensitivity for more resource-intensive devices such as Uno and Mega than Duemilanove. Comparison of pulse width wise sensitivity versus serial clock rate is shown in Fig. 6.

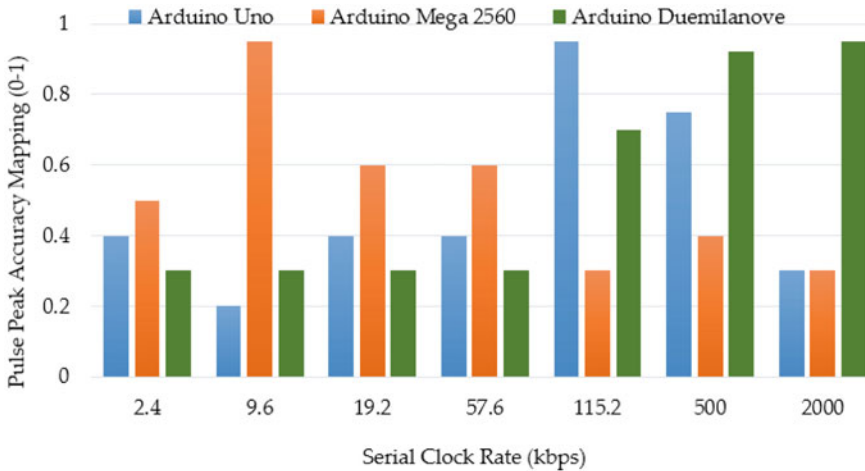


Fig. 5 Comparison of pulse peak accuracy versus serial clock rate

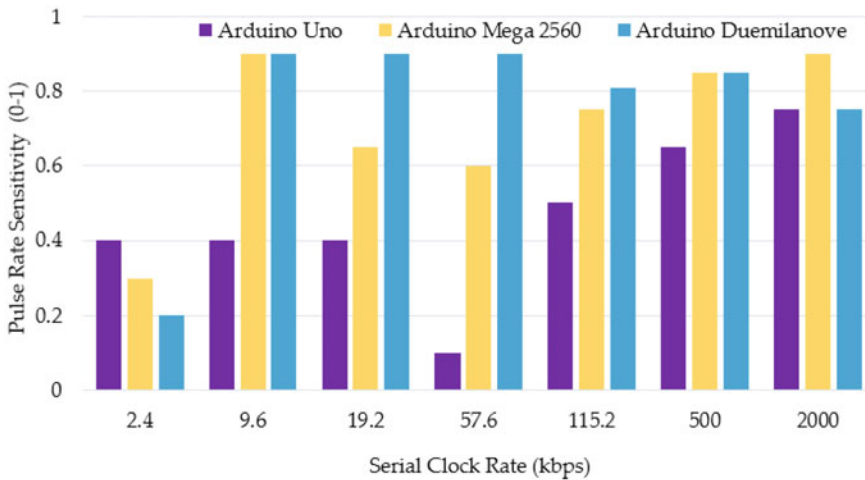


Fig. 6 Comparison of pulse width wise sensitivity versus serial clock rate

Novelty of Work. Thus, this paper aims to cover an important research gap that leads to the real-time OS aware multithreading aspect for the IoT-based devices. We find no work that advocates for such augmentation in the IoT-aware ecosystem, especially for extremely resource-constrained devices. Our proposed work demonstrates that FreeRTOS-enabled kernel specific performance evaluation on low-memory foot print IoT devices. The proposed RiotModel framework is capable to infer the sleeping period analysis and power consumption to validate the efficiency of the low-resource microcontrollers for appropriate IoT-based application deployment.

4 Conclusion

This paper proposed RioTModel framework to validate and analyze the sleep period and power consumption of the real-time OS-enabled IoT device pool. The results obtained advocate over the relationship between sleep period modes and power consumption for sustainability in futuristic IoT-based scenario. Pulse sensor assisted services to the RiotModel has successfully confirm the necessity and appropriateness of harnessing of IoT-based e-healthcare scenario in coming days. We perform comparison of power consumption versus watchdog power saver mode in this work. We notice that Arduino Uno consumes lowest power among all the three microcontroller boards in all watchdog mode, e.g., 15, 30, 60, 120, 250 and 500 ms. Various active and sleep modes are illustrated in this paper. We show that pulse peak accuracy mapping (0–1) is highest for 2000 kbps. Arduino Duemilanove has highest pulse peak accuracy of all for higher kbps. Arduino Mega has highest pulse peak accuracy mapping at 9.6 kbps. Arduino Uno shows higher pulse peak accuracy at 115.2 kbps. Pulse rate sensitivity analysis shows Arduino Duemilanove has higher pulse rate sensitivity at 9.6, 19.2, 57.6, 115.2, and 500 kbps. Arduino Uno has higher pulse rate sensitivity at 2.4 kbps. In the future, we will investigate how to automate the whole process by using dashboard integration.

References

1. Noor-A-Rahim M et al (2021) Robust and real-time state estimation of unstable microgrids over IoT networks. *IEEE Syst J* 2176–2185
2. Shubham, Reza M, Choudhury S, Dash JK, Roy DS (2020) An AI-based real-time roadway-environment perception for autonomous driving. In: 2020 IEEE international conference on consumer electronics—Taiwan (ICCE-Taiwan). IEEE, Taoyuan, Taiwan, pp 1–2
3. Kumari M, Kumar A, Singhal R (2020) Design and analysis of IoT-based intelligent robot for real-time monitoring and control. In: 2020 International conference on power electronics & IoT applications in renewable energy and its control (PARC). IEEE, Mathura, India, pp 549–552
4. Kanan R, Arbess H (2020) An IoT-based intelligent system for real-time parking monitoring and automatic billing. In: 2020 IEEE international conference on informatics, IoT, and enabling technologies (ICIoT). IEEE, Doha, Qatar, pp 622–626

5. D'Aloia M et al (2020) Low cost IoT sensor system for real-time remote monitoring. In: 2020 IEEE international workshop on metrology for Industry 4.0 & IoT. IEEE, Roma, Italy, pp 576–580
6. Madhu, Harikrishnakumar R, Nannapaneni S, Krishnan K (2020) Real-time event-based dynamic multi-facility vehicle rerouting. In: 2020 IEEE 23rd international symposium on real-time distributed computing (ISORC). IEEE, Nashville, TN, USA, pp 152–153
7. Maru V, Nannapaneni S, Krishnan K (2020) Internet of things based cyber-physical system framework for real-time operations. In: 2020 IEEE 23rd international symposium on real-time distributed computing (ISORC). IEEE, Nashville, TN, USA, pp 146–147
8. Duman Ü, Aydin E (2020) IOT based baby cradle system with real time data tracking. In: 2020 5th international conference on computer science and engineering (UBMK). IEEE, Diyarbakir, Turkey, pp 274–279
9. Wang S, Yan H (2020) Design of real-time monitoring platform for internet of things based on cloud platform. In: 2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC). IEEE, Chongqing, China, pp 61–64
10. Li R, Zhang T, Wei C, Liu Y, Tao L (2020) Real-time monitoring and comprehensive analysis framework for rainwater and soil pollution based on big data feature extraction and IoT. In: 2020 3rd international conference on intelligent sustainable systems (ICISS). IEEE, Thoothukudi, India, pp 242–245
11. Yew HT, Ng MF, Ping SZ, Chung SK, Chekima A, Dargham JA (2020) IoT based real-time remote patient monitoring system. In: 2020 16th IEEE international colloquium on signal processing & its applications (CSPA). IEEE, Langkawi, Malaysia, pp 176–179
12. Hu CC, Pan JS (2020) Maximum profit of real-time IoT content retrieval by joint content placement and storage allocation in C-RANs. *IEEE Trans Cloud Comput*

A Text Analytics Approach of Exploratory Visualization of Legal Parameters of Dowry Death Cases



Souraneel Mandal, Sajib Saha, and Tanaya Das

Abstract Natural Language Processing (NLP) extracts relevant information from unstructured data. It is mostly used to recover words or terms from unstructured text to comprehend the meaning present within the documents. Indian judicial systems rely on unstructured textual sources. Retrieving information by analyzing legal documents is a time-consuming task for legal professionals. Natural Language Processing and its techniques, i.e., text analytics can support in analyzing cases to a greater extent. Hence, in this paper, the authors have proposed an exploratory visualization approach to retrieve most relevant parameters found in dowry death documents, which is one of the heinous women centric crimes in recent times. The authors have applied the text analytics concepts like weighted mean, correlation matrix and clustering techniques to visualize the major parameters present in dowry death cases. The relevant comprehensive parameters assist legal professionals in deriving insights from legal cases on dowry death by clearly displaying the major parameters specified in Indian Penal Code Section 304B.

Keywords Exploratory visualization · Dowry death · Legal data retrieval · Text analytics · Term frequency-inverse document frequency

1 Introduction

Text analytics is a Natural Language Processing (NLP)-based technique that helps in analyzing the unstructured text to generate some meaningful information from the text. As a result, it should come as no surprise that text analytics technique will take their appropriate place in Indian Legal System. As most of the documents in Indian Legal System are unstructured in nature, it is quite difficult for legal professionals to analyze and extract major parameters present in legal documents. To retrieve the major parameters from legal documents, legal professionals proceed manually which is time-consuming practice. Legal Data Retrieval (LDR) aims to identify the relevant

S. Mandal · S. Saha · T. Das (✉)
Adamas University, Kolkata, India
e-mail: tanayadas.das23@gmail.com

information from legal cases. The objective of Legal Data Retrieval (LDR) aims different records like court transcripts, decisions, legislation archives and judgments that are created amid the course of a legal process. These archives are essential assets for the translations of the law of any legal executive and henceforth are expected by a legitimate proficient for decision making as well as argumentation. Explicit attributes of authoritative archives like report size, record inside structure, temporal properties, specific legal terminology, polysemy and heterogeneity make LDR different and complicated when contrasted with different areas.

Recently, it has been observed that in Indian Legal System, text analytics is working on sectors like document summarization [1], referencing statutes using linguistic [2], semantic information from archives of case decisions [3], sentiment analysis [4], text mining [5], keyword extraction [6], search engine [7], big data analytics [8], multitask learning based on text classification [9], etc. Rather than concentrating on all kinds of legal sections, the authors have focused on cases related to crime against woman. According to National Crime Records Bureau (NCRB) data [10], dowry death is an increasing significant crime due to illegal dowry demand, leading to death of woman. Hence, in this paper, authors are mainly visualizing major parameters of dowry death cases and trying to capture the specific parameters as mentioned in dowry death legal rule, i.e., Indian Penal Code Section 304B. The authors have used conventional TF-IDF-based vector to show the dispersion of words within the case records. This strategy is displaying relationship among words or concepts that represents co-occurrence, semantic closeness, etc., present in dowry death cases.

Section 2 portrays the related study on Natural Language Processing and Term Frequency-Inverse Document Frequency procedures applicable for the legal system. Section 3 describes the details of mathematical modeling of TF-IDF. Section 4 outlines the proposed methodology to calculate the weighted mean and relationship between two cooperative words or terms through representation in legitimate cases. Section 5 discusses the results of the proposed work with the help of correlations matrix, heatmap visualization and hierarchical clustering to explore the major parameters. Finally, Sect. 6 concludes the paper and its future use in Indian Legal System.

2 Literature Survey

In NLP, there are many techniques like Bag of Words, Binarization, Term Frequency-Inverse Document Frequency (TF-IDF), etc., among of them utilization of the TF-IDF is widely approachable because of logarithmic approach on frequency. In 1972, the TF-IDF calculation was first proposed by Sparck Jones [11] from Computer Scientist at the University of Cambridge. Many advancements were made gradually to upgrade the TF-IDF ideas, and it was applied in various fields for improved solutions.

Numerous researchers proposed upgrades to the calculation, for example, in the year 2016, few analysts dealt with Term Frequency-Inverse Document Frequency

(TF-IDF) to figure out what words in a corpus of archives may be better to use in a question. As the term infers, TF-IDF computes values for each word in a report through a reverse extent of the recurrence of the word in a specific record to the level of archives the word shows up in. Words with high TF-IDF numbers infer a solid relationship with the report they show up in, recommending that word was to show up in a question, and the record could hold any importance with the client. Thus, they gave proof that this basic calculation effectively classifies pertinent words that can upgrade inquiry recovery [12]. After a few time 2018, a few investigate worked on the framework named “FEVER 2018 Shared Assignment.” They portrayed as DeFactoNLP. The point of this errand was to conceive a framework that cannot as it were naturally evaluate the veracity of a claim but to recover prove supporting this evaluation from Wikipedia. Their approach on the Wikipedia records whose Term Frequency-Inverse Document Frequency (TF-IDF) vectors are most comparative to the vector of the claim and those archives whose names are comparable to those of the named substances (NEs) specified within the claim are identified as the archives which might contain prove [13]. During the same school year, other students worked on implementing the TF-IDF algorithm on a website to determine whether or not a “word” is included in every document and to calculate the total number of documents in hand. To do this, they simply retrieved data from websites, then used preprocessing to eliminate HTML/CSS elements, StopWords (all unneeded words are irrelevant in this case), and lastly counted the total number of words and their occurrences across all texts [14]. In 2019, an article has been distributed on order framework to proficiently uphold the paper arrangement, which is fundamental to give clients quick and effective quest for their ideal papers. The proposed framework consolidates TF-IDF and LDA plans to ascertain a significance of each paper and gatherings the papers with comparative subjects by the K-implies grouping calculation. It can in this way accomplish right characterization results for clients’ intriguing papers [15]. Recently, few individual analysts propose STF-IDF, a novel semantic touchy strategy based on the ordinary TF-IDF. The key thought is rearranging the ordinary TF-IDF scores based on the semantic representation of most important words. Subsequently, they proposed that in the event that a set of terms is considered vital by TF-IDF, and all the semantically comparable words related to this set ought to be considered more vital than those ones with less semantic pertinence to the setting [16]. To address this scenario about deviation of DF esteem, the researchers have proposed the assortment recurrence factor by adding Normal Record Recurrence (ADF) into computation to resolve the issues referenced previously [17].

From the overall literature review, it has been found that most of the cases are based on Sentiment Analysis, Keyword Extraction, Summarization, etc. Very limited literatures are there that uses exploratory visualization of the parameter of any legal documents. Hence, authors of this paper applied TF-IDF to all dowry death’s legal archive to analyze and implement TF-IDF method to extract the primary words or concepts that legal professionals mostly concerned in dowry death cases as stated in legal rule.

3 Background Work

TF-IDF or weighted mean was acquainted with address the significant disadvantages of Bag of words by presenting a significant idea called inverse document frequency. It is a score which the machine keeps where it assesses the words utilized in a sentence and measures its use contrasted with words utilized in the whole archive. As such, it is a score to feature each word's significance in the whole record.

The detailed description of the mathematical concepts behind the Term Frequency-Inverse Document Frequency as follows:

Term Frequency: TF of a word is the times the word shows up in a sentence contrasted with the absolute number of words in the sentence.

$$TF(w, s) = \frac{\text{no of repetition of } \mathbf{words} \text{ in a } \mathbf{sentence}}{\text{no of } \mathbf{words} \text{ in a } \mathbf{sentence}} \text{ or } \frac{\sum_{w \in s} w}{|s|} \quad (1)$$

where, w and s represent *word* and *sentence*, respectively.

Inverse Document Frequency: IDF is basically the complementary of sentence recurrence. Along these lines, it does the trick to show what sentence recurrence is, since IDF would promptly follow from document frequency. Practically speaking, by applying a logarithm to forestall the IDF score from detonating. Additionally, by adding some smoothing to forestall division by nothing. There is by all accounts numerous varieties of how smoothing is carried out.

$$IDF(w, S) = \frac{\text{no of } \mathbf{sentence}}{\text{no of } \mathbf{sentence} \text{ containing } \mathbf{words}} \text{ or } \log\left(\frac{|S|}{\sum_{w \in S} w}\right) + 1 \quad (2)$$

where, w and S represent *word* and *sentences*, respectively.

Term Frequency-Inverse Document Frequency: Then TF-IDF is calculated by multiplying Eqs. 1 and 2.

$$TF - IDF(w, s, S) = TF(w, s) \times IDF(w, S)$$

where, w , s and S represent *word*, *sentences* and *sentences*, respectively.

4 Work Methodology

The methodology includes the origin of this research work and the extension using Natural Language Processing technique. The details of the proposed methodology is described in the subsequent section.

4.1 Origin of Work

As origin [18] of this research work, authors have used dowry death legal text documents containing unstructured data as the input data. The authors have used some preprocessing techniques like tokenization, lemmatization, etc. The authors have removed stop words like ‘and,’ ‘he,’ ‘she,’ etc., from the tokens and also removed all the punctuation and single letters and words containing any number of digits, last names, website addresses, tags to extract the most influential features from the unstructured legal data. The authors have stored the tokens of each document in separate lists. The authors have applied countvectorizer on each list of tokens and found the frequency of each token in a list and applied outer join on the lists and found the countvector matrix which shows the frequency of the words in each document. The major words present in most of the documents are ‘*person,*’ ‘*accused,*’ ‘*evidence,*’ ‘*husband,*’ ‘*dowry,*’ ‘*death,*’ etc., as shown in Fig. 1. In this paper, the authors have proposed an approach of TF-IDF as extension of this work as explained in Sect. 4.2.

As author of this paper used countvector matrix to perform frequency of the words for all documents case which is expressed in previous work [18] is as shown in here.

	word	Case1	Case2	Case3	Case4	Case5	Case6	Case7	Case8
0	person	65.0	22.0	1.0	39.0	8.0	10.0	10.0	0.0
1	accused	56.0	15.0	4.0	35.0	21.0	9.0	33.0	8.0
2	evidence	49.0	15.0	37.0	34.0	45.0	44.0	9.0	6.0
3	husband	46.0	16.0	2.0	27.0	19.0	12.0	8.0	4.0
4	dowry	46.0	12.0	16.0	28.0	12.0	21.0	14.0	16.0
...
127	mother	1.0	0.0	11.0	15.0	12.0	16.0	6.0	4.0
128	examined	1.0	8.0	3.0	9.0	2.0	0.0	4.0	1.0
129	place	0.0	8.0	1.0	14.0	8.0	9.0	2.0	6.0
130	told	0.0	7.0	6.0	8.0	11.0	4.0	2.0	1.0
131	could	0.0	6.0	5.0	3.0	5.0	3.0	5.0	12.0
132 rows × 9 columns									

Fig. 1 Stepwise proposed methodology for frequency matrix. Source [18]

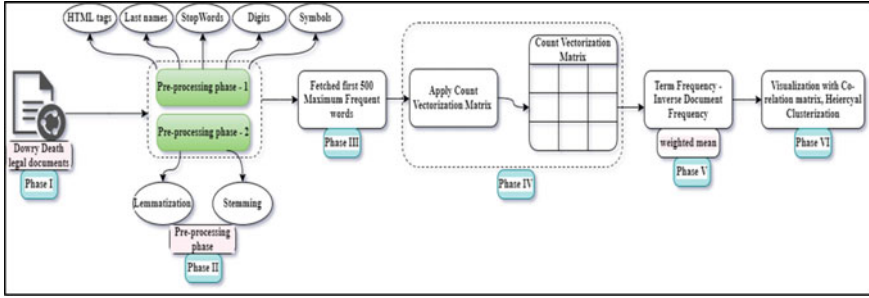


Fig. 2 Stepwise proposed methodology for weighted mean

4.2 Proposed Approach

In this paper, authors proceed with the help of TF-IDF to understand the context of sentence and visualized with multi-words' correlation which describe with algorithmic approach in subsequent section. Stepwise proposed methodology for weighted mean is shown in Fig. 2.

Proposed Algorithm :

- Step 1: **Start.**
- Step 2: Import the appropriate packages, modules like *nltk*, *pdftotext*, *glob*, *seaborn*, *CountVectorizer*, *pandas*, *TfidfVectorizer*.
- Step 3: Initialize *max_count* to 500 and stored all Dowry death legal documents in sorted manner with the help of *glob* module.
- Step 4: **Phase I–IV** is already explained as work methodology in previous work [18].
- Step 5: As an extended work, in **Phase V**, the output received from previous research work that extracted the major parameter found in dowry death cases was further represented as TF-IDF Vectorization using the mathematical equation as explain in I and II.
- Step 6: Applied *TfidfVectorizer* in proposed DataFrame to find out weighted mean.
- Step 7: Then visualization with *seaborn* module expressed the correlation matrix and hierarchical cauterization to find out how to much two words are associated with each other which are the extension of previous work in Sect. 4.1.
- Step 8: **End.**

The authors have shown the architecture of NLTK libraries which are used which are used to preprocess the legal documents as shown in Fig. 3.

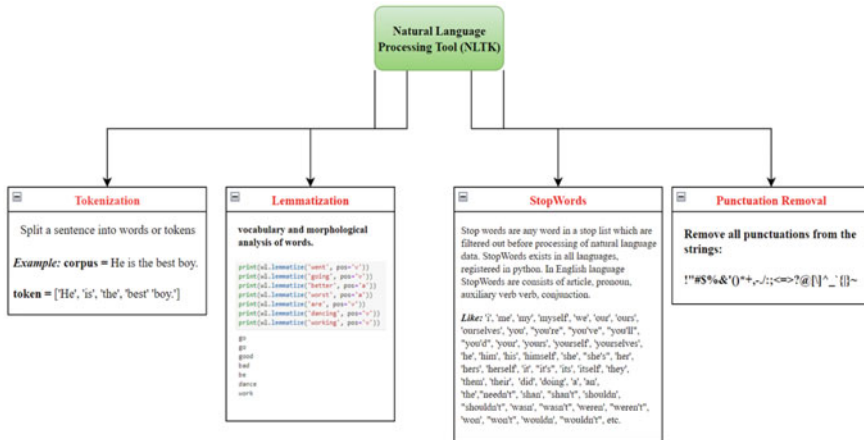


Fig. 3 Proposed NLTK architecture

5 Results and Discussion

As shown in Fig. 4, authors have obtained the output as weighted mean of the most frequent words found in dowry death cases. The matrix comprises words like ‘person,’ ‘accused,’ ‘evidence,’ ‘husband,’ ‘dowry,’ etc., which are the significant parameters on which legitimate experts are generally focused in Indian Penal Code Section 304B [19] which describes whether the individual is found accused for any dowry death case or not. The authors have found that the main words like ‘accused,’ ‘husband,’ ‘dowry’ have highest weighted mean than normal parameters. For example, as in case 1, the word ‘dowry’ has a higher weighted mean as compared to case 4. Hence, in case 1, the word ‘dowry’ has much higher impact of dowry demand by the accused persons from rest of the cases. Similarly, the weighted mean matrix is showing the mean of the words present in dowry death cases, as this is the type of information, and the legal professionals are mostly concerned with.

As shown in Fig. 5, authors have visualized the TF-IDF vector matrix using a seaborn heatmap. Here, the authors have visualized the top 20 most frequent words have as a map plot and the higher values of the parameters are shown in a lighter color, and lower values are shown in a darker color which shows how the parameters are correlated with each other corresponding to the dowry death cases.

	Word	Case1	Case2	Case3	Case4	Case5	Case6	Case7	Case8
0	person	0.799505	0.278934	0.012971	0.490737	0.099149	0.124880	0.128727	0.000000
1	accused	0.696143	0.192208	0.052436	0.445097	0.263039	0.113589	0.429326	0.105673
2	evidence	0.501329	0.158194	0.399200	0.355862	0.463907	0.457049	0.096368	0.065229
3	husband	0.750328	0.269020	0.034402	0.450540	0.312276	0.198728	0.136567	0.069329
4	dowry	0.687359	0.184832	0.252120	0.428016	0.180675	0.318587	0.218936	0.254044
...
126	mother	0.034532	0.000000	0.400568	0.529895	0.417536	0.560951	0.216838	0.146773
127	examined	0.073226	0.603848	0.231660	0.674196	0.147567	0.000000	0.306542	0.077809
128	place	0.000000	0.381295	0.048760	0.662225	0.372720	0.422500	0.096782	0.294792
129	told	0.000000	0.413889	0.362935	0.469442	0.635770	0.232948	0.120063	0.060951
130	could	0.000000	0.357017	0.304368	0.177160	0.290823	0.175821	0.302065	0.736057

Fig. 4 TF-IDF vector matrix or weighted mean for dowry death cases

As shown in Fig. 6, the authors have visualized the TF-IDF vector in a seaborn cluster map. The lower values are represented in darker color and the higher values are represented in lighter color. Each correlated word is represented as clusters, also known as dendrograms. As shown in dendrograms, merging two closest clusters like in the row section, the value of ‘person’ and ‘husband’ has the minimum distance respect to all the columns, so they are merged to one cluster which is merged with the cluster ‘stated’ and then again merged with ‘accused’ which again got merged with the cluster made by merging of ‘cruelty’ and ‘harassment.’ This shows which parameters are closest to each other and how each parameter is correlated with another corresponding to the legal documents.

6 Conclusion and Future Work

Visualization has a greater impact and aids in deriving insights. In this paper, the authors have proposed some text analytics approaches to visualize the major words and their linkages with other parameters present in dowry death cases. From the proposed work, a legal professional can easily understand the relevant words present while analyzing any dowry death cases. It relieves the legal professionals from the time-consuming process of manually analyzing legal cases. The legal professionals quickly extract the major words from the weighted mean, heatmap and cluster-based

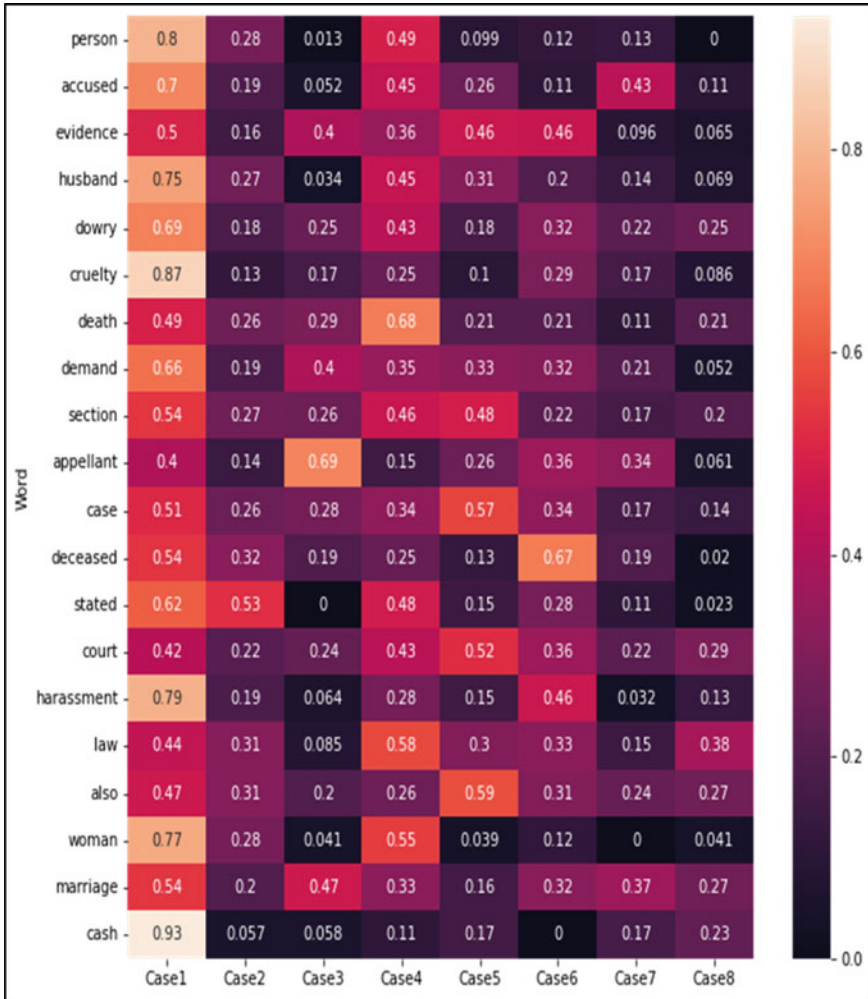


Fig. 5 Correlation matrix for dowry death cases

technique to extract key parameters present in dowry death cases. Also, this work is clearly stating the major concepts found in Indian Penal Code Section 304B. In the future, this work will help to generate a structured dataset on dowry death cases for any machine learning-based work, as it states the important parameters of dowry death cases.

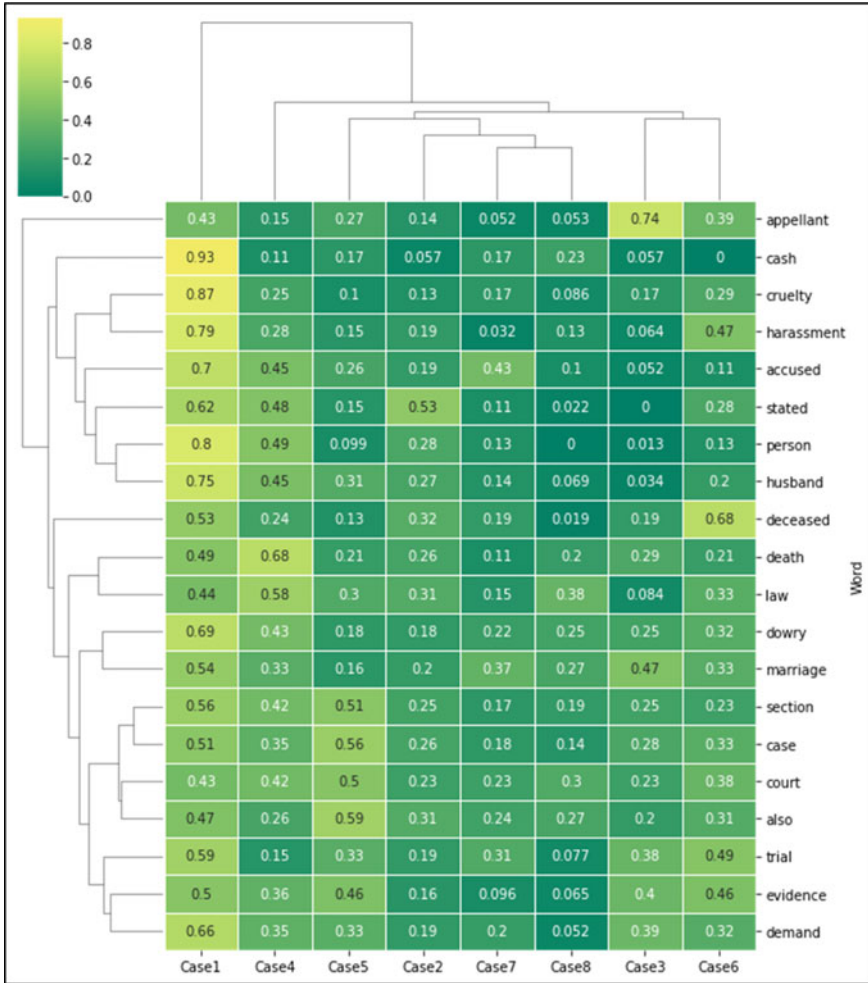


Fig. 6 Hierarchical cluster analysis for dowry death cases

References

1. Sullivan D (2001) Document warehousing and text mining. Wiley, New York
2. Capotorti F (1979) Study on the rights of persons belonging to ethnic, religious, and linguistic minorities. United Nations, New York
3. Yadla HK, Rao DPVRDP (1970) Machine learning based text classifier centered on TF-IDF vectoriser: semantic scholar. Undefined [Pdf]. Retrieved 2 May 2022, from <https://www.semanticscholar.org/paper/Machine-Learning-Based-Text-Classifier-Centered-On-Yadla-Rao/55f834d9ed747bc2c9d51f74f8ed54fb13fa8aed>

4. Malik U (2019) Twitter sentiment analysis using TF-IDF approach. I am Machine Learning and Data Science expert currently pursuing my PhD in Computer Science from Normandy University. GoTrained Python Tutorials. Retrieved 28 April 2022, from <https://python.gotrainned.com/tf-idf-twitter-sentiment-analysis/>
5. Yu N (2018) A visualized pattern discovery model for text mining based on TF-IDF weight method. In: 2018 10th international conference on intelligent human-machine systems and cybernetics (IHMSC). <https://doi.org/10.1109/ihmsc.2018.10148>
6. Ganesan K (2021) Python keyword extraction tutorial using TF-IDF. Retrieved 28 April 2022, from <https://kavita-ganesan.com/python-keyword-extraction/>
7. Akella K (2020) Implementing the TF-IDF search engine. Medium. Retrieved 28 April 2022, from https://medium.com/@kartheek_akella/implementing-the-tf-idf-search-engine-5e9a42b1d30b
8. Bestak R, Smys S (2019) Big data analytics for smart cloud-fog based applications. *J Trends Comput Sci Smart Technol (TCSST)* 1(02):74–83
9. Jacob IJ (2020) Performance evaluation of caps-net based multitask learning architecture for text classification. *J Artif Intell* 2(1):1–10
10. National Crime Records Bureau (n.d.) Retrieved 2 May 2022, from <https://ncrb.gov.in/en/crime-india-2018>
11. Sparck Jones K (1972) A statistical interpretation of term specificity and its application in retrieval. *J Documentation* 28(1):11–21. <https://doi.org/10.1108/eb026526>
12. Text mining: use of TF-IDF to examine the relevance of ... (n.d.) Retrieved 24 April 2022, from https://www.researchgate.net/profile/Shahzad-Qaiser/publication/326425709_Text_Mining_Use_of_TF-IDF_to_Examine_the_Relevance_of_Words_to_Documents/links/5b4cd57fa6fdcc8dae245aa3/Text-Mining-Use-of-TF-IDF-to-Examine-the-Relevance-of-Words-to-Documents.pdf
13. Reddy AJ, Rocha G, Esteves D (2018) DeFactoNLP: fact verification using entity recognition, TFIDF vector comparison and decomposable attention. arXiv.org. Retrieved 28 April 2022, from <https://arxiv.org/abs/1809.00509v1>
14. Qaiser S, Ali R (1970) Text mining: use of TF-IDF to examine the relevance of words to documents: semantic scholar. Undefined. Retrieved 24 April 2022, from <https://www.semanticscholar.org/paper/Text-Mining%3A-Use-of-TF-IDF-to-Examine-the-Relevance-Qaiser-Ali/f3b9ab8573a8b0f4de6b1dd1387e175dada698e0>
15. Kim S-W, Gil J-M (2019) Research paper classification systems based on TF-IDF and LDA schemes—human-centric computing and information sciences. SpringerOpen. Retrieved 28 April 2022, from <https://doi.org/10.1186/s13673-019-0192-7>
16. Jalilifard A, Caridá VF, Mansano AF, Cristo RS, da Fonseca FPC (1970) Semantic sensitive TF-IDF to determine word relevance in documents. SpringerLink. Retrieved 28 April 2022, from https://doi.org/10.1007/978-981-33-6987-0_27
17. Jiang Z, Gao B, He Y, Han Y, Doyle P, Zhu Q (2021) Text classification using novel term weighting scheme-based improved TF-IDF for internet media reports. *Math Probl Eng*. Retrieved 28 April 2022, from <https://www.hindawi.com/journals/mpe/2021/6619088/>
18. Mandal S, Saha S, Das T (2022) An approach to extract major parameters of legal documents using text analytics accepted at the ICT for sustainable development conference, 2022. GOA, INDIA. Paper accepted at 7th Edition ICT4SD International ICT Summit & Awards, paper id: 130
19. Khastgir J (2012) Criminal manual, p 469

Okra Cultivation Using IoT and Disease Detection



D. Abishek, S. Hareesh Chandran, S. Mayukha, and B. Sabitha

Abstract Okra is also commonly called ladies finger, bhindi and scientifically known as *Abelmoschus esculentus*. Okra is a remarkable success among traders because of its short cultivation time, with a substantial number of yields produced. In cultivation of okra, farmers face difficulties in maintaining the moisture in soil, detecting the diseases since okra is prone to acquire diseases, harvesting during the right period, and involving expensive labor. Scientists are trying hard to inculcate technology in farming to increase productivity in order to compensate for the growing population. One such technology used in farming is IoT. The most promising methods from the existing methods have been proposed in this paper. The IoT system consists of sensing, reporting, monitoring, and controlling phases. The IoT provides the information regarding the humidity, soil moisture, and temperature collected through MQTT protocol, and it is displayed on the system as a user interface (UI). This information is passed on to farmers' mobile where they can check the readings. The detection of diseases using image processing is also done using IoT.

Keywords Okra cultivation · Internet of Things (IoT) · Image processing · Message queuing telemetry transport (MQTT) · Smart irrigation

1 Introduction

Agriculture is important for India's economic growth. India is the world's second farm producer. Approximately, 17% of the total GDP is contributed by agriculture

D. Abishek · S. Hareesh Chandran · S. Mayukha (✉) · B. Sabitha
Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India
e-mail: svmayukha@gmail.com

D. Abishek
e-mail: abishek.durai2001@icloud.com

S. Hareesh Chandran
e-mail: haree2001@gmail.com

B. Sabitha
e-mail: sabitha.b.mce@kct.ac.in

in our country and gives employment to 60% of the population [1]. “Okra” is a crop that is considered primarily as a vegetable in most places, but also a fruit because of its sweet seed pods that are contained within. This crop is originally native to the Nile in Abyssinia (Ethiopia), later migrated to Africa, Egypt, and also has its wing in India. It is one among the most heat and drought tolerant species in the vegetable category across the world. It tolerates soil with high clay content and is adaptive to intermittent rainfall. But, drastic changes in climate could cause less production or eventually loss to the farmer. Other factors like high density planting, crop diseases, flies, and viruses also pose a threat to the better yield of the crop. It also plays a vital role in many countries as a major cash crop. Okra cultivation can be done even during water constraints. In most of the neighboring countries, okra is served as a main meal, side dish, and also even in salads. Okra is a great success among traders because of its short cultivation time, with a large number of yields produced. And a fact state that okra has the prettiest name—ladies’ finger, it also belongs to the hibiscus family [2].

Okra preferably grows in the soil having acidic nature, pH valued from 6 to 7. An average native okra crop grows up to 9 feet in height and some farmers, on their own method of cultivation, limit the height of the plant [3]. The optimum temperature at which the okra crop grows well ranges from 25 to 35 °C. Optimum soil pH for okra crop ranges from 6 to 6.8 [4]. The seeds of okra can be directly sown in soil. In common seeds are sown in a spacing of 60 cm × 45 cm during all seasons. The seeds can also be soaked overnight before sowing to germinate faster. The cultivating land is plowed at a depth of approximately 20–25 cm. Okra feeds on nutrients heavily. So, after the land is plowed, the fertilizers are added to the soil. The irrigation depends on the climatic condition. Generally, the land must be irrigated 4–6 days in hot season. Weeds must be removed in regular intervals of time to get better yield. Frequent removal of weed is the main problem faced by the farmers, as it includes high labor cost and time consuming. It can be harvested in 40–45 days [5]. The average yield is 10–12.5 t ha⁻¹. It can be stored at 7–10 °C temperature [6].

Nowadays technology gives a supporting hand to farmers in order to feed the population that is growing at a massive rate. With the help of modern technology, the burden that farmers have during harvesting, pest spraying, weed removal, disease detection, and lots can be reduced to a greater extent. Moreover, the extra expenses that farmers face during the wage day of laborers are totally lessened. One of the main advantages of using IoT technology in farming is that the farmer need not stay in land all the time surveying and monitoring, whereas all the activities in the farm can be controlled via sensor system, electronics and devices under any temperature, climate, unless there is any sort of natural calamity. The increase in global population is urging a shift toward smart agricultural practices [7]. Internet of Things (IoT) plays a major role in modern technology in farming such as calculating productivity, data analysis, precision farming, and predicting climate conditions. This system is used in detection of some vital needs such as sensing the temperature, humidity, crop health, and soil moisture for the crops to grow efficiently. It also contributes in irrigating the crops in the proper ratio, less time, with more conservation of water. By using IoT, the farmer can use the available resources in a more efficient manner [8]. The use of IoT enhances

the operational efficiency, quality of work, and productivity in the agriculture sector. The Internet of Things (IoT) is a promising technology that is generating a plethora of creative ideas for improving the agriculture economy. Research institutions and scientific organizations are constantly working to develop solutions and products that use IoT to tackle a wide range of agricultural problems. Besides sensing and monitoring systems, the key IoT applications in agriculture include precision farming, livestock, and greenhouses, which are classified into separate monitoring regions [9]. All of these are monitored with the use of various IoT-based sensors/devices via wireless sensor networks, which allows farmers to collect important data via sensing devices.

Sensors used in the IoT system will sense the data from soil, atmosphere, and moisture and will send the data to the reporting system using MQTT protocol [10]. The image processing is done to identify the diseases in the okra plant. Through this the farmers can get the information about the plant and can use the desired fertilizer and pesticides for the disease. The smart irrigation system using IoT will benefit the farmers in great amounts. The prototype measures the quantity of moisture in the soil. When the system is turned on and off, a message is sent to the user via MQTT protocol, updating the status of the water pump and soil moisture. This activates the irrigation system, pumping water to the plants. This helps in irrigating the field in a regular interval of time.

2 Methodology

2.1 Sensing System and Reporting System Using MQTT Protocol

The sensing system consisting of the temperature sensor and soil moisture sensor is connected to the microcontroller and it forms a module. The code has been fed into the microcontroller. The readings from the sensors are collected and sent to the microcontroller through integration. After that, the individual parameter readings are interfaced together and finally sent to the central hub (Raspberry Pi) [11]. Figure 1 illustrates the block diagram of the sensing and reporting module.

The information collected by the sensors is forwarded to the microcontroller. The soil moisture content and signal of water level will be sent to the central hub through MQTT protocol which is operated through Wi-Fi. Through this protocol, the farmers will message. “A microcontroller is the key device for the entire system to function along with the other electronic components.” The microcontroller then sends the gathered information to the MQTT broker using Wi-Fi. An MQTT broker is an intermediate member in receiving the data from the client that acts as the sender, filtering it by categories and finally sending the data to the subscriber, which acts as the receiver [12]. Figure 2 illustrates the integration with modules.

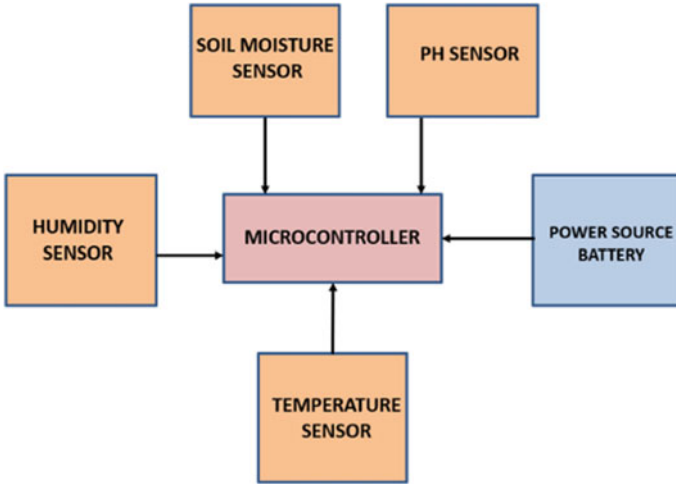


Fig. 1 Block diagram of the sensing and reporting module

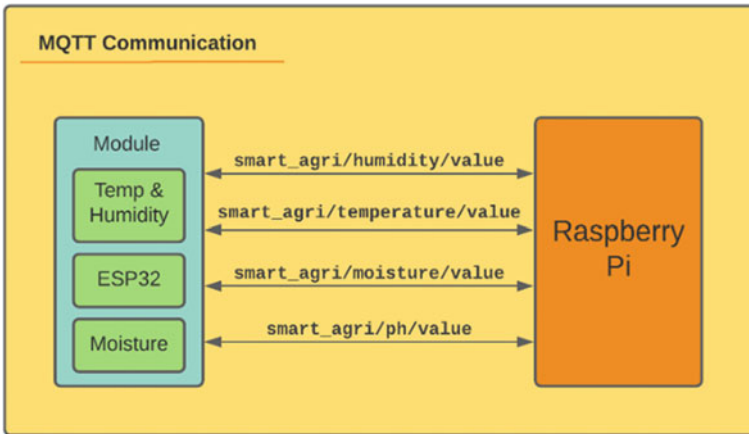


Fig. 2 Integration with modules

2.2 User Interface Using Node-RED

The data from the module is transferred via MQTT to the central hub over Wi-Fi. The Information is displayed as a user interface (UI) using Node-RED. In Node-RED, the UI displays the value of temperature and humidity, soil moisture in the form of line graph and gage meter [13]. The information is sent to the farmer’s mobile application, where they can check the periodic activity of the crops. The schematic representation of the monitoring and control system is depicted in Fig. 3.

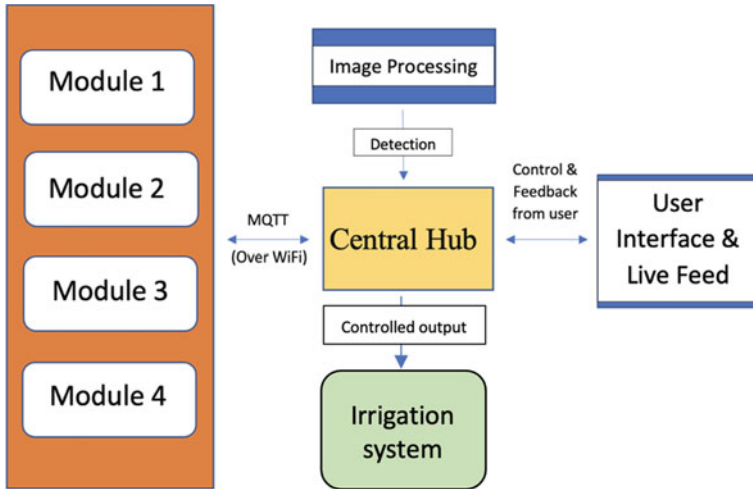


Fig. 3 Monitoring and control system

2.3 Disease Detections

As mentioned above, okra is prone to several diseases like yellow vein mosaic virus, enation leaf curl, etc. This type of disease can be sensed via the looks of the okra leaves. More images for the various diseases in okra are captured. Using the captured image, a dataset is created. So, a CNN model is trained using TensorFlow to classify the diseased okra from the good ones and thereby combining the data from modules and the trained model [14]. The proposed convolutional neural networks, or CNNs, have an excellent picture recognition technique built into its design. This architecture is used to extract text data from picture input [19]. The model is trained against the dataset of okra diseases. This model contains several layers including CNN, normalization, dropout, activation, etc. This model contains millions of trainable neurons where each neuron’s weight is tuned to give the expected results [15]. After training, the binaries of the model are stored locally for later use. The stored model is ported to the central hub which is connected to a camera [18]. The camera detects the disease in the okra plant. To cure the disease, fertilizer/pesticide is poured manually into the irrigation system so that it can reach the plants that are affected. The fertilizers are poured manually depending on the disease identified in optimum proportion to ensure that the excess amount of fertilizer/pesticides is not wasted, and the plant and soil are not affected. Figure 4 illustrates the flowchart of data process in microcontroller.

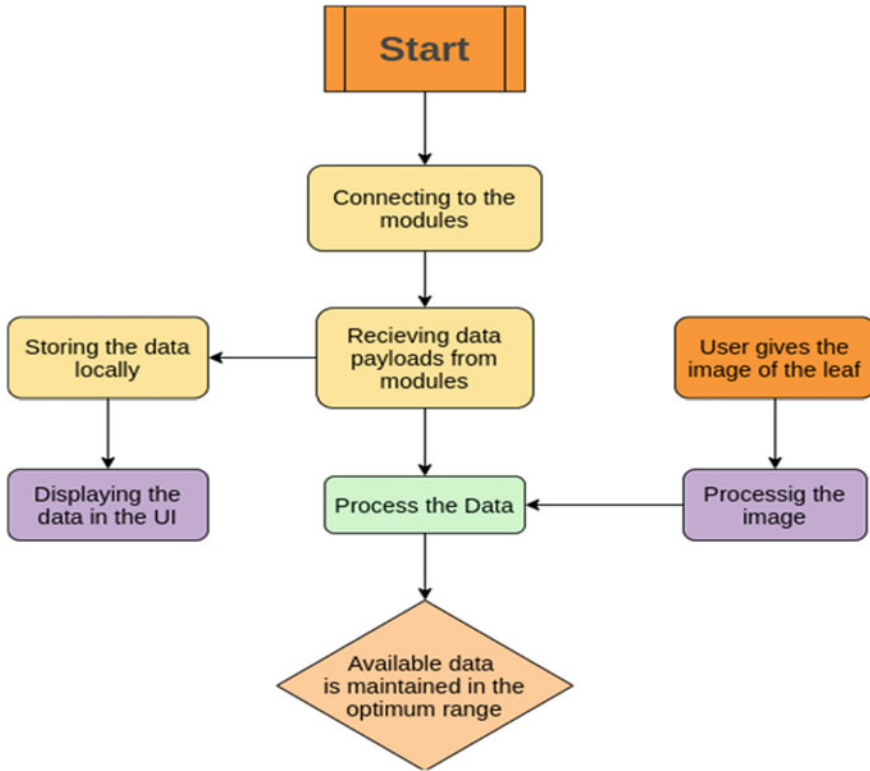


Fig. 4 Flowchart of data process in microcontroller

2.4 Smart Irrigation System

The irrigation system will automatically start if the soil moisture level is less than 50%. The soil moisture sensor senses the quantity of moisture in the soil. A preset range of soil moisture is established, which can be modified based on soil type or crop variety [16]. If the moisture content of the soil deviates from the specified range, the watering system is activated or deactivated. When the system is turned on and off, a message is sent to the user via MQTT protocol, updating the status of the water pump and soil moisture. This activates the irrigation system, pumping water to the plants. A flow of water will be directed to the plant when there is low soil moisture content through drip irrigation. Drip irrigation is the irrigation system employed in this project. The soil moisture sensor senses the moisture level initially, and if the soil is dry, the indication will be sent to the microcontroller [17]. Then the signal will be sent from the microcontroller to the relay module through integration, from where the signal will be amplified and sent to the solenoid valve, then the valve opens to let the water flow. The microcontroller development board's relay board module is utilized to regulate higher current loads. The signals are transmitted

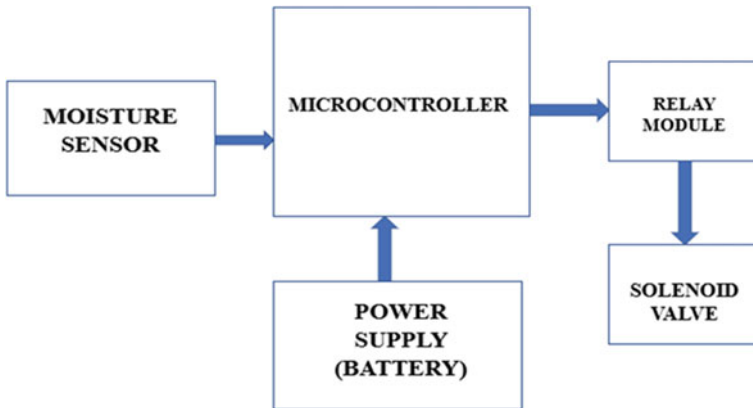


Fig. 5 Block diagram of smart irrigation system

through the MQTT protocol to the central hub. The irrigation system is programmed to deliver the right amount of nutrients and pesticides, as well as water, to maximize production. Smart irrigation system which will help the farmers to make maximum use of water and to reduce excess supply of water to the crops which causes damage to the plants, which will eventually increase yielding of the crops. To treat the disease, fertilizer/pesticide is manually placed into the irrigation system, allowing it to reach the diseased plants. The fertilizers are applied in the proper proportions to guarantee that no fertilizer/pesticide is wasted and that the plant and soil are not harmed. The block diagram of the smart irrigation system is shown in Fig. 5.

3 Results and Discussion

Okra cultivation has been effectively automated with great yield efficiency, decreased time consumption, substantially reduced human involvement, pest control, and other benefits. As a result, raw data from the sensors is acquired. The data is subsequently sent through the MQTT protocol to the reporting system. The IoT devices are used to create an automated irrigation system. Moisture details for the cultivation of crops were obtained from both dry and wet lands. The identification of diseases in the plant is done by image processing. The three most common diseases seen in okra are trained in this project, and a database model is developed. The values from the temperature and soil moisture sensor in dry land are shown in Fig. 6, and values of sensors in moist land are shown in Fig. 7. Figure 8 shows the training and validation accuracy. The training and validation loss is shown in Fig. 9.

Here, a diseased leaf is collected, and the system detects yellow vein mosaic virus infection. Then a damaged leaf is captured, and the system reveals that it is infected with cercospora leaf spots. In another picture, a sick leaf is caught, and

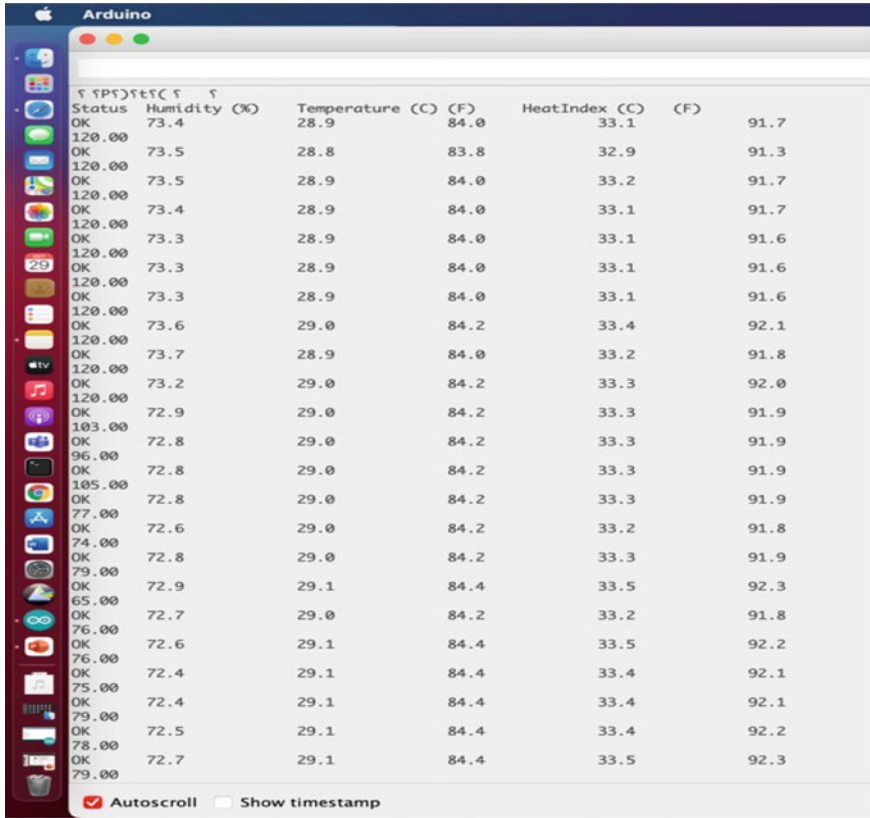


Fig. 6 Values from the temperature and soil moisture sensor in dry land

the system detects powdery mildew infection. A regular okra leaf is also captured, and the system indicates that it's a fresh leaf. In the images below, the accuracy and loss for validation and training are depicted. The detection of disease through image processing is shown below. The image of diseased leaf will be captured using camera and it will be compared with similar diseased images and the percentage of accuracy will be displayed below the picture as shown in Figs. 10, 11, 12, 13 and 14. Figure 10 shows the yellow vein mosaic virus. Figure 11 shows the cercospora leaf spots. The disease named powdery mildew and fresh leaf is shown in Figs. 12 and 13. Sensors value displayed in Node-RED (user interface) is shown in Fig. 14.

The accuracy of image processing in detecting the disease in okra is around 90%. By doing this, labor costs are decreased, production rates are increased, and market value is realized. Fertilizer is fed into the irrigation system in the proper proportions when the disease has been identified. This disease detection aids farmers in determining whether or not their plants are infected with the disease in question. This also aids in the early detection of any diseases.

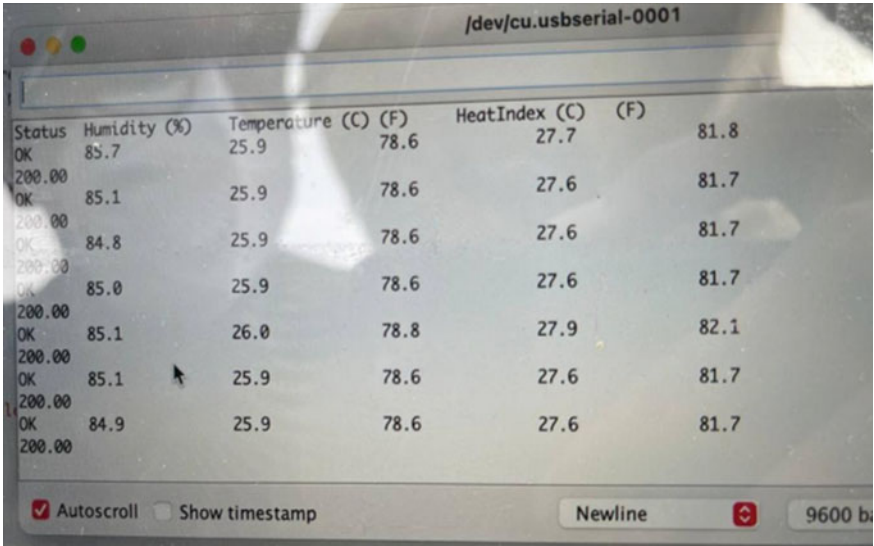


Fig. 7 Values of sensors in moist land

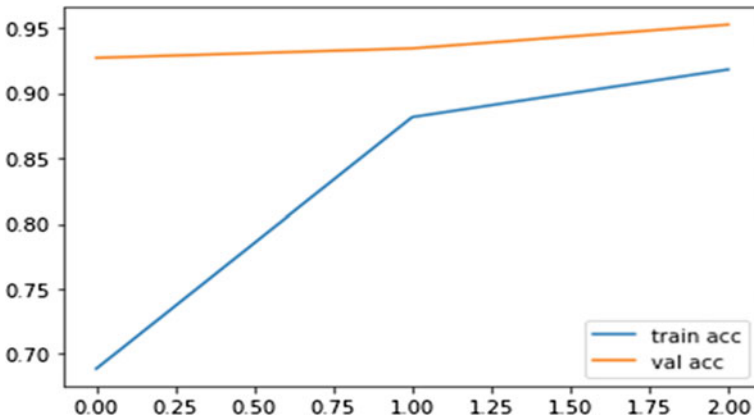


Fig. 8 Training and validation accuracy

4 Conclusion

In general, traditional crop cultivation is a prevalent style of practice, but it also has its own set of inconsistencies. This type of cultivation necessitates a great amount of labor and takes longer. Crop cultivation automation is already underway, but automating the bhindi crop remains a hurdle. The old strategy will be unable to balance demand in the time frame required. As a result, IoT technology in crop cultivation benefits farmers. To top it off, the traditional irrigation system has been

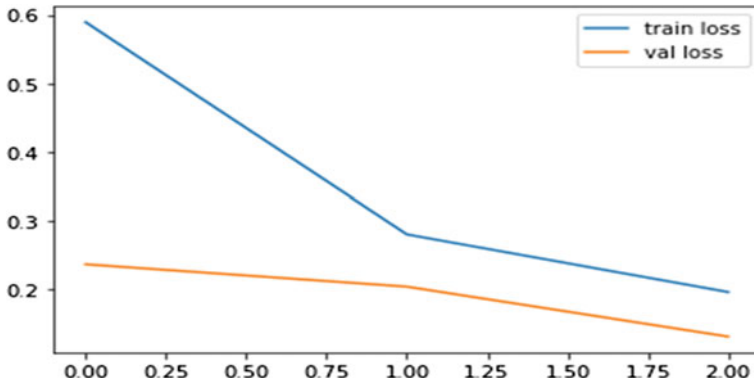


Fig. 9 Training and validation loss

Fig. 10 Yellow vein mosaic virus



Fig. 11 Cercospora leaf spots



Fig. 12 Powdery mildew

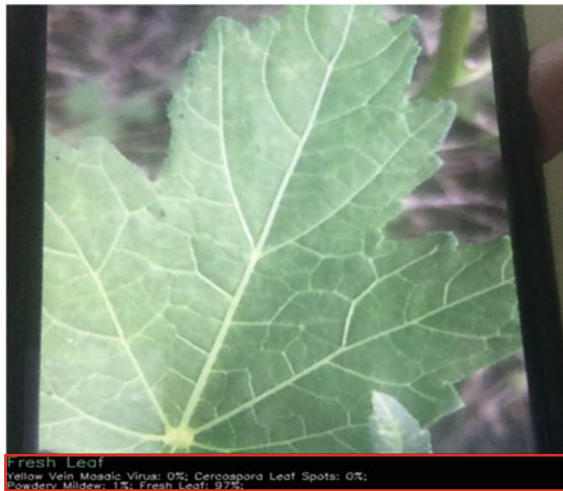


Fig. 13 Fresh leaf

upgraded with technology, making one of a farmer’s most tough chores in crop cultivation easier. Detecting illnesses in the plant is one of the most difficult aspects of okra production. Okra plant diseases come in a variety of forms, and identifying a specific disease is another issue. The Internet of Things (IoT) facilitates the detection procedure. The method of image processing aids in the detection of the type of disease that affects the plant. It’s a time-consuming procedure that requires a large number of photos to be loaded and trained. Image processing, on the other hand, is the most effective way for detecting the disease in the okra plant. When a larger number

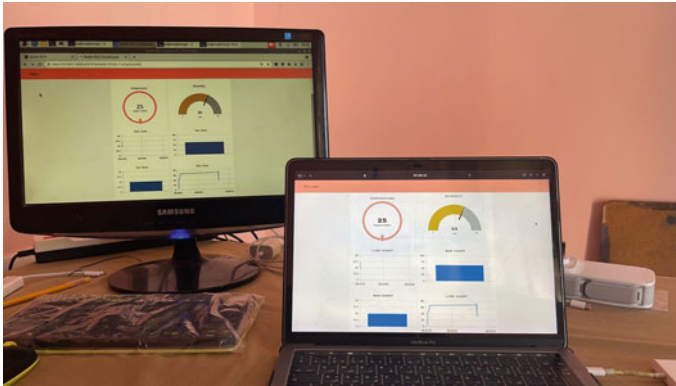


Fig. 14 Sensor's value displayed in Node-RED (user interface)

of images are taught, it improves accuracy and efficiency. In the production of the okra plant, the irrigation system is critical. IoT technology is used to automate the irrigation system. It automatically transmits a signal and irrigates the area when the moisture level in the soil drops.

5 Future Works

In the upcoming years, this same technology can be implemented in large areas of land. The proposed work that has been conducted is up to a limited area of land; in the future it is planned to expand and implement these modules in a wider scale of land, which may result in an increased yield of okra. This similar method can also be used in the farming of various other crops with a slight modification in the module. In the future, the system can be made more efficient by using a moving rover consisting of a camera. This rover moves inside the field and captures the picture of the disease in the plant in a closer range. Through this the disease can be captured with more clarity. By collecting a greater number of images, in the future we can get the cent percent accuracy in the detection of diseases in okra. The image processing can be implemented in detecting the growth of the fruit. It can also be helpful in monitoring the health of the crop.

References

1. Arjun KM (2013) Indian agriculture-status, importance and role in Indian economy. *Int J Agric Food Sci Technol* 4(4):343–346
2. Mateus R (2011) Evaluation of varieties and cultural practices of okra (*Abelmoschus esculentus*) for production in Massachusetts

3. Benchasri S (2012) Okra (*Abelmoschus esculentus* (L.) Moench) as a valuable vegetable of the world. *Ratarstvo i povrtarstvo* 49(1):105–112
4. Kumar V, Ramjan Md, Das T (2019) Cultivation practices of okra
5. Kumar A, Choudhary AK (2014) Scientific cultivation of okra (*Abelmoschus esculentus* L.)
6. Chanchal DK, Alok S, Kumar M, Bijauliya RK, Rashi S, Gupta S (2018) A brief review on *Abelmoschus esculentus* linn. Okra. *Int J Pharm Sci Res* 9(1):58–66
7. Elijah O, Rahman TA, Orikuhmi I, Leow CY, Hindia MN (2018) An overview of internet of things (IoT) and data analytics in agriculture: benefits and challenges. *IEEE Internet Things J* 5(5):3758–3773
8. Tomičić-Pupek K, Pihir I, Furjan MT (2020) The role of perception in the adoption of digital platforms in agriculture. In: 2020 43rd international convention on information, communication and electronic technology (MIPRO). IEEE, pp 1429–1434
9. Dewhurst R, Tian GY (2008) Sensors and sensing systems. *Meas Sci Technol* 19(2):020101
10. Punjabi HC, Agarwal S, Khithani V, Muddaliar V, Vasmatar M (2017) Smart farming using IoT. *Int J Electron Commun Eng Technol* 8(1):58–66
11. Subashini MM, Das S, Heble S, Raj U, Karthik R (2018) Internet of things based wireless plant sensor for smart farming. *Indonesian J Electr Eng Comput Sci* 10(2):456–468
12. Syafarinda Y, Akhadin F, Fitri ZE, Widiawan B, Rosdiana E (2018) Precision agriculture based on wireless sensor network with MQTT protocol. *IOP Conf Ser Earth Environ Sci* 207(1):012059
13. Devi KK, Premkumar J, Kavitha K, Anitha P, Kumar MS, Mahaveerakannan R (2021) A review: smart farming using IoT in the area of crop monitoring. *Ann Rom Soc Cell Biol* 3887–3896
14. Pandurng JA, Lomte SS (2015) Digital image processing applications in agriculture: a survey. *Int J Adv Res* 46
15. Pujari JD, Yakkundimath R, Byadgi AS (2015) Image processing based detection of fungal diseases in plants. *Proc Comput Sci* 46:1802–1808
16. Kansara K, Zaveri V, Shah S, Delwadkar S, Jani K (2015) Sensor based automated irrigation system with IoT: a technical review. *Int J Comput Sci Inf Technol* 6(6):5331–5333
17. Islam MM, Al-Momin M, Tauhid ABM, Hossain MK, Sarker S (2020) IoT based smart irrigation monitoring & controlling system in agriculture
18. Putungan DB, Santos AB, Natividad C, Kim D (2009) Quantitative measurement of discolorations on Okra (*Abelmoschus esculentus*) leaves through digital image processing
19. Jacob IJ, Ebby Darney P (2021) Design of deep learning algorithm for IoT application by image based recognition. *J ISMAC* 3(3):276–290

Identifying Pitfalls and Solutions in Parallelizing Long Short-Term Memory Network on Graphical Processing Unit by Comparing with Tensor Processing Unit Parallelism



Aswathy Ravikumar, Harini Sriraman, S. Lokesh,
and P. Maruthi Sai Saketh

Abstract In deep learning, an important class of algorithms is recurrent neural networks (RNN). RNN experiences high latency and high computational cost, which makes it difficult. RNN is used mainly for sequential data, but the dependencies are complex, making it difficult to parallelize. Resource utilization is inefficient for RNN in GPU compared to TPU. This analysis of RNN for hate speech application is done in GPU and TPU, and the data reuse is identified. The main reasons for the unsatisfactory utilization of GPU for RNN are synchronization overhead and data reuse. In TPU, the global memory accesses and overhead that occur in synchronization are less than GPU. The main objective of this work is to identify the gaps and find solutions for them; analysis is conducted to answer the questions identified regarding the acceleration of RNN in GPU and TPU and to overcome the pitfalls that occur during the parallelizing of RNN in GPU. For implementing the LSTM, we have chosen the hate speech detection application because it is an application that needs to be parallelized for the timely detection of hate speech. This system works as a closed loop, and the attackers are aware of the detection process, so they try all possible ways to evade detection, so timely detection is needed to avoid the problem. The pitfalls and solutions in parallelizing LSTM on GPU are identified.

Keywords Long short-term memory · Acceleration · Tensor processing unit · Graphical processing unit · Speech processing · Neural networks · Recurrent neural networks

A. Ravikumar (✉) · H. Sriraman · S. Lokesh · P. Maruthi Sai Saketh
School of Computer Science and Engineering, VIT, Chennai 600127, India
e-mail: aswathy.ravikumar2019@vitstudent.ac.in

H. Sriraman
e-mail: harini.s@vit.ac.in

1 Introduction

Speech and text processing is an important application area in deep learning, which is overgrowing. In the normal DNN, the inputs and outputs are independent of each other, which is ineffective in processing sequential data applications like the following word prediction. In applications like movie script writing, the prediction of the next word, the memory of the previous work, and events used are needed [1, 2]. For sequential data processing applications, a specialized class of deep learning model called recurrent neural networks (RNN) is used [3]. In RNN, the output depends on the previous computation and is developed based on the human learning capability. RNN is widely used in fields like machine translation speech processing, reading comprehension, and natural language models [4]. In RNN, there are two steps: model weight training and predict the following new result. The initial training phase of RNN is a throughput-oriented task that needs high computation and massive parallelism to improve it. Usage of GPU and TPU helps improve the RNN network's throughput using parallelism. In applications that require faster results like a prediction for the requests made online, hate speech prediction, to impose the goals and overcome social challenges. In such applications, the output needs to be delivered fastly. Otherwise, it affects customer satisfaction and the economy. Many applications in RNN cannot be effectively processed due to the high latency and cost. The study on RNN applications raised many questions like the workload, RNN structure, and use of effective hardware. RNN is a field that has a high scope for research.

In RNN, the computations are dominant from the first iteration, and they have a GEMM structure that must be parallelized for data reuse with low cost and less data movement. The data movement from the L3 cache to the L2 store in GPU is a bottleneck done for minimum data reuse. The parallelizing capability of GEMM is not exploited fully in RNN. In RNN, the computation increases as the sequence length increases and more steps are required for the sequential data processing. It is difficult to parallelize sequential data due to the dependencies across the grades. The GPU is better for RNN than the CPU for large batch sizes but small batch sizes; the GPU is not suitable due to the ideal nature of the many cores. This study discusses the limitations of RNN in the CPU, GPU, and TPU and optimized platform for RNN execution for data reuse. The performance and training time in all the forum are analyzed in this work, and it helps to understand the limitation of LSTM/RNN in GPU compared to TPU. A novel method is required to make effective use of the hardware to accelerate RNN. The evaluation should be done on a standard benchmark real-world application to see if the technique has low latency and high throughput while effectively utilizing the hardware for acceleration. In the modern data center workloads, RNN plays a vital part, and the workload is executed in the TPU of Google.

Social networking is medium for connecting to a group of people and spreading data. The data available in this medium is vast, and many researchers are done on it like sentiment analysis. Sentiment analysis uses data from social media platforms like Twitter [5]. The other fields of research in social media data are detecting negative

comments on social media and product recommendations based on customer interest [6]. Social media is now a platform for sharing celebrities' political and religious views, which often leads to many issues if the opinion does not match with others and leads to the spread of hate speeches and abusive content on the sites. The reports have shown that it can even lead to riots in many cities, leading to the need for timely detection of hate speech on social media. The verbal or nonverbal data shared among individuals and groups with aggression is hate speech [7]. Hate speech can be abusive language in terms of political views, race, sexual content, etc., and this negatively impacts people and society [8]. In this paper, the hate speech is identified using an RNN network, and this acceleration is done using GPU and TPU for timely detection of it. The open-source dataset available is taken for this study, and an LSTM network is used for the prediction. The CNN-based analysis on GPU and TPU was performed earlier [9, 10].

LSTM network is superior to the RNN because of the constant back propagation logic for error correction compared to the RNN. The increased learning of the LSTM enables the user to train models utilizing sequences with hundreds of time steps, which the RNN struggles to perform. Something that wasn't addressed while discussing the gates is that it is their responsibility to determine the relevance of information that is kept in the cell and concealed states so that, when back propagating from cell to cell, the passed error is as near to 1 as feasible. This assures that there is no disappearance or explosion of gradients. For sequential modeling applications such as language modeling and time series forecasting, LSTMs are the most used approach. Such activities often exhibit both long-term and short-term patterns; hence, it is essential to understand both practices for accurate forecasts and estimates. There is an increase in transformers-based approaches that simulate long-term dependencies much better than LSTMs; yet, transformers cannot be employed for all applications because of their data-intensive training and deployment requirements.

The paper is structured as the related study given in Sect. 2. Section 3 is the proposed method with details of RNN, LSTM, implementation in GPU, and TPU. In Sect. 4, the implementation and result details are discussed. In Sect. 5, the future scope is identified, and the paper is concluded.

2 Related Study

Recurrent neural networks are widely used in many sequential data applications, and the training time is a crucial factor that affects the performance of RNN. Facebook used GPU for the training of RNN to have better training time than CPU. An RNN library named DeepCPU was designed to implement RNN effectively and fastly in CPU compared to two GPU implementations. But the GPU implementations need to be considered and developed because most of the datacenters are equipped with GPU [11]. The basic GPU structure is very effective for the GEMM operations in RNN. GPU-based implementations of the RNN network are mainly executed using TensorRT, TensorFlow, and cuDNN [12]. The latency is very high in TensorFlow's

GPU compared to CPU for the exact model sizes. This is mainly due to the repeated loading of model weights in GPU from global memory, which caused low throughput and high latency. The same problem exists in TensorRT, but in CuDNN, the latency is reduced. In cuDNN, the efficiency is reduced, and the GPU parallelism is not effectively utilized.

Research in hate speech detection gained significance in recent years [13] due to the spread of harmful content on social media, causing massive damage to society. Since 2012, there has been a growth in representation learning, the handcrafted feature representation and vocabulary approach is neglected, and the dense representations are used by specialized techniques like paragraph2vec, node2vec, and word2vec. Hate speech detection is handled by researchers in different ways, like natural language processing models using the unsupervised method or traditional machine learning supervised models. In the NLP approach, the data is split as hate or not hate speech. The NLP models are sometimes slower when compared to the ML and DNN models. It is not widely used. In ML and DNN models, the labeled dataset is used, and here, the dataset collected from social media is labeled before the algorithms are applied. Waseem and Hovy [14] analyzed many tweets and the best features that give good performance for hate speech identification and also explained the significance of annotators' knowledge of the hate speech model. The main classification was into binary class, either hate speech or not, but multi-class sort can also be done with a degree of hatred. Mishra et al. [6] use social graph information with graph-based CNN for hate speech detection. Hate speech detection was done for different languages. LSTM networks are also used for speech recognition applications, and sequence labeling is used for the training process in speech recognition. Graves et al. [15] used connectionist temporal classification logic for sequence labeling in RNN, and this method avoids the need for pre-segmentation in the training dataset and output post-processing. He also used LSTM for speech recognition.

2.1 GPU Versus TPU

GPU is efficient for irregular computations, easy to program, and highly flexible for nonMatMul and small-batch computations. GPU is better for medium and small batch sizes. The cost of TPU is high compared to a GPU but the training expenses are highly reduced in TPU and compensate for the programming expenditure. TPU is mainly used for matrix computations due to the systolic array structure and is effective for large batch sizes and complex networks. TPU is better for processing the complex neural network models associated with NLP and image processing applications. The speed of execution and ease of use favor TPU compared to GPU for machine learning applications.

2.2 Scope Identified

From the literature study done, the following gaps were identified, and in this work, the analysis is conducted to find an answer to the questions identified.

- Is GPU suitable for accelerating RNN?
- Is TPU better than GPU for accelerating RNN?
- Identify the pitfalls in parallelizing RNN in GPU
- Suggestions to improve the performance by accelerating RNN

3 Proposed System

From the literature, it was clear the existing research questions were low latency, high throughput, and practical resource utilization implementation possible for RNN in GPU. Can acceleration be achieved in RNN through TPU? Effective data parallelism is potential for RNN in GPU and TPU. For small and large batch sizes. Keeping these questions as the motivation, the RNN network (LSTM) was implemented for a standard hate speech application, and its working and parallelism in GPU and TPU were analyzed.

3.1 RNN

The deep neural networks, which are feed-forward networks, are unidirectional in which the output of the current layer is fed into the next layer in the forward direction. In feed-forward networks, the past information is not stored, but in applications like speech processing, there is a need for previous information and temporal dependencies. RNN is a bidirectional network with loops in the hidden layer, which helps store the last information layer and helps in the future predictions. In RNN, the short-term dependencies are only handled due to the exploding and vanishing gradient problem. RNN is widely used for sequential data and applications like speech processing, machine translation, image captioning, etc. The major drawback faced in RNN is the need for a pre-segmented training dataset and the output to be post-processed to labeled sequences. The connectionist temporal classification method was applied to RNN to overcome this. In RNN, the long-term dependencies were challenging to handle, so the long short-term memory network (LSTM) was introduced. In LSTM, there are memory cells to identify the long-term dependencies. The recursive cells in RNN maintain the previous layer information in the hidden layers [15]. In each step, the input is an element from the sequence given, and the secret state value during the last layer updates the hidden state and finally produces the output. The number of iterations of a cell is determined based on the length of the input. LSTM has special memory cells that store data for the long term. The model was designed for both the

```
Model: "sequential_2"
```

Layer (type)	Output Shape	Param #
embedding_2 (Embedding)	(None, 85, 200)	4793200
lstm_2 (LSTM)	(None, 85, 128)	168448
global_max_pooling1d_2 (Glob	(None, 128)	0
dense_4 (Dense)	(None, 64)	8256
dense_5 (Dense)	(None, 1)	65

```

Total params: 4,969,969
Trainable params: 176,769
Non-trainable params: 4,793,200
    
```

Fig. 1 RNN model for binary classification

```
Model: "sequential_3"
```

Layer (type)	Output Shape	Param #
embedding_3 (Embedding)	(None, 85, 200)	4793200
lstm_3 (LSTM)	(None, 85, 128)	168448
global_max_pooling1d_3 (Glob	(None, 128)	0
dense_6 (Dense)	(None, 64)	8256
dense_7 (Dense)	(None, 3)	195

```

Total params: 4,970,099
Trainable params: 176,899
Non-trainable params: 4,793,200
    
```

Fig. 2 RNN model for multi-classification

multi-class, as shown in Fig. 1, and binary class classification, as shown in Fig. 2, with embedding, LSTM, pooling, and dense layer.

3.2 RNN in GPU

In GPU hardware, there are many streaming multi-processors (SMs), GPU global memory, L2 cache (helps in data reuse), and the network for interconnection which aids in acceleration of the neural network processing. Kernels are the function currently running in GPU, and the thread blocks launched by the GPU driver execute the same kernels. The main bottleneck faced in GPU is the inter-thread communication which depends on the thread hierarchy. Threads that lie in different blocks communicate through the global memory, leading to the large number of

global memory access in GPU, which in turn increases the latency and reduces the throughput of GPU, whereas for threads in the same block, communication is possible through shared memory and in the same wrap through registers. In GPU, thousands of arithmetic processing units are there, which help speed up the summation and multiplication in the RNN. In GPU, massive pipelining and parallelism are possible by merging the multiple data access request. The main drawback of GPU architecture is the von Neumann bottleneck, which occurs due to register access. Direct communication between ALU is not possible. It happens through main memory, which causes high memory and energy and reduces throughput. In the latest architectures of NVIDIA Volta, the shared memory is available in SM, L1, and L2 cache, four partitions, for scheduling which can independently execute many threads. The network designed for RNN was implemented in GPU, and the steps are clearly explained in the figure below. The word embedding, LSTM, pooling, and dense layers are executed in GPU, but there is a delay caused due to the global memory access, and finally, the results are updated to CPU, as shown in Fig. 3.

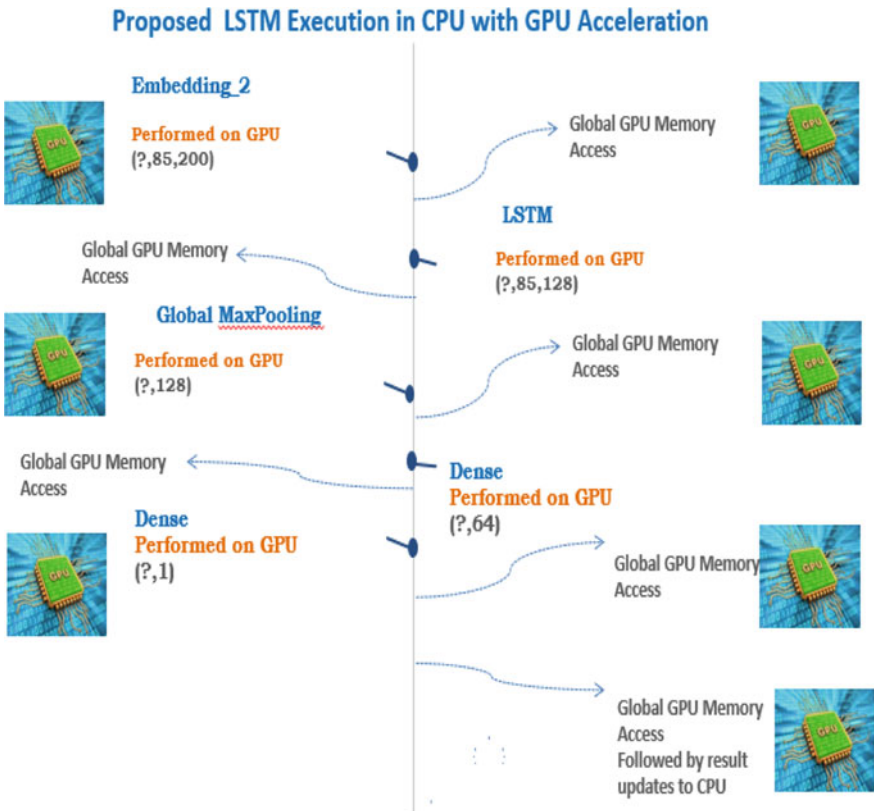


Fig. 3 LSTM execution in GPU acceleration

3.3 RNN in TPU

The TPU is a custom-made ASIC with a matrix processor created specifically for the neural network. TPU efficiently performs multiplication and summation operations in neural networks at a high-speed rate while using little power. Each multiplication is completed and communicated to the next multiplier, while simultaneous addition is conducted. With TPU, memory access is not required for parallel calculations, enabling high computational throughput and a reduction in neural net power consumption. The TPU aids in the acceleration of generic multiplications of floating-point data—GEMM—which is the primary component of CNN [6]. Systolic array in TPU facilitates data reuse, which improves the speed and energy efficiency of LSTM execution. With fewer memory accesses, the systolic arrays aid in accelerating the process. TPU, unlike CPU and GPU, eliminates characteristics not used by the neural network, hence conserving energy. In TPU, embeddings, LSTM, pooling, and dense layers are processed, making the execution quicker than GPU, as seen in Fig. 4. In TPU, the parallel memory access is done in the local memory to threads, and the global memory access is limited.

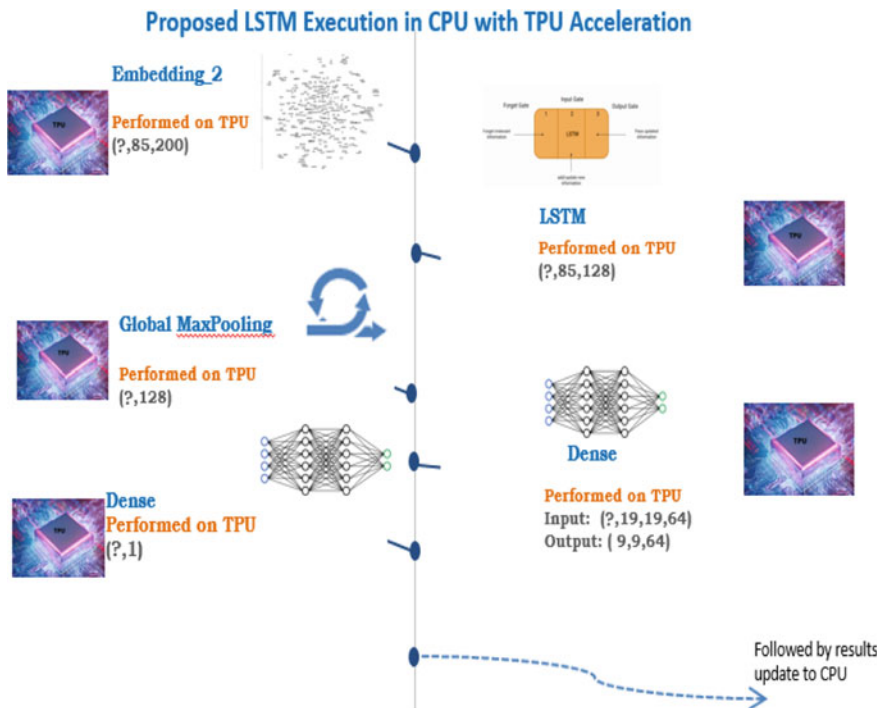


Fig. 4 LSTM execution in TPU acceleration

4 Implementation Details

The hate speech dataset used 21,000 sentences, of which 13,400 sentences were used for training, 6600 for validation, and 1000 for testing. This hate speech dataset contains three classes: hate speech, offensive language, and none. We combined hate speech and offensive speech as one class and second class as non-hate speech. The number of people using the Internet is getting more every day. There are many applications where people can share their opinions and views, and this next generation is getting huge. One side of these comments can be motivating, educative, and helpful, and the other side is spreading hate among people. The Internet should be the one that helps people motivate and grow. But there are many cases where people get bullied and trolled about their culture, gender, race, or disability which causes hate and anger. It is difficult to inspect all the comments, whether they are toxic. The proposed model uses the word embeddings, the long short-term memory network (LSTM) used to improve classification efficiency, followed by the max-pooling layer and two dense layers.

Implementation in the CPU platform is done in CPU in Google Colab Intel® Xeon®, 2.30 GHz CPU frequency, 2 CPU core, 12 GB RAM, and 25 GB disk space. The parameters are loaded from storage into matrix adders and matrix multipliers inside the TPU. While multiplication is being performed, the product of each multiplier is sent to the subsequent multiplier. The outcome would be the total of all parameters multiplied by values. Because memory access is not needed during this massive computations and data transmission process, TPUs may achieve outstanding computational throughput on neural networks. Because the Colab TPU has eight cores, the training data is dispersed over all eight cores, which improves the performance.

5 Result Analysis

The testing accuracy for each application was compared for both GPU and TPU for batch sizes 16, 32, 64, and 128. The analysis was done for both the binary and multi-class classification of the hate speech application. The layer-wise timing analysis uses the just-in-time (JIT) compiler. Just-in-time compilation helps to improve the performance of the integrated programs. In the execution phase, the program is compiled as native code for performance improvement and is known as dynamic compilation. The `timeit` function in python measures the time of the code snippets and each layer in the model.

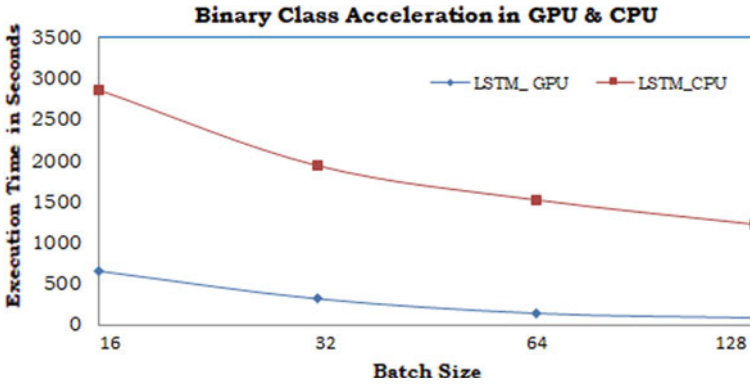


Fig. 5 Acceleration of binary class model in CPU and GPU

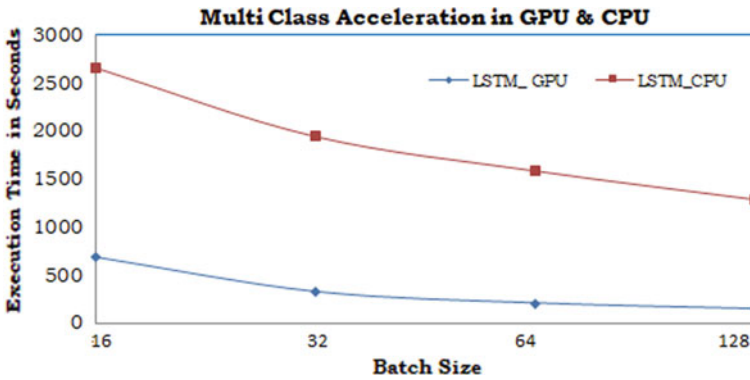


Fig. 6 Acceleration of multi-class model in CPU and GPU

5.1 Is GPU Suitable for Accelerating RNN?

Yes, GPU is suitable for the acceleration of RNN compared to CPU in terms of execution time and parallelism. The analysis was done for the multi-class and binary class models in GPU and CPU. From Figs. 5, 6, 7, 8, 9 and 10, it is obvious that GPU performs well for RNN compared to CPU.

5.2 Is TPU Better Than GPU for Accelerating RNN?

Yes, TPU is better compared to GPU for accelerating RNN. The TPU gives better accuracy and faster execution time.

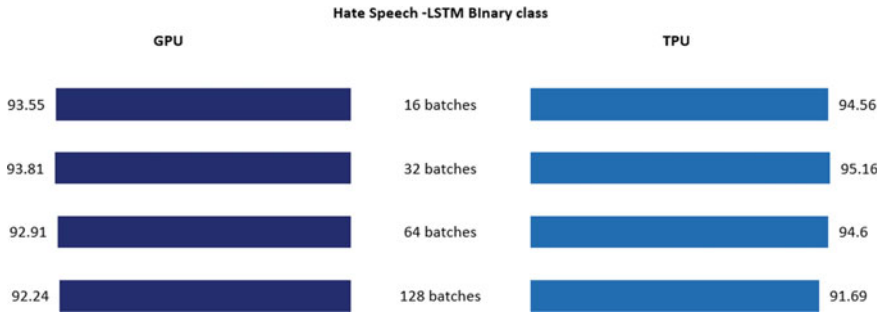


Fig. 7 Accuracy of the binary class model in TPU and GPU

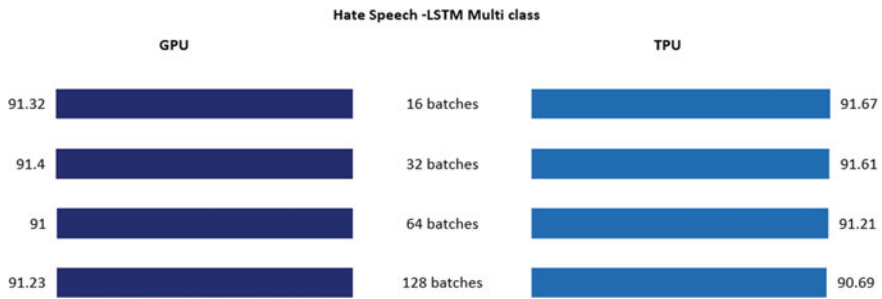


Fig. 8 Accuracy of multi-class model in TPU and GPU

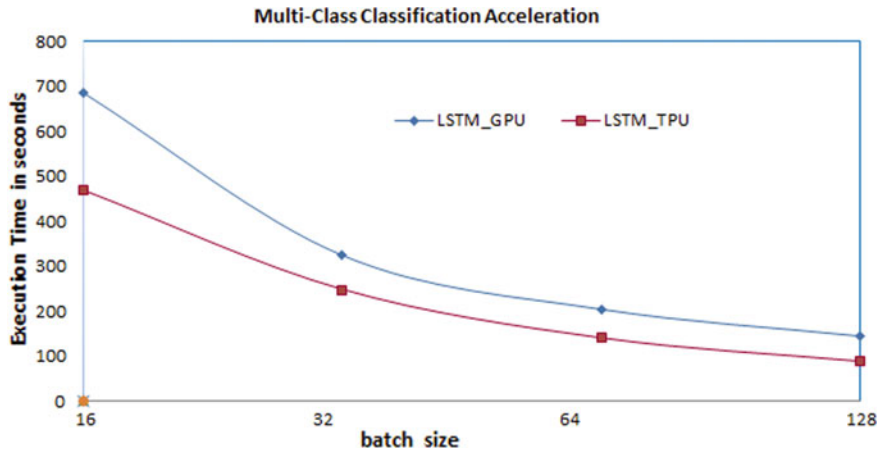


Fig. 9 Acceleration of multi-class model in TPU and GPU

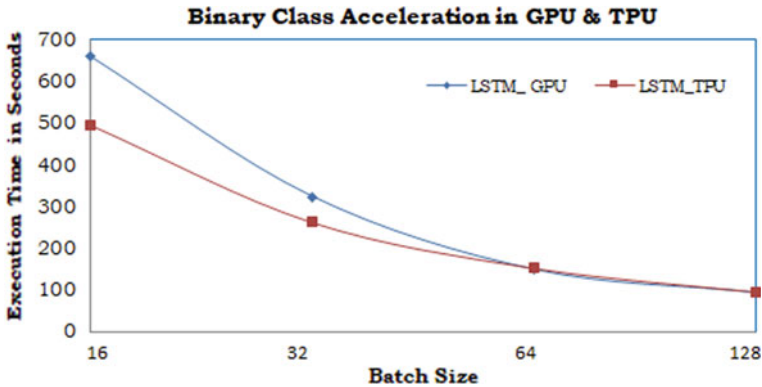


Fig. 10 Acceleration of the binary class model in TPU and GPU

5.3 Identify the Pitfalls in Parallelizing RNN in GPU

The main pitfall identified in the parallelizing of RNN in GPU is the global memory access in the GPU. The global memory access causes a bottleneck in the execution of RNN in GPU. The other pitfall identified is the global max pooling applied in the model for the reduction of the size is creating a delay in the execution process in TPU. The max-pooling layer was avoided in TPU execution, giving better performance, and the execution time is at par with the max-pooling layer. So, for TPU, the max-pooling layer can be avoided to obtain better accuracy and remove the bottleneck. But in GPU, the removal of max-pooling layers leads to a decrease in the performance, and the execution time increases, which shows down sampling is necessary for GPU. The max-pooling can be replaced with the multiple local kernels in GPU, and the kernels communicate with each other to update the global result. The impact max-pooling layer in the model was analyzed for the CPU, GPU, and TPU.

5.4 Suggestions to Improve the Performance by Accelerating RNN Through GPU

Suggestion 1: Optimized memory movement—unified memory access, one of the latest technologies, can be utilized, and the memory movement can be reduced among the dataset.

Suggestion 2: Streaming process—using the latest streaming options within the kernel operations and memory movement. The data movement and the process can be executed in parallel.

Suggestion 3: Utilization of dynamic parallelism—using GPU with the kernel-to-kernel call can be executed without going to the CPU. It was using more vectorized functions in the execution in GPU. The task-based analysis was done for each layer,

Table 1 Task timings

Layer details	Task ratio	Time ratio	Number of tasks	Time per layer	Time per task in the layer
Layer 3—LSTM	0.610451	0.165133172	257	34.1	0.132684825
Layer 4—max pooling	0.07601	0.169491525	32	35	1.09375
Layer 5—dense	0.306413	0.167070218	129	34.5	0.26744186
Layer 6—dense	0.007126	0.165617433	3	34.2	11.4

and the time for each layer was analyzed with its impact on each layer. The tasks for each layer were identified.

- For neural network layer, the tasks are identified using the basic $y = \text{summation}(wix + b)$ equation, so it has multilocation and addition tasks at each node. Each node in the neural network layer has two tasks and, finally, an overall summation task, which can be generalized as the number of tasks $= m * 2 + 1$, where m is the number of nodes in the neural network layer.
- LSTM layer of the network has 128 nodes, so tasks $= (128 * 2) + 1 = 257$ tasks.
- Max-pooling layer, the number of nodes is 128, and the kernel size is $2 * 2$, tasks $= 128/4 = 32$ tasks. It can be generalized as the number of tasks $= m/n$, where m is the number of neurons and n is the kernel size.
- The layer 5: dense layer with 64 neurons, the number of tasks $= 64 * 2 + 1 = 129$ tasks.
- The layer 6: dense layer with 1 neuron, the number of tasks $= 1 * 2 + 1 = 3$ tasks.

The time for each layer based on the task was analyzed for the existing system and shows the task ratio, time ratio, time per layer, and time per task in the layer.

Table 1 shows the task timings. From the analysis, it was clear that any layer with the least number of tasks does not benefit from the conventional existing sequential execution, and the penalty caused to the memory is more. This can be effectively solved using dynamic parallelism. The impact of dynamic parallelism is explained theoretically using Amdahl's law. This law is used to give the theoretical speedup in latency of the task executed for a workload concerning improved resources. The total time taken for each task without parallelism is 137.8 ns. The total time calculated with dynamic parallelism is calculated theoretically and is 55.86031128 ns. The overall speedup can be calculated using Amdahl's law is 2.5 times.

6 Conclusion

In this paper, the recurrent neural networks were implemented in GPU and TPU, and the reasons for the poor performance in GPU were explained. The low data reuse

and the bottleneck in GPU were defined. The main reasons for the unsatisfactory utilization of GPU for RNN are synchronization overhead and data reuse. In TPU, the global memory accesses and overhead occurred in synchronization are less than GPU. The bidirectional long short-term memory RNN was studied in this and implemented in the hate speech application. The acceleration of LSTM is achieved using the TPU. From the analysis, it is clear that TPU has faster execution compared to GPU for both binary and multi-class classification. The pitfalls in RNN for GPU are identified, and suggestions are given. The common pitfalls are identified, and the solutions for the same are given with detailed analysis.

References

1. Ravikumar A, Saritha R, Chandra V (2013) Recent trends in computational prediction of renal transplantation outcomes. *Int J Comp Appl* 63:33–7
2. Robin M, John J, Ravikumar A (2021) Breast Tumor Segmentation using U-NET. In: 2021 5th international conference on computing methodologies and communication (ICCMC) pp 1164–1167. <https://doi.org/10.1109/ICCMC51019.2021.9418447>
3. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
4. Bengio Y, Ducharme R, Vincent P, Janvin C (2003) A neural probabilistic language model. *J Mach Learn Res* 3:1137–1155
5. Jemai F, Hayouni M, Baccar S (2021) Sentiment analysis using machine learning algorithms 12(9):775–779. <https://doi.org/10.1109/iwcmc51323.2021.9498965>
6. Mishra P, Del Tredici M, Yannakoudakis H, Shutova E (2019) Abusive language detection with graph convolutional networks. In: NAACL HLT 2019—Proceedings of the 2019 annual conference of the North American chapter of the association for computational linguistics: human language technologies (NAACL-HLT), vol 1, pp 2145–2150. <https://doi.org/10.18653/v1/N19-1221>
7. Sreelakshmi K, Premjith B, Soman KP (2020) Detection of hate speech text in Hindi-English code-mixed data. *Proc Comput Sci* 171:737–744
8. Watanabe H, Bouazizi M, Ohtsuki T (2018) Hate speech on twitter: a pragmatic approach to collect hateful and offensive expressions and perform hate speech detection. *IEEE Access* 6:13825–13835. <https://doi.org/10.1109/ACCESS.2018.2806394>
9. Ravikumar A, Sriraman H, Sai Saketh PM, Lokesh S, Karanam A (2022) Effect of neural network structure in accelerating performance and accuracy of a convolutional neural network with GPU/TPU for image analytics. *PeerJ Comput Sci* 8:e909. <https://doi.org/10.7717/peerj-cs.909>
10. Ravikumar A (2021) Non-relational multi-level caching for mitigation of staleness & stragglers in distributed deep learning. In: Proceedings of the 22nd international middleware conference: Doctoral symposium (Middleware '21). Association for computing machinery, New York, NY, USA, 15–16
11. Zhang M, Rajbhandari S, Wang W, He Y (2018) DeepCPU: serving RNN-based deep learning models 10x faster. In: 2018 {USENIX} annual technical conference ({USENIX} {ATC} 18), pp 951–965. [Online]. Available: <https://www.usenix.org/conference/atc18/presentation/zhang-minjia>
12. Abadi M et al (2016) TensorFlow: a system for large-scale machine learning. In: 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16), pp 265–283. [Online]. Available: <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi>

13. Kumar R, Ojha AK, Malmasi S, Zampieri M (2018) Benchmarking aggression identification in social media. TRAC 1:1–11
14. Waseem Z, Hovy D (2016) Hateful symbols or hateful people? Predictive features for hate speech detection on twitter 88–93. <https://doi.org/10.18653/v1/n16-2013>
15. Graves A, Mohamed A, Hinton G (2013) Speech recognition with deep recurrent neural networks 3
16. Chetlur S et al (2014) cuDNN: efficient primitives for deep learning, pp 1–9. [Online]. Available: <http://arxiv.org/abs/1410.0759>

Supply Chain Management in Blockchain Using Hyperledger



Arun Kumar Kushwaha, Soumya Singh, Anshay Rastogi, Vipin Deval,
Pravesh Singh, and Vikas Goel

Abstract A track of the products is kept from producer to consumer as a test of combining the digital and physical worlds. The final consumer has access to a comprehensive record of information and can be confident that the information is correct and precise. Because it employs a distributed public book of account, one of the emerging technologies, blockchain could be a useful approach in handling supply chain management. As a result, building supply chain on blockchain helps improve overall supply chain management as it eliminates counterfeit products, improves visibility and transparency as well in the supply chain. The following qualities of blockchain technology make it useful in the supply chain sector: it lowers errors, avoids product delays, eliminates fraudulent activities, improves management, promotes consumer/supplier confidence, and so on. As a result, we use blockchain to deploy supply chains to improve overall supply chain management. Blockchain records information on a wide range of products and/or service transactions, which are tracked in real time to address a lack of transparency.

Keywords Supply chain · Blockchain · Hyperledger fabric · Hyperledger composer · Hyperledger playground

A. K. Kushwaha (✉) · S. Singh · A. Rastogi · V. Deval · P. Singh · V. Goel
KIET Group of Institutions, Delhi NCR, Ghaziabad, Uttar Pradesh 201206, India
e-mail: arun.1822cs1043@kiet.edu

S. Singh
e-mail: soumya.1822ec1157@kiet.edu

A. Rastogi
e-mail: anshay.1822it1027@kiet.edu

V. Deval
e-mail: vipin.deval@kiet.edu

P. Singh
e-mail: pravesh.singh@kiet.edu

V. Goel
e-mail: vikas.goel@kiet.edu

1 Introduction

Whenever someone refers to supply chains, they are basically bearing on a network of different parties through which a merchandise or product goes through, in order to reach the consumer from the site of its manufacturing. Most of the supply chains are developed to scale back the cost, time and also remain competitive in the market. These supply chains may be for any reasonable level. You might be working with car manufacturers where the mechanical system is functioning as a supply chain. It could even be related to a quick food joint where something like McDonald's may be working as a supply chain. It could even be associated with grocery markets where the products are being transported from warehouses to retail houses. Some of the major industries which use supply chain are the car manufacturers, the food joints like the Walmart, Reliance; the pharmaceutical industry, the automobile industry, jewelery, brick and cement manufacturers, and so on.

Now, the present supply chain process that provides the solution from the manufacturer's end to the consumer's end has a lot of problems. One of the major problems you will be able to learn through an example is what Walmart faced back in 2018. They came across some disease in a number of batches of spinach but they could not determine their origin. They wanted to track back only that batch which had the disease but they did not have a soundproof system. The only way was to recall all the spinach back from their retail houses. Tracking and recalling spinach spread across the country was going to cost them a lot.

Now, such kind of problems can easily be avoided if you are using blockchain. Walmart concluded the same when they conducted a proof of concept and they figured that by using blockchain they could have reduced the tracking time from 21 d to 2.1 s. Some other problems with the present supply chains are counterfeit products. This is a very serious problem in the drugs [9] and automobile industries [31]. There are a lot of counterfeit products in the market, and present supply chains are not able to track them. Then, there also exist problems like less visibility and less transparency which means people are not able to see what is happening within the system, how the whole communication is carried out, etc. This creates a huge backlog because now one cannot trust the system as there is no credibility maintained. Also, the traditional process of supply chain involves a lot of paperwork and administrative costs which again hampers the purpose of it all. Now, let us have a look at what blockchain powered supply chain brings to the table.

One of the primary benefits of using blockchain would be increased traceability [4, 16]. It would be very easy to track the entire course of an item from source to destination because audit trails are maintained within blockchain itself. Blockchain dictates the whole history of transactions so it becomes very easy to travel back and cross-check the whole transaction. One is also able to remove the counterfeit trading because now if everything is put in blockchain then nothing can be edited or updated. The products supplied by the participants of the supply chain will be the sole products available in the market. Nobody can add in their products, nobody can remove the products from the supply chain. This also improves the visibil-

ity as well as the credibility to the system [8]. Blockchain provides immutability to the data stored on supply chain. It also reduces paperwork and administrative costs because most of the things are now available on computers. Finally, this helps to engage stakeholders because providing an emerging technology like blockchain as a solution for your supply chain would be a very good standing in the market and state to your customers that you are on par with emerging technologies.

The rest of the paper is divided accordingly. Section 2 talks about the related work done by others in this field. Section 3 describes the preliminary steps that need to be taken before one starts building the supply chain management system, and also the objectives of this study. Section 4 contains the proposed solution—how we are going to implement the project and its various other aspects. Section 5 discusses the end result—the working application, and finally, Sect. 6 concludes this study.

2 Related Work

Public blockchain, private blockchain, and permissioned blockchain are the three types of blockchain available [23]. While public blockchains are decentralized peer-to-peer networks, private blockchains have a centralized authority that controls the ledger: the main distinction is the level of access that users have. Permissioned blockchains, on the other hand, run a blockchain among a set of identified and known participants. In almost all the cases, it is the private blockchain that is used to implement supply chain management systems. This is because in order for members of a supply chain to be able to determine the source and quality of their merchandise, each unit must be tightly coupled with the identity of its specific owner at every step along the way.

Research done on the application of blockchain to supply chain management and traceability intended to showcase the benefits that this technology could provide. The authors introduced a blockchain-based framework in [18], which uses an unclonable ID created from an SRAM-based PUF to ensure the validity of electronics. Using blockchain to improve and secure the integrity of electronic supply chains, Xu et al. [32] gave a complete solution and summary. These two systems, nevertheless, do not provide detailed device tracing and ownership data. Feng [29] introduced a supply chain traceability system in 2016 that integrated blockchain and RFID technologies to identify, track, and monitor the whole supply chain on a transparent and traceable platform accessible to all system members.

Quite a lot of efforts have been made to modernize the pharmaceutical supply chains using blockchain-based solutions. Stafford and Treiblmaier [28] have stressed the importance of securely storing patients' medical and health records by suggesting a blockchain-based network to do so. The use of the G-Coin blockchain network for drug supply chains has been proposed by Tseng et al. [30]. This is a digital gold currency based on blockchain technology that is resistant to “double-spend”

attacks in particular. The paper “Blockchain ready manufacturing supply chain using distributed ledger” [1] describes how blockchain can be used to build a worldwide supply network.

We will be using hyperledger technology in order to implement blockchain-based supply chain management systems. Hyperledger Fabric, an open-source [3] blockchain platform, is one of Hyperledger’s [10] initiatives under the umbrella of Linux Foundation. You can record transactions in the hyperledger, but they cannot be altered or erased without the approval of each peer [26]. As a result, the supply chain benefits from a more secure and reliable service, as well as greater data efficiency [15].

3 Prerequisites and Objectives

Hyperledger composer is an application development framework which simplifies and expedites the creation of smart contracts and Hyperledger fabric blockchain applications. It allows you to model your business network and integrate existing systems and data with your blockchain applications. By using Hyperledger composer, you do not require as much experience in development as in other platforms. It offers its own set of file structures and languages, which are much easier to learn and grasp compared others. It allows you to create your business model and integrate with the existing systems present on top of your blockchain. Composer does not allow you to create a blockchain by itself, it is not a standalone tool. It is more like a predefined blockchain where you will be creating business applications as per your use cases. Also, it makes sure that you do not require extensive programming language experience in Node.js or Java as it has its own structures and templates that you can directly use to create business networks and smart contracts on top of Hyperledger Fabric blockchain.

Now that we have gone through the introduction of Hyperledger, we are going to list out the steps which we have to perform to complete our blockchain network with Hyperledger Composer. So to start off with, we create a business network structure using the Command Line Interface (CLI) on a Linux-based system. This business network structure will include all the different files which we are supposed to create. This will include model file, transaction processor file, permission files, and query file. After that we are going to define the business network. This definition is regarding the modeling language which we are going to provide, the business logic which we are going to execute, and the permissions which we are going to provide with permission files.

Once the business network is defined, we are going to generate the business network archive file using the Composer CLI. This business network archive file will have the syntax of BNA which can be deployed on top of the run time environment. Then, we are going to deploy the business network over the run time environment. Here, we are going to use Hyperledger Fabric and you are going to learn how we are going to deploy a business network archive file on top of Hyperledger fabric. Once we have deployed that we are going to test out our network and run a Rest

server for business network to generate the APIs. These APIs can be utilized with our front-end application. Finally, we are going to generate a front-end application from the Composer CLI and there we can execute different transactions and test out our business network.

The objective of this study is to develop and deploy a working model of supply chain management in blockchain. Anyone would be able to access the model via its public URL from a browser and add themselves as a participant. They could then track their product in the simulated supply chain in real time. Besides enhanced traceability, other objectives this model fulfills are lowering losses that result from counterfeit trading, improving visibility over manufacturing of products, reducing paperwork and administrative costs. It also improves credibility and trust in the system thanks to its decentralized nature.

4 Proposed Solution

Figure 1 shows the detailed structure of our model. The participants and assets are defined separately to build a business network over hyperledger composer. Hyperledger Composer includes an object oriented language which is used to define the domain model for a business network (similar to a database model) where different participants, assets, and transactions will be defined. This language provides a primary key in order to identify each participant and asset independently. This network is then run on hyperledger fabric to create a rest server and a front-end application which can easily be used by laymen. Let us now have a closer look at the various aspects of the algorithm used for defining this model.

4.1 Defining Participants

For any standard supply chain network, we are going to have the following participants:

- (a) supplier
- (b) manufacturer
- (c) distributor
- (d) retailer
- (e) customer.

Each of these participants is going to have a company name and an address. Using the concept of inheritance, we can create an abstract participant named “Trader” for example and define all the attributes related to it, like company name and address. The participants supplier, manufacturer, distributor, retailer, and customer will then extend this abstract participant “Trader” and inherit all its properties. Each participant will further have a primary key in order to uniquely identify it.

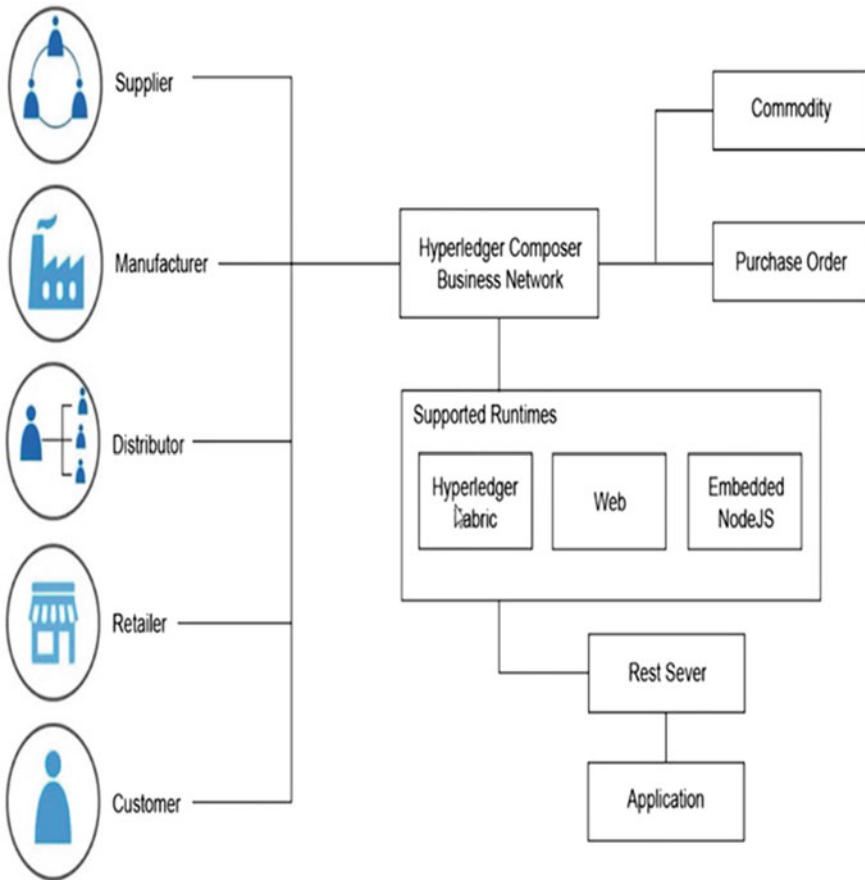


Fig. 1 Proposed model

4.2 Defining Assets

Assets are the commodities which are going to be traded by these participants. As we are dealing with a supply chain, we will have only two assets:

- (a) commodity which is going to be traded
- (b) purchase order which the participant needs to raise for trading commodities.

Each asset will also have a separate primary key. A commodity can have attributes like name, description, quantity, price, etc. Every commodity is going to be linked with a purchase order on top of which the commodity will be traded. We also need to state who is the current owner as well as the issuer of a given commodity. A single purchase order can be made for a variety of commodities, i.e., there is no need to

make separate purchase orders for every commodity. We can define the order status of an asset as:

- (a) initiated
- (b) confirmed
- (c) delivering
- (d) delivered.

Another functionality that can be added here is that of tracking the assets, so we know the timestamp and location of different commodities at different intervals.

4.3 Defining Transactions

Transactions will be executed by the participants in regard to the assets. A transaction can be used to initiate a purchase order, transfer a commodity, etc., by a participant. These transactions help us to manage our blockchain network. Now, we will go on and define the two transactions mentioned above.

4.4 Defining Initiate Purchase Transaction

We will create a function inside which the initiate purchase order will be defined. This transaction is supposed to create a new asset that is going to be a purchase order. New resources will be created using factory API. The vendor in this transaction will be the participant who is calling this function(that is he is the one creating this purchase order) and the orderer will be the participant on whom this function is called upon. Once this new asset is created, we need to store it in our asset registry.

4.5 Defining Transfer Commodity Transaction

Now, we will create another function inside which the transfer of commodity transaction will be defined. Its purpose is to track the trade of a commodity from one participant to another. Consider the following scenario: the supplier provides raw materials to the manufacturer who then manufactures finished products and sells them to the distributor. The distributor in turn hands these products to many retailers from where the end-user can eventually buy them. Every participant in this scenario maintains a record (or receipt) of what was bought sold and by whom. The raw materials and finished products are the commodities here whereas the receipts are the purchase orders. The business logic behind this will be to receive a commodity inside this transaction and assign a new owner to it. The participant performing the transfer of commodity will be the one to call this transaction function. We have to

make sure that only the participant who is issuer of the commodity is able to perform the transaction. The tracking and ownership details of the commodity will have to be updated as per the current status. All the changes will then be saved in the asset registry.

4.6 Defining Access Control for Participants

Here, we define the rules and regulations for transaction assets and admin inside the business network in a permissions file made available to us by the hyperledger modeling language. A participant will have reading access to commodities owned by him. Vendors and orderers should also have reading rights for their respective purchase orders. All participants can access the transfer commodity transaction. A participant can transfer only those commodities which he owns. Furthermore, he will have reading access to only his transactions. Participants will be allowed to create their own commodities (if they are owners) as well as their own purchase orders (if they are orderers). Vendors and orderers can read their respective transaction records.

4.7 System Access Control Definition

Now that we have defined what the access control rules for the admin will be like, here we define what kind of feature sets or access control rules other participants will have in the network. We will create a system access control rule which grants all permissions and access to each and every participant, so that they can log into the business network and perform transactions. Restrictions to these rules are already defined in access control rules for participants in the subsection above, based on the type of participant.

4.8 Defining Queries for Participants and Assets

Now here will define the queries for our business network in order to retrieve information about participants and assets. A typical query inside the query file in hyperledger composer follows JSON format. Every query will have two attributes: description and statement. Description attribute elaborates what the query is doing, whereas the statement attribute defines the rules of the query and can have many operators associated with it (quite similar to SQL).

5 Result

After deploying the business network, it is time to go into generating a Rest server. Hyperledger Composer offers you a Rest API which can be generated for your business network. When you paste the URL in a browser you will see the Hyperledger Composer reservoir will start and generate the API as shown in Fig. 2. You can browse the API and use the get, post, head, put, delete, etc., methods. So in this way, you will be able to interact with the business network. You will be able to create new commodity, customers, and distributor. You can also call in the transactions using the API. You can initiate a purchase order and pass in the information to generate a

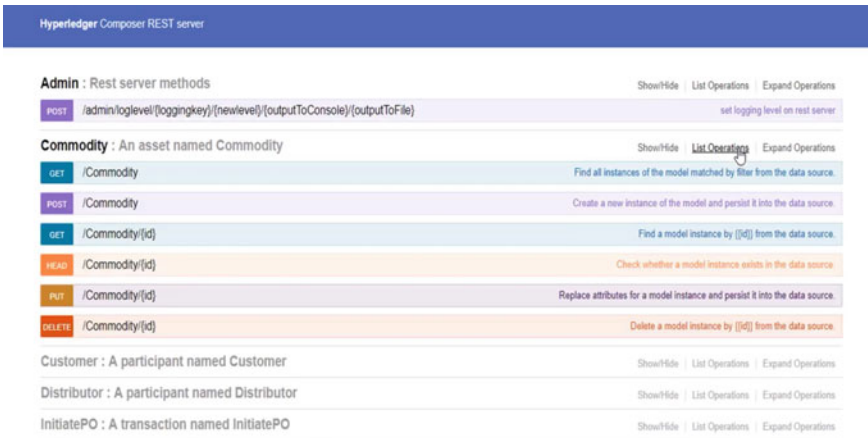


Fig. 2 Rest API

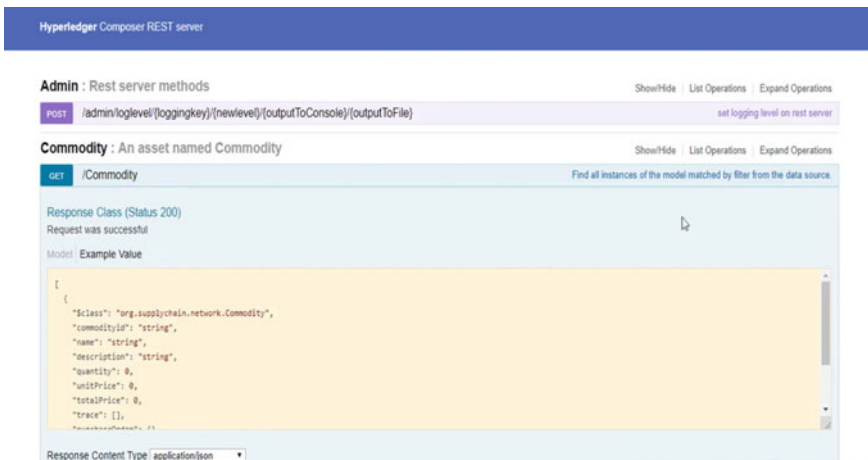


Fig. 3 Example data



Fig. 4 Front-end application

transfer commodity. All these different operations have been listed and are available as APIs to use. These APIs can be utilized with the front-end application to execute different functionalities with the business network. All the different operations related to different types of assets and participants are listed. Moreover, if you expand operations you will see a query field available for all the APIs. You can try these APIs out at your end, try to submit a new value, and create a new community or distributor or a participant. To do that you can use example data as shown in Fig. 3, modify the different contents and finally click on “try it out” to try out the API’s. Make sure you have the correct JSON response while you are trying out the API’s. The final front-end application will look like Fig. 4.

6 Conclusion

The primary goal this study achieves is increasing traceability in supply chains. Blockchain validates data sharing without the need to change the systems that each entity has and thus enabling greater traceability of products across multiple partners, locations, and facilities. Each stakeholder is able to view the same data on a product’s life cycle. Since every product is traceable and the system transparent, duplicate products do not enter the supply chain which significantly reduces the losses incurred due to them. Moreover, one-up/one-down visibility can limit supply chain networks. Blockchain-based supply chain solutions offer authorized players with improved visibility across all supply chain activities thanks to distributed ledger technology that gives a shared, single version of the truth. There is no central authority or any other third party involved in this supply chain; hence, no unnecessary paperwork nor

any extra cost is involved. Ultimately, this model aids participants in keeping track of price, date, location, quality, certification, and other pertinent data in order to better manage the supply chain.

The proposed model has several merits and virtues over other models. It provides immutability as the data is stored on blockchain. You will be able to track the commodities because the model maintains audit trails within the blockchain itself. It maintains the identities of participants within the supply chain. Finally, it provides built-in security throughout the system as a characteristic feature of Hyperledger.

References

1. Abeyratne SA, Monfared RP (2016) Blockchain ready manufacturing supply chain using distributed ledger. *Int J Res Eng Technol* 5(9):1–10
2. Agrawal D, Minocha S, Namasudra S, Gandomi AH (2022) A robust drug recall supply chain management system using hyperledger blockchain ecosystem. *Comput Biol Med* 140:105100. ISSN 0010-4825. <https://doi.org/10.1016/j.compbiomed.2021.105100>
3. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, Enyeart D, Ferris C, Laventman G, Manevich Y, Muralidharan S (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains. In *Proceedings of the thirteenth EuroSys conference*, pp 1–15
4. Aung MM, Chang YS (2014) Traceability in a food supply chain: safety and quality perspectives. *Food Control* 39:172–184
5. Baralla G, Ibba S, Marchesi M, Tonelli R, Missineo S (2018) A blockchain based system to ensure transparency and reliability in food supply chain. In: *European conference on parallel processing*. Springer, pp 379–391
6. Baralla G, Pinna A, Corrias G (2019) Ensure traceability in European food supply chain by using a blockchain system. In: *2019 IEEE/ACM 2nd international workshop on emerging trends in software engineering for blockchain (WETSEB)*, pp 40–47. <https://doi.org/10.1109/WETSEB.2019.00012>
7. Bateman AH (2015) Tracking the value of traceability. *Supply Chain Manag Rev* 9:8–10
8. Bhalerao S, Agarwal S, Borkar S, Anekar S, Kulkarni N, Bhagwat S (2019) Supply chain management using blockchain. 456–459. <https://doi.org/10.1109/ISS1.2019.8908031>
9. Bryatov SR, Borodinov AA (2019) Blockchain technology in the pharmaceutical supply chain: researching a business model based on hyperledger fabric. In: *Proceedings of the international conference on information technology and nanotechnology (ITNT)*, pp 21–24, Samara, Russia
10. Caliper H (2019) Available online: <https://www.hyperledger.org/projects/caliper>. Accessed on 15 Jan 2019. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license. <http://creativecommons.org/licenses/by/4.0/>
11. Caro MP, Ali MS, Vecchio M, Giaffreda R (2018) Blockchainbased traceability in agri-food supply chain management: a practical implementation. In: *IoT vertical and topical summit on agricultureTuscany (IOT Tuscany)*. IEEE, pp 1–4
12. Casado-Vara R, Prieto J, la Prieta FD, Corchado JM (2018) How blockchain improves the supply chain: case study alimentary supply chain. *Procedia Comput Sci* 134:393–398. In: *The 15th international conference on mobile systems and pervasive computing (MobiSPC 2018)/the 13th international conference on future networks and communications (FNC-2018)/affiliated workshops [Online]*. Available: <http://www.sciencedirect.com/science/article/pii/S187705091831158X>

13. COTEU European Parliament (2002) Regulation (ec) no 178/2002 of the european parliament and of the council. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002R0178&from=EN>
14. Cui P, Dixon J, Guin U, Dimase D (2019) A blockchain-based framework for supply chain provenance. *IEEE Access* 7:157113–157125. <https://doi.org/10.1109/ACCESS.2019.2949951>
15. Dweekat AJ, Hwang G, Park J (2017) A supply chain performance measurement approach using the internet of things: toward more practical SCPMS. *Indus Manage Data Syst*
16. Galvez JF, Mejuto J, Simal-Gandara J (2018) Future challenges on the use of blockchain for food traceability analysis. *TrAC Trend Anal Chem*
17. Gao K, Liu Y, Xu H, Han T (2020) Hyper-FTT: a food supply-chain trading and traceability system based on hyperledger fabric. In: Zheng Z, Dai HN, Tang M, Chen X (eds) *Blockchain and trustworthy systems*. BlockSys 2019. Communications in computer and information science, vol 1156. Springer, Singapore. <https://doi.org/10.1007/978-981-15-2777-7-53>
18. Guin U, Cui P, Skjellum A (2018) Ensuring proof-of-authenticity of IoT edge devices using blockchain technology. In: 2018 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, pp 1042–1049
19. Leng K, Bi Y, Jing L, Fu H-C, Van Nieuwenhuysse I (2018) Research on agricultural supply chain system with double chain architecture based on blockchain technology. *Fut Gener Comput Syst*
20. Malik S, Kanhere SS, Jurdak R (2018) Productchain: scalable blockchain framework to support provenance in supply chains. In: *IEEE 17th international symposium on network computing and applications (NCA)*. IEEE, pp 1–10
21. Marchesi M, Marchesi L, Tonelli R (2018) An agile software engineering method to design blockchain applications. In: *Proceedings of the 14th Central and Eastern European software engineering conference Russia*. ACM, p 3
22. Min H (2019) Blockchain technology for enhancing supply chain resilience. *Bus Horiz* 62(1):35–45. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0007681318301472>
23. Pilkington M (2016) *Blockchain technology: principles and applications*. In: *Research handbook on digital transformations*. Edward Elgar Publishing
24. Porru S, Pinna A, Marchesi M, Tonelli R (2017) Blockchainoriented software engineering: challenges and new directions. In: 2017 IEEE/ACM 39th international conference on software engineering companion (ICSE-C), pp 169–171
25. Rice S (2015) The total scope of supply chain management. [Online]. Available: <http://www.apics.org/sites/apics-blog/think-supply-chain-landing-page/thinking-supply-chain/2015/03/11/the-total-scope-of-supply-chain-management>
26. Schatsky D, Muraskin C (2015) Beyond bitcoin. *Blockchain is coming to disrupt your industry*
27. Shakya S (2019) Efficient security and privacy mechanism for block chain application. *J Inform Technol* 1(2), 58–67. Bhalaji N (2019) QOS and defense enhancement using block chain for fly wireless networks. *J Trend Comput Sci Smart Technol (TCSST)* 1(1):1–13
28. Stafford TF, Treiblmaier H (2020) Characteristics of a blockchain ecosystem for secure and sharable electronic medical records. *IEEE Trans Eng Manage* 67(4):1340–1362
29. Tian F (2016) An agri-food supply chain traceability system for china based on rfid & blockchain technology. In: 2016 13th international conference on service systems and service management (ICSSSM). IEEE, pp 1–6
30. Tseng JH, Liao YC, Chong B, Liao SW (2018) Governance on the drug supply chain via gcoin blockchain. *Int J Environ Res Publ Health* 15(6):1055
31. Wang K, Liu M, Jiang X, Yang C, Zhang H (2020) A novel vehicle blockchain model based on hyperledger fabric for vehicle supply chain management. In: Zheng Z, Dai HN, Tang M, Chen X (eds) *Blockchain and trustworthy systems*. BlockSys 2019. Communications in computer and information science, vol 1156. Springer, Singapore. <https://doi.org/10.1007/978-981-15-2777-7-59>
32. Xu X, Rahman F, Shakya B, Vassilev A, Forte D, Tehranipoor M (2019) Electronics supply chain integrity enabled by blockchain. *ACM Trans Des Autom Electron Syst (TODAES)* 24(3):1–25

Jyoti: An Intelligent App for Women Security



Pooja Pathak and Parul Choudhary

Abstract In today's world of rapid urbanization, a new research chapter in social relations and women's safety has begun. With this wave of urban transformation comes a growing sense of loneliness and helplessness among certain social groups. Women's safety is a major issue in this situation, not only for the poor but also for civilized women. Weaknesses in the situation women's safety plans, practices, and guidelines exist to reduce violence against women and their fear of crime. While alone walking or traveling, women should be aware of their surroundings. The advancement of technology may also contribute to women's safety. Our proposed application is "Jyoti" for women's safety. This application helps to prevent female homicide. When a female is in a bad situation, the application will send a message to trusted people and the police will go to current location. This application has unique feature that it sends messages to trusted people at regular intervals while shaking the phone. It also calls for the first trusted contact to assist in an emergency. It will help to reduce female crime.

Keywords Audio · Global positioning system (GPS) · Video recorder · Security app · Switch

1 Introduction

Women are not safe today. Whenever she travels alone at night, if something goes wrong, she always needs a help and wanted family members nearby to protect her, which is impossible. The smart phone is one mode of communication, but she can't

P. Pathak (✉)

Department of Mathematics, Institute of Applied Sciences, GLA University, Uttar Pradesh, Mathura, India

e-mail: Pooja.pathak@gla.ac.in

P. Choudhary

Department of Computer Application and Engineering, GLA University, Uttar Pradesh, Mathura, India

e-mail: parul.choudhary_phd.cs20@gla.ac.in

always call or ask for help. If her phone has a security application that allows her to connect with her family, it may help women overcome insecurities.

According to various sources, 93 young ladies are assaulted daily in India.

- The number of assaults in 2013 was 1636, double that of 2012.
- 15,556 assault victims were aged between 18 and 30 in 2013.
- 16.5% of all crimes against women were committed while traveling in 2014.
- From 2010 to 2014, 15.95% (212,880) of all cases (1,333,973) involved young ladies being accused of wrongdoing.

This app is for women's safety. Unlike other similar apps, this app's location and background sounds are updated frequently.

Authors developed an android model that sends an email with the message [1]. Saranya and Karthik [2] features "Alarming neighbors by loud noise," "Auto dealing," and "Finding location of nearby police station and hospitals." Authors designed an app that sends the location of a place via GPS with a single click [3]. This application sends messages to trusted contacts every five minutes until the "stop" button is clicked. This work [4] gives an overview of Android's architecture and component models. It analyzes an Android app's anatomy. Murphy [5] explains the concept of Surekha, a woman in trouble receives a message from the app. Bhardwaj and Aggarwal [6] and Hagen [7] explores how a queer security analysis reveals how heteronormativity and CIS privilege sustain the current gendered violence analysis gap. When using an Android application, the device and the phone are synchronized via Bluetooth. It also records audio for further investigation.

2 System Design

2.1 Existing System

An existing application are discussed below.

1. **Safetipin:** The app has three main features: crisis contact numbers, GPS tracking, and safe area bearings. This app also saves the safe zones and their well-being scores for quick access [8, 9].
2. **Women security:** This application will send an SMS to a specific phone number and then take three pictures with the back and front cameras, which are stored in the database.

Himmat: To use the app, the user must first register with the Delhi Police. On completion of the application process, an OTP is sent and checked. As soon as the client sends an SOS alert from the application, GPS data and video/sound information are sent to the police control room, and then the police respond.

3 Proposed System

The framework is designed to integrate multiple options such as GPS. However, it also includes all the features of the most popular applications. Registration with a login id and a secure password is required. As a result, users should regularly update their emergency contact numbers. Google Map location with volume button will be enabled once the user travels. Once GPS is activated, the application will track the user and send an alert message to the crisis contacts. The client can use the protected telephone application to call the free helpline numbers. If the client wants to change the profile picture, they can do. The framework is modified to meet the needs of the anticipated application.

3.1 Features

Sign up and Maintenance

1. First, the user must register with their full name, phone number, and email address in the sign-up page.
2. An one-time password (OTP) is sent to the user's email. Then enters the OTP (verification code) to complete the registration. The user is then sent a message to complete the registration.
3. Users can manage their emergency contacts.

3.2 In an Emergency

- Sending user location.
- A phone shake alerts your recipient.
- Addition of audio recording.

It can capture video footage of the crime scene. Then it's sent to the contacts.

3.3 Workflow Chart (WFD)

The proposed application has four modules that are shown in Fig. 1: the login, sign up, alert message, and danger location. The sign-up module accepts name, phone, email, and password. The user is informed that "registration was successful". Then re-login to the home page to save up to five trusted contacts. In case of an emergency, the contact persons will be notified. The user can update these five contacts at any time with proper authentication.

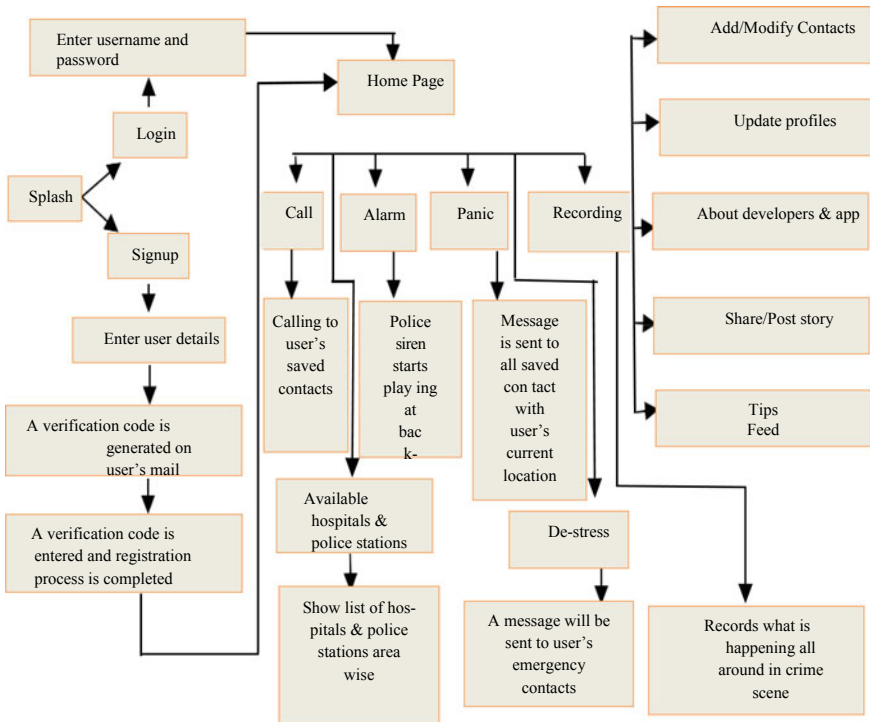


Fig.1 Workflow diagram

3.4 Data Flow (DFD)

An information development framework is depicted and investigated using data flow diagram (DFD). It shows the data stream and the changes as it moves from contribution to yield.

3.4.1 DFD 0

Figure 2 shows the system model description.

3.4.2 DFD 1

Figure 3 depicts the separation of all modules and the application system. This module contains DFD sign-up module with four options. After registering, you can send messages, save contacts, and manage contacts.

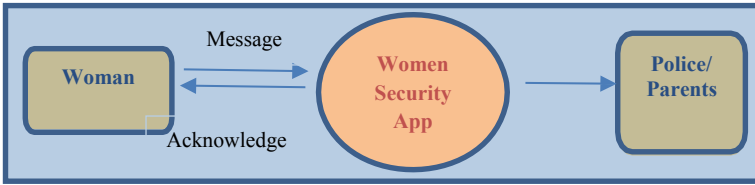


Fig. 2 Model description

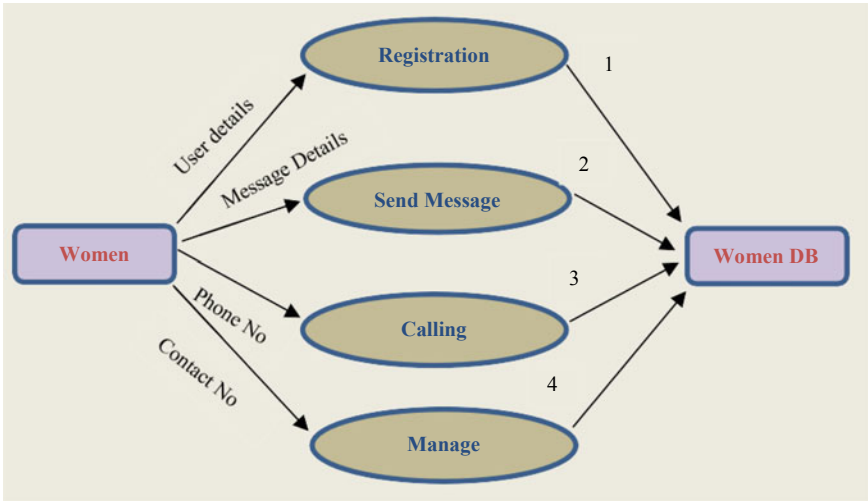


Fig.3 Network architecture of Jyoti

3.4.3 DFD 2

Figure 4 is differentiating the modules between front end and back end. In this module, the details are fetched from the database.

4 Technology Used

4.1 Firebase

The database is live. Firebase is a cloud-based client database validation framework that stores a confirm client’s details. No SQL database is in Firebase because it uses node; its administration is faster than other web administrations.

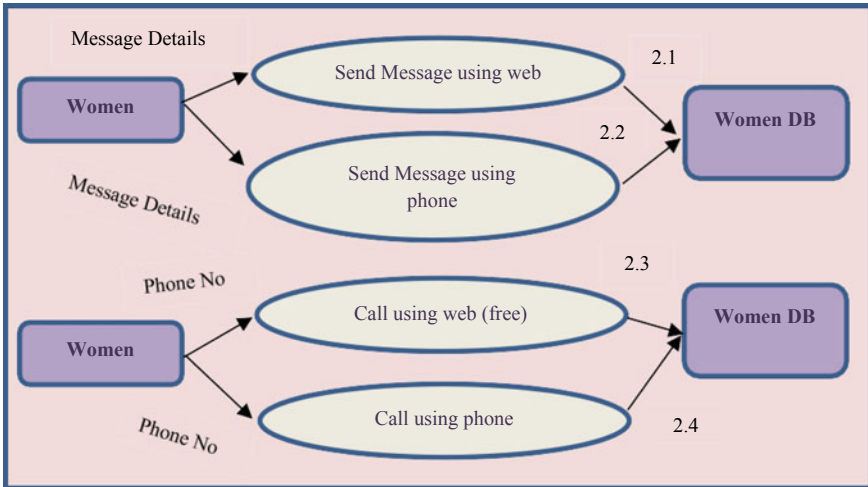


Fig. 4 Women database details

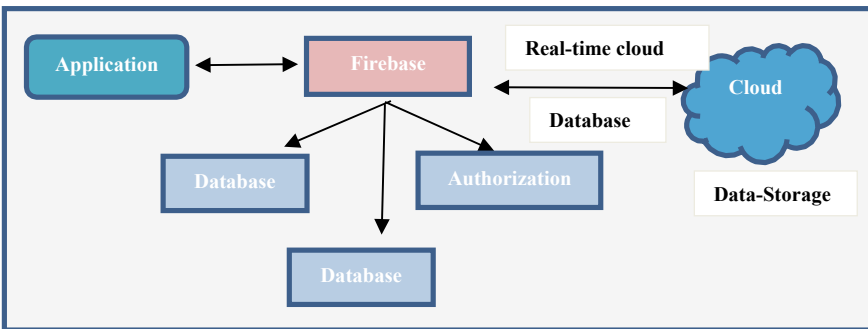


Fig. 5 Schematic diagram of Firebase database

Figure 5 shows how the application network connects to Firebase. It stores data in a database. This will upload the data to the cloud. We can also see how data is saved into our database and who updates the data.

4.2 Conditions for Optimal Function Application

- Location updates are preferred on phones with long battery life and high Internet speed.
- Every smartphone must have GPS.

Table 1 Comparative study between existing women security app and Jyoti

App features	Safetipin	Jyoti	Himmat	Women	Proposed application
SOS	■	■	■	■	■
Shortcuts		■			■
Police alert			■		■
Audio		■	■		■
Camera		■		■	■
Real-time database	■	■			■
GPS tracking	■	■			■

4.3 Comparison Table of Existing Applications

Such application has all the details but what are the features are broken? Therefore, all features in our application work safely. We can add more options like calling, tracing, etc., in the future. So, our application is very flexible and easy to use. Table 1 lists the app’s features.

5 GUI of Jyoti Application

5.1 Login Page

Figure 6 is the login page. Our application requires an email and password to login. Then email verification is required [9], and an OTP is sent to that email.

5.2 Sign up

In Fig. 7, the sign-up page is 7, and its registration requires name, email, password, and mobile number. User gets alert after successful registration (Registration is done successfully).

5.3 Contact Page

See Fig. 8 for details on how to contact your parents and other trusted people in case of any emergency. We can also save the cops. Currently, we can save five numbers, but in the future, we can save many more. Following that we notify the numbers. Anyone can change the numbers by selecting edit contacts on the danger location page of proposed application [10].

Fig. 6 Login page



Fig. 7 Sign-up page

Fig. 8 Contact page



5.4 Location Page

Figure 9 shows the user's location in a dangerous situation. On the location page, you can see your coordinates. This coordinate will help locate the user or the area where they are currently located. Each contact in our database receives an offline message [11, 12].

Figure 9b shows a red danger button. Pressing that button sends a text message to the trusted contact numbers that store the coordinates and the message as shown in Fig. 9a.



Fig. 9 a, b Location page of Jyoti

6 Conclusions

This is the “Android application” for women’s safety. Women can use this app at anytime and anywhere when they feel unsafe, such as alone in a taxi at night. So, they can easily share their location with their loved ones. When they click the button, it updates all authorized contacts with the location. This application ensures women’s safety.

References

1. Pasha S, Kavana J, Mangala GKR, Nischitha K, Surendra BK, Rakshitha MS (2016) B-secure for women: an android application. *Int J Innov Res Comput Commun Eng* 4(5):8073–8080
2. Saranya N, Karthik K (2015) Women safety android application using mobile phone. *Int J Eng Sci Comput* 1317–1319
3. Yarrabothu RS, Thota B (2015) Abhaya: an android app for the safety of women. In: Annual IEEE India conference 2015 (INDICON). IEEE, pp 1–4
4. Shu X, Du Z, Chen R (2009) Research on mobile location service design based on Android. In: 5th International conference on wireless communications, networking and mobile computing 2009. IEEE, Beijing, pp 1–4
5. Murphy ML (2010) Android programming tutorials. Commons Ware
6. Bhardwaj N, Aggarwal N (2014) Design and development of “Suraksha”—a women safety device. *Int J Inf Comput Technol* 4(8):787–792
7. Hagen JJ (2016) Queering women, peace and security. *Int Aff* 92(2):313–332
8. Monisha DG, Monisha M, Pavithra G, Subhashini R (2016) Women safety device and application-FEMME. *Indian J Sci Technol* 9(10):1–6
9. Patel V, Kapadia D, Ghevariya D, Pappu S (2020) All India grievance redressal app. *J Inf Technol Digital World* 2(2):91–99
10. Sungheetha A, Sharma R (2020) Real time monitoring and fire detection using internet of things and cloud based drones. *J Soft Comput Paradigm (JSCP)* 2(03):168–174
11. Pathak P, Farheen N, Dubey A (2021) Prediction of students’ performance in education system based on artificial intelligence. *IOP Conf Ser Mater Sci Eng* 1116(1)
12. Choudhary P, Pathak P (2021) A review of convolution neural network used in various applications. In: 2021 5th international conference on information systems and computer networks (ISCON). IEEE

Performance Analysis of Machine Learning Algorithms in the Systematic Prediction of Chronic Kidney Disease on an Imbalanced Dataset



M. Revathi, G. Raghuraman, and J. Visumathi

Abstract The applications of artificial intelligence in the field of medicine is growing rapidly. Huge volume of patient related data is available and if this data is properly utilized, it will be helpful in the diagnosis as well as treatment of many diseases. Chronic kidney disease (CKD) is a progressive reduction in the function of kidney which is not reversible. During the progress in disease development, patients may suffer from other health issues like diabetes, high blood pressure, anaemia etc. At the end stage, dialysis will be a temporary solution and kidney transplantation is the permanent solution. The progress of the disease towards end stage may stop, if we identify the disease at the earlier stage and accurate treatment is given. Nowadays, machine learning as well as deep learning algorithms helps the medical experts in the prediction of many diseases. In this work, various algorithms have been used in the prediction of CKD and a comparison of results have been done. It is found that artificial neural network (ANN) is able to predict the disease with higher accuracy than the other algorithms being used.

Keywords Machine learning · Chronic kidney disease · Classification · Artificial neural network

1 Introduction

In general, food is taken by a human being and it is converted into energy [1]. The needed energy will be absorbed and sent into the blood and unwanted waste will be sent out. Our body organs will release chemicals into blood which needs to be

M. Revathi (✉)
St. Joseph's Institute of Technology, Chennai, India
e-mail: gmrevathigopinath@gmail.com

G. Raghuraman
SSN College of Engineering, Chennai, India
e-mail: raghuramang@ssn.edu.in

J. Visumathi
Veltech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, India

filtered out. Kidney is an important organ in our body which is useful for filtering the wastages from our blood. The wastes are sent along with the water in the form of urine. Blood cells [2] are produced in bone marrow and its lifetime will be approximately 120 days. After that the blood cells will be destroyed and new cells must be created for the successful body function. Erythropoietin is a hormone produced by kidney and it helps in the generation of blood from the human body. In case of kidney failure, erythropoietin will not be produced, new blood cells will not be created and thus the person will be anaemic.

Kidney is the organ that helps in the production of Vitamin D which is needed to absorb calcium from the food taken. Calcium is an important vitamin which is important for the growth of bones and teeth in a healthier way. If Kidney fails, Vitamin D will not be produced. Without Vitamin D, calcium will not be absorbed and all the calcium in our food will be sent out as waste. Potassium is the most dangerous waste formed in the blood and its level should not go above 7. Normal value of potassium should be 3.6–5.2 mmol per litre (mmol/L). In case of kidney failure, the level of Potassium in blood will increase. If it goes above 7, the function of heart will be stopped, brain enters into coma and muscles will not work.

Hydrogen (H^+) Ion is the next dangerous chemical formed from blood. If Kidney function fails, the amount of H^+ ions will increase. When H^+ ion increases, the level of PH decreases and acidic level of our body gets increased and thus the body organs will be damaged. Ammonia is another dangerous chemical formed from blood and its normal value should be 60. Increased level of ammonia will make the person to enter into unconscious stage. Ammonia is insoluble in water and cannot be converted into urine. Ammonia will be converted into urea and sent out as urine with the help of kidney and liver. In case of kidney failure, the amount of Ammonia in blood will also increase.

There are many causes for kidney disease [3] which includes age, hypertension, diabetes, increased body mass index (BMI), smoking habit, food habits, hereditary kidney diseases and infections due to bacteria, virus, and parasites etc. Hypertension causes blood to flow forcefully into kidney and thus kidney will be affected. Diabetics increases osmolality and can cause kidney damage.

Due to increased osmolality, more amount of urine will be generated. Most of the blood vessels in kidney will be damaged and thus the needed oxygen will not be supplied to kidney. The chemicals in the pain killer tablets that we take also affects the nephrons in kidney. The change in food habits can also affect the kidney. If BMI (obesity) is increased, the work of kidney will be increased and thus it will be affected. If enough water is not taken by a person, stones will be formed in kidney and thus affects kidney function. Thus, kidney acts as a filter [4], retains the needed energy from blood and unwanted materials will be sent out. If kidney is damaged, needed energy will be sent out and unwanted materials will be retained. One of the wanted material is albumin, and it will be sent out when kidney is damaged. Since needed materials are sent out, our immunity will be reduced and there is an increased chance of diseases.

Chronic kidney disease (CKD) [5] is defined as a progressive reduction in the function of kidney which is irreversible. There are two methods [6] to identify the

Table 1 Categories of CKD

Different stages	Filtration value	Explanation
I	Greater than or equal to 90	Slight damage without loss in the function of kidney
II	Sixty to eighty nine	Damage with mild loss in the function of kidney
III	Thirty to fifty nine	In range of low to severe loss in the function of kidney
IV	Fifteen to twenty nine	Heavy loss in the function of kidney
V	Less than fifteen	Failure of kidney

prevalence of CKD: reduced glomerular filtration rate (GFR) and increased urea content in albuminuria test. Using urinalysis, biopsy and imaging, the structural as well as the functional abnormalities in kidneys could be identified. The estimation of GFR is done based on the concentration of creatinine in serum along with features such as age, sex, and origin using Modification of Diet in Renal Disease (MDRD) equation. Based on the value of GFR, the CKD can be categorized as stages of five as illustrated in Table 1.

When GFR value is less than 15 [5], the kidneys will be fully damaged and the ability of filtering the blood cannot be done in the usual way and this stage is known as end stage renal disease. When this end stage is reached, two options are available for the patient to survive: Hemo-dialysis and kidney transplantation. Hemo-dialysis is a technique of filtering the blood which is done for a duration of 4 h and it should be done three times per week. Kidney transplantation provides a better solution than Hemo-dialysis to patients. With kidney transplantation, the patient will be able to lead a normal life with improved renal function.

The amount of medical data grows exponentially on day-to-day basis and this in turn increases the need to analyse those data [7]. If medical data is properly analysed, it can be used to cure many diseases at earlier stages. Artificial Intelligence (AI) is a technique in which Intelligent machines that mimic human are created. Nowadays, AI is an important technique that helps doctors to accurately predict the disease and also to analyse and suggest patient specific treatment. In AI, machines will perform the operations for which it is programmed.

There is a sub-area of AI called machine learning (ML) [8] which helps in producing machines that can automatically learn from examples and does not require to be programmed explicitly. Supervised machine learning, unsupervised machine learning and reinforcement machine learning are three broad classification of ML techniques. The supervised learning algorithms takes the datasets with pre-defined labels as input and learns from it. The two main categories of supervised learning are classification and regression. The unsupervised learning algorithms takes unlabelled datasets as input and they try to extract the pattern and features from those input datasets. The main example of unsupervised learning includes clustering. Reinforcement learning works on basis of trial and error and it uses the concept of reward and punishment.

Deep learning [9] is a subset of ML which imitates human brain in processing the records. DL is preferred over ML for the following reasons: (i) the performance of DL is directly proportional to input data size. Since huge volume of medical data is available, DL performs better than ML. (ii) Medical data can be of any type including images, Genetic expressions, text, or signals. DL can process all input types more efficiently than ML. (iii) In ML, fundamental information about the dataset must be processed manually while in DL it is obtained automatically.

In this paper, ML methods are applied to the CKD dataset. Amongst those methods taken for analysis, it is found that ANN performs better than other algorithms. A feature selection method based on dispersion measure is chosen in order to retain the attributes that are highly relevant to the classification task. Also the attributes that occurs repeatedly (redundant) and irrelevant to the classification task needs to be eliminated from the dataset considered as input. Then the new CKD dataset is given as input to ANN which shows improved results than dataset with original set of features.

2 Related Works

In [10], Zhang et al. uses ANN for predicting the survival of patients with CKD. Since multiple inter-related health factors such as high blood pressure, cardio-vascular disease can also influence the patients survival, authors have used ANN to establish a mapping from inter-related health factors to the survival of patients.

In [11], Yildirim investigates how the performance of the neural network algorithms are affected by using imbalanced datasets. Sampling techniques were applied on the dataset to overcome the problem of imbalanced distribution of dataset. A comparative study of using different under sampling and over sampling techniques was done. Back propagation neural networks was used to find the patterns in the dataset for predicting CKD. The learning rate of neural networks is varied and the corresponding results were compared.

In [12], Islam and Ripon have evaluated the effective execution of various boosting classifiers, Ant-Miner procedures, and J48 decision tree in predicting CKD. The also demonstrates the relationship between attributes that are used to predict CKD. In [13], Bhaskar and Manikandan devised an algorithm known as correlational neural network (CorrNN) and it is combined with SVM to enhance the precision of the devised algorithm. In this work, authors found that the overall computational time is reduced by 9.85%.

In [14], Pradeepa and Jeyakumar suggested that the earlier detection and stage categorization of CKD is important. But due to the redundant and high dimensional data, earlier detection, and categorization is difficult. In this work, authors introduced a self-tuning spectral clustering algorithm to overcome the effects of high dimensionality and redundancy. The resulting new data dataset is given as input to various

algorithms like deep neural network (DNN, random forest (RF)), ANN, support vector machine (SVM), and K-Nearest Neighbour (KNN). The authors found that SVM and KNN shows better results.

In [15], Ma et al. devised heterogeneous modified artificial neural network (HMANN) for the identification and classification of kidney damage in the earlier stages. The concept of SVM and Multi-Layer Perceptron (MLP) is also used. In [16], Almansour et al. uses mean of attributes to all the values that are found to be missing in the dataset. The SVM and ANN algorithms were used to predict the prevalence of CKD and it is found that ANN performs better than SVM.

In [17], Akter et al. used various feature selection methods like Wrapper Method—CFS (Correlation Based Feature), Recursive Feature Elimination (RFE) and LASSO regression are used to find the most relevant features from the CKD dataset. The resultant dataset is given as input to various algorithms like ANN, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bi-Directional LSTM and MLP. Experiments shows that ANN and MLP has produced higher accuracy of prediction. In [18], Elkholy et al. modified Deep Belief Network with different activation function and loss function. The activation function used is softmax and the loss function used is cross-entropy. The experiments shows that the prediction accuracy is 98.5%.

In [19], Alloghani et al. applied various algorithms like logistic regression, neural networks, RF and decision trees for the prediction of CKD. Multiple factors along with Estimated GFR (eGFR) are considered for the prediction of CKD and the results shows better performance. In [20], Revathi et al. analyses the performance of various algorithms in the prediction of CKD and found that Gradient Boosting performs better than other algorithms.

3 Methodology

Artificial neural network (ANN), Naive Baye’s, logistic regression, K-Nearest Neighbour (KNN), decision tree, and support vector machine (SVM) are the classification techniques used in this paper. Each of these techniques has been briefed. Figure 1 shows the steps generally followed by a supervised machine learning algorithm.

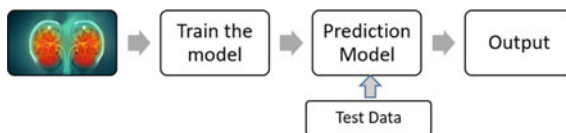


Fig. 1 Working steps of supervised ML algorithms

3.1 Artificial Neural Network

Human brain is made up of billions of neurons and interconnection amongst all neurons makes brain as a complex system which is capable of many new discoveries. Axons, body cell, and dendrites are the three important components in a neuron. The use of dendrites is the reception of incoming data that arrives from other neurons and axons are used to transmit information to other neurons. Artificial neural network (ANN) [21] is a technique for implementing the working of brain in machines. ANN imitates the structure as well as function of the neurons in human brain [22].

ANN is made up of many neurons (basic processing unit) [21] and these neurons are organized into layers which is categorized into three types: input layer, hidden layer and output layer. Each of the layer is an aggregate of group of nodes referred as neurons. Figure 2 shows a simple model of ANN which has only one hidden layer.

- (i) **Input Layer:** There should be only one input layer. Data in the form of patterns, images, signals are sent as input to this layer. The data are sent in the form of values in vectors. Some form of pre-processing of data is done for getting better results.
- (ii) **Hidden Layer:** There can be one or more hidden layer. The inputs are sent along the weighted edges to the hidden layer. First the hidden layer computes the sum of product of each of the input values and its corresponding weights, and the result is summed with the bias value given by Eq. 1.

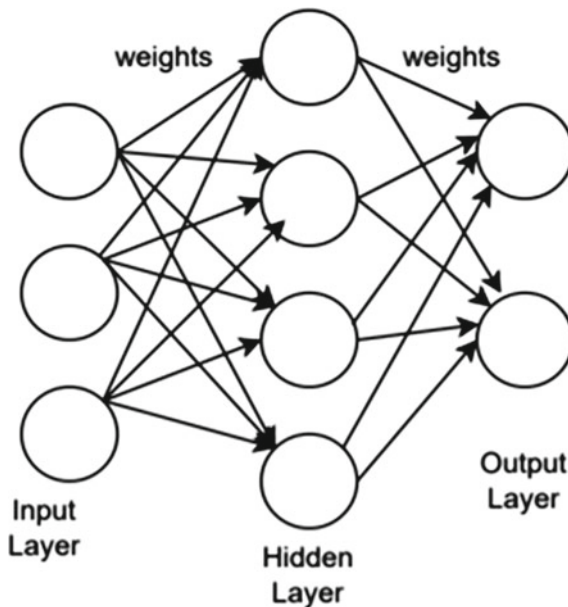


Fig. 2 ANN with 1 hidden layer

$$f(X) = \sum_{i=0}^n (x_i * w_i) + \text{bias} \quad (1)$$

The value of $f(X)$ is then given to activation function which has some threshold value. There are many activation functions like Binary, Sigmoid, ReLu etc., and these activations determine whether the neurons can be fired or not. The neurons that can be fired will send its value along with the weighted edges to the next layer. This process continues for several layers of hidden layers and finally the values will be passed as input to the last layer i.e. the output layer.

- (iii) **Output Layer:** There can be only one output layer. This is the last layer and it is responsible to determine the results. The results are calculated based on the values received from the previous hidden layer.

Activation function [23] has a main important role in the working of ANN. It helps the ANN to make use of relevant features while concealing the effect of irrelevant features in the input dataset. Activation function introduces the property of non-linearity to the network that ANN builds. The following are some of the commonly used activation functions:

- (i) **Binary Step Function:** Here, a value will be chosen as threshold and the input value will be compared against the threshold value. The neuron will fire, i.e. it sends its computed value to neuron of next layer only if it exceeds the threshold as given in Eq. 2.

$$f(x) = \begin{cases} 0 & \text{for } x < \text{threshold} \\ 1 & \text{for } x \geq \text{threshold} \end{cases} \quad (2)$$

- (ii) **Sigmoid Function:** This is commonly used when the value of output is decided based on probability. The function defined is differentiable and this derivative value will be used by the neural network learning algorithm. The input to the sigmoid function will be any real value and the value of output is from 0 to 1 and the function is given by the Eq. 3.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

- (iii) **TanH Function:** This function is similar to the above defined sigmoid function but the difference is that the value which is real in nature is given as input to the tanH function and the value of output is from -1 to 1 and the function is given by the Eq. 4.

$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (4)$$

When compared to sigmoid function, the use of tanH activation function is more. This is because of the efficiency of the tanH activation function in the artificial neural network.

- (iv) **Softmax Function:** This function is developed to overcome the drawbacks found in sigmoid function. Softmax function uses the concept of relative probability. The binary class categorization problem makes use of sigmoid function while for classification problem with numerous classes Softmax function is used. The function is given by the Eq. 5.

$$\text{softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (5)$$

- (v) **ReLU Function:** This function produces better performance when compared to any activation functions. It uses the concepts of linear transformation function and it also uses the concept of threshold. Here, if the input value is less than 0, the output will be zero else the output will be same as input value as given in the following Eq. 6.

$$f(x) = \max(0, x) = \begin{cases} x_i & \text{if } x_i \geq 0 \\ 0 & \text{if } x_i < 0 \end{cases} \quad (6)$$

3.2 *K-Nearest Neighbour (KNN) Classification*

One of the popular supervised learning algorithm is KNN [24] in which labelled datasets are used. It is used for the applications of classification as well as regression. KNN relies on the observation that the instances that are similar in characteristics will appear nearer to each other. Distance is an important measure in determining how similar the objects are. If distance is smaller, the objects are closer to each other and if the distance is higher, the objects are dissimilar.

KNN is called as Lazy Learner, since the algorithm will not do training in prior [25]. The working of the algorithm starts only after the test data set is received. The value of K should be initialized in prior. Once the test data is received, the distance between the new unknown data points and all of the data points in training data should be calculated. The distance should be sorted in ascending order and the top K training data points should be considered as neighbours of the test data. The majority class amongst the neighbours should be assigned to the test data.

There is no predefined method for selecting the value of k [26] and it is selected based on trial and error method. The most frequently used measure to calculate the distance between instances is Euclidean distance which is given by Eq. 7 in which the instances are (x_1, y_1) and (x_2, y_2) ,

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (7)$$



Fig. 3 Step-by-step execution concept of KNN

Figure 3 shows the Step-by-Step Execution Concept of K-Nearest Neighbour classification. The value of k is assumed to be 3. At first, the data with class labels (circle and triangle) will be placed in its appropriate position. Next when a new data point arrives, distance between the new incoming unknown data point and all of the existing data points needs to be estimated. Since k is 3, the first three data points that is at smallest distance from new object is considered (two circle and one triangle). Since the major class amongst three neighbours is circle, the new object is assigned to the class circle.

3.3 Naive Baye’s Classification

Another popular supervised learning algorithm is Naïve Baye’s [27], and it depends on the concepts of Baye’s Theorem (Probabilistic theorem) which is given by Eq. 8.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{8}$$

In which A and B are events and $P(B) \neq 0$. $P(A)$ is termed to be Prior probability [28] and it represents the chance in which event A will happen and $P(B)$ represents the chance in which event B will happen. $P(A|B)$ is known as Posterior probability and it demonstrates the chance in which A occurs only when event B occurs. The equation is developed by assuming that there will independence amongst the features used in the input dataset. It uses the concept of probability to find the class of the new data objects. There are three types of Naïve Baye’s Classification [29].

- (i) Multinomial Naïve Baye’s: It is used in classification problem where the test data may belong to different types of classes. When the input dataset has discrete set of features (such as word counts), Multinomial Naïve Baye’s approach is used.
- (ii) Bernoulli Naïve Baye’s: It is similar to Multinomial type but the classification will be boolean (yes or no).

- (iii) Gaussian Naïve Baye's: It is applied in a situation where the predictor variables are continuous data. Gaussian distribution is used to sample these values. A curve which is bell shaped is formed using the data points and with the average value of the attributes the curve will be symmetric.

3.4 Logistic Regression

One of the famous supervised learning algorithm is logistic regression [30], and it acquires its name from the fact that it uses logistic function. The logistic function/sigmoid function is given by Eq. 9.

$$\text{sig}(z) = \frac{1}{1 + e^{-z}} \quad (9)$$

This function resembles a curve in the shape of *S* and it will map the input into value between 0 and 1, i.e. the logistic function tries to map the input values to probabilities. This methods is used find the relationship between the dependent feature and one or more predictor feature. It is well suited for binary class classification [31]. Since logistic function is used, it is difficult to understand the predicted output as a linear combination of inputs. The output feature in logistic regression will be of categorical type. So, the output will be either true (0) or false (1) similar to the prediction of cells as cancer affected or not. The input feature can be in the form of either discrete or continuous. Logistic regression can effectively find the features in the input dataset that are helpful in the accurate classification.

The idea behind both the linear regression and logistic regression are almost same [31]. The difference is that for regression problems, linear regression is used and for classification type of problem, logistic regression is used. In linear regression, we will try to position the new unknown data point in the constructed regression line. But in Logistic regression, we will try to position the new unknown data point in a 'S' shaped curve with the help of the logistic function.

The logistic regression model relies on the use of threshold value. Initially, a suitable threshold value is chosen. The output of the logistic function is matched against the threshold value. If the value of the output exceeds the threshold, then it will be placed in class true (1). If the value of the output lies below threshold, then it will be placed in class false (0). Figure 4 represents the logic behind the working of logistic regression.

3.5 Decision Tree

One of the most commonly used supervised learning algorithm is decision tree [32] and it constructs a tree based structure for classification. The leaf node in the tree

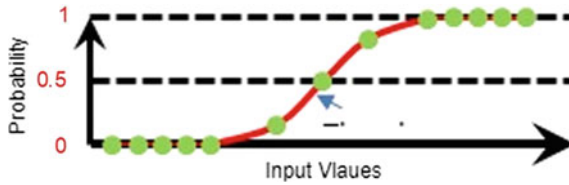


Fig. 4 ‘S’ curve of logistic regression

represent class labels and every node that is internal (not leaf node) will indicate a test that applies to the attributes. The major issue in building the tree is the identification of attribute nodes at each level which is referred to as the selection of attributes. Techniques such as information gain as well as Gini Index are the two popular methods employed for attribute selection.

Two kinds of nodes are identified in decision tree, namely decision node and leaf node [33]. The decision node represents the attributes whose value is to be compared and based on the test on attribute the nodes will follow downwards towards the leaf node. The leaf node represents the class of the new input data. There will be links/branches from decision node but the leaf node act as terminal node of the tree. Decision tree is preferred due to the fact that the tree structure will give easy and better understanding. The working flow of the decision tree which starts from top root node towards the leaf node.

The algorithm works as follows [34]: Initially, all the data points will be kept in single node called root node. Next, we will use attribute selection techniques to find the best attribute to classify the objects. The best attribute will be made as the decision node. Relying upon the value of that best identified attribute, the original set of whole objects is divided into two subsets. The above steps needs to be replicated for the obtained subtrees. This procedure continues until we reach a stage in which the dataset cannot be further classified.

Information gain and Gini index are widely used methods for selecting best attribute. Information is a mechanism that calculates the amount of data/information that each attribute contains about the class. The attribute with the value of higher information gain will be chosen as first decision node. The process will continue until the information gain of all child node reaches 0. Equation gives the calculation of Information gain where ‘S’ indicates the ‘set of all objects’, ‘A’ indicates ‘attributes’, ‘S_v’ indicates the ‘objects in S with v as value for attribute A’ and ‘values(A)’ indicates ‘the set of all possible values for the attribute A’. Equations 10 and 11 shows the calculation of entropy which is used the calculate the level of impurity.

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \tag{10}$$

$$\text{Entropy}(S) = \sum_{i=1}^c -P_i \log_2 P_i \tag{11}$$

Gini Index is another attribute selection measure that calculates how often the incoming new data is classified incorrectly. The attribute with lower value of Gini index is first chosen and the process continues. Equation 12 shows the calculation of Gini Index.

$$\text{Gini - Index} = 1 - \sum_j P_j^2 \tag{12}$$

3.6 Support Vector Machine (SVM)

The most well-known technique for classification problem is support vector machine which is a supervised machine learning approach. SVM can also be used for regression type of problems. A graph with n -dimensions is constructed where ‘ n ’ represents the value of total number of attributes found in the dataset considered as input. Initially, all the data objects will be placed as points in the graph of n -dimension. Many lines can be drawn in the graph between objects. The goal of this SVM is to find the suitable line that will be used to distinguish the group of one of objects from another group of objects. This kind of identified suitable line is referred to as hyperplane. Figure 5 shows an example of hyperplane in SVM.

Hyperplanes can be many dimensions and the number of dimensions depends on the number of features in input dataset. If two features are present in input dataset, then hyper-plane will be the kind of straight line. If three features are present, then hyper-plane will be a plane with two dimension and so on. The objects that appears closer to the hyperplane and influence the location of hyperplane is referred as support vectors. SVM generally belong to two categories:

- (i) Linear SVM: In linear SVM classifier, it is possible to find a single straight line to distinguish the group of one class of objects from the group of another class of objects.

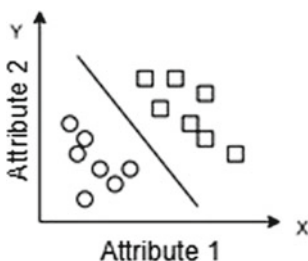


Fig. 5 SVM in 2D

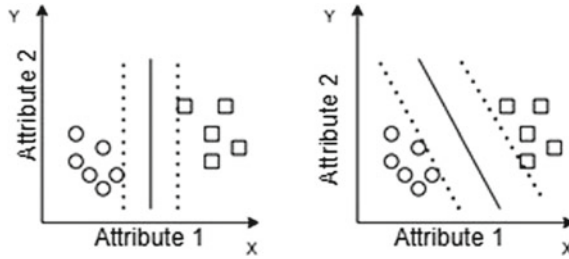


Fig. 6 Selection of best boundary line

- (ii) **Non-Linear SVM:** In non-linear SVM classifier, it is not possible to find a single straight line to distinguish the group of one class of objects from the group of another class of objects.

Figure 6 shows two possible lines that can be drawn to separate the class of one object from another class of object. The line in right hand side is preferred since it has maximum marginal distance (distance form hyperplane to margin). SVM uses a special type of function called kernel to convert the objects in non-linear form to linear form. The kernel function will take the data objects as input and transform them in the needed form.

The following are some types of frequently used kernel functions:

- (i) **Linear Kernel:** When there are lot of attributes in input dataset, this kernel function is preferred and it is computed using the Eq. 13 where x_1 and x_2 represents data samples.

$$K(x_1, x_2) = \sum (x_1 \cdot x_2) \tag{13}$$

- (ii) **Polynomial Kernel:** This is considered to be the generalised view of linear kernel and it has its applications in image processing and is calculated using the Eq. 14.

$$K(x_1, x_2) = ((x_1 \cdot x_2) + 1)^d \tag{14}$$

- (iii) **Gaussian Radial Basis Function Kernel:** This is considered to be the preferred kernel function in SVM classifier. It does not demand for any information about the data in prior and gives better classification results and is found by the Eq. 15.

$$K(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2) \tag{15}$$

The variable gamma can take any values between 0 and 1 and the most frequently used gamma value is 0.1.

- (iv) Sigmoid Kernel: The sigmoid kernel is very much preferred in the area of neural networks and is given by the Eq. 16.

$$K(x_1, x_2) = \tanH(\alpha x_1^T x_2 + c) \quad (16)$$

4 Results and Discussion

4.1 Dataset Description

Figure 7 shows the distribution of disease amongst the persons considered. From 400 persons taken into consideration, 250 has CKD and 150 doesn't have CKD. Analysis of different attributes of the input dataset has been done to find the importance of attributes in prediction of CKD. Proper preprocessing and normalization techniques have been applied to remove missing values and to keep the value of attributes within specified range. This is done because the presence of missing values and large range of values may affect the prediction of output and thus the prediction accuracy may not be obtained as expected.

The input dataset for the prediction of chronic kidney disease dataset is retrieved from UCI machine learning repository [22]. There are a total of twenty five attributes and four hundred objects in the UCI dataset. Out of 400 total instances, 250 instances are of class 'CKD' and 150 instances are of class 'Not CKD'. Out of the 25 attributes

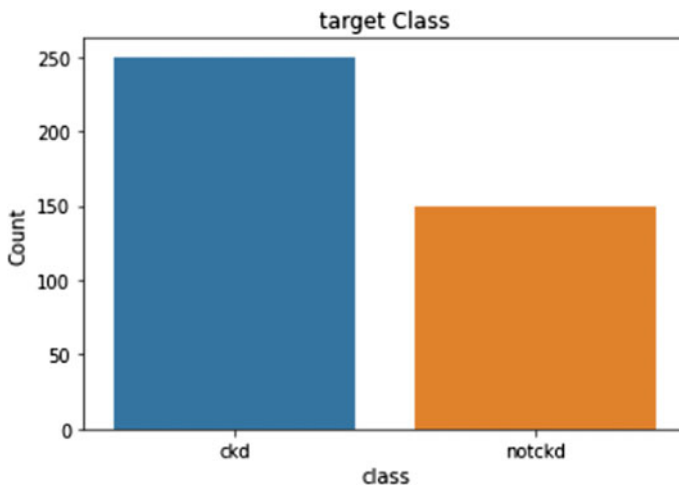


Fig. 7 CKD class distribution

available, 11 attributes belongs to the category of numerical type and 14 attributes belongs to the category of nominal type.

Figure 8 shows the correlation amongst various attributes in the input dataset. Correlation analysis is a technique to discover the relationship/Patterns amongst the attributes in the input dataset. Correlation can be either positive values or negative values. Positive correlation values demonstrates the fact that when the value of one correlated attribute gets shoot up, the value of another attribute also gets increased (directly proportional). Negative correlation values demonstrates the fact that when the value of one correlated attribute gets declined, the value of another attribute also gets decreased (inversely proportional). If value of correlation is 0, we can infer that there is no relationship between the two attributes.

Table 2 shows the analysis of the impact of red blood cells in the cause of chronic kidney disease. Totally there are 77 persons with abnormal red blood cell. Out of 77 persons, 76 has CKD and 1 is non-CKD. So if red blood cells are abnormal, there are high chances of occurrence of chronic kidney disease. There are totally 323 persons with normal red blood cells. Out of 323 persons, 174 has normal red blood cells and 149 has abnormal red blood cells. It has been observed that in persons with normal



Fig. 8 Correlation amongst attributes

Table 2 Analysis of the impact of RBC

Red blood cells	Class	Count	Mean	Min	Max
Abnormal	CKD	76	4.38	2.5	6.5
	notCKD	1	4.9	4.9	4.9
Normal	CKD	174	4.27	2.1	8.0
	notCKD	149	5.35	3.7	6.5

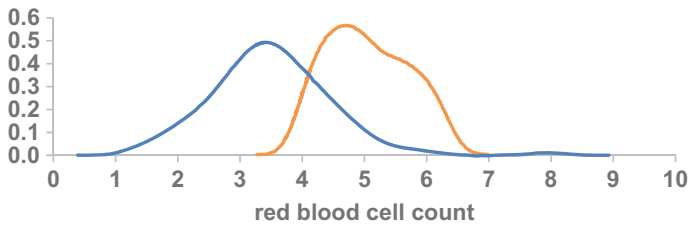


Fig. 9 Impact of RBC

red blood cells if minimum value is 2.1 and maximum value is 8, there are chances of chronic kidney disease. Figure 9 shows the graphical representation of the analysis.

Figure 10 represents the density of people in each age group who were considered to collect the chronic kidney disease dataset. From the graph, it can be understood that the persons considered are mostly in the age group of 35–75. Since the age group of 35–75 has high chances of chronic kidney disease, data has been collected from such group of peoples. Also, some persons from other age group is also considered for collection of data.

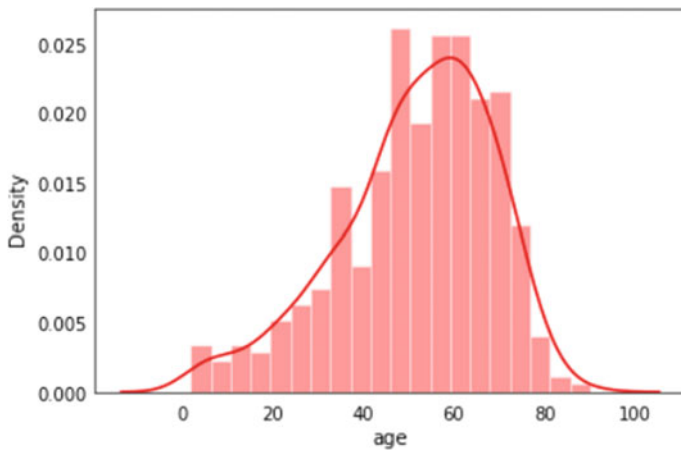


Fig. 10 Collection of data amongst different age groups

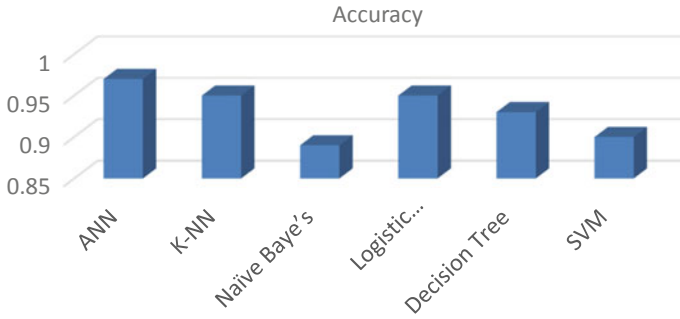


Fig. 11 Analysing the accuracy of various algorithms

4.2 Accuracy

For getting good results, it is necessary to determine how well the classifier performs. Accuracy is a commonly used parameter to access the performance. Accuracy is calculated by dividing the ‘total predictions which are correctly done’ by the ‘total possible number of data objects considered for prediction’ as given in Eq. 17.

$$\text{Accuracy} = \frac{\text{Number of Correct Prediction}}{\text{Total Number of Input Samples}} \tag{17}$$

Figure 11 depicts the analysis of accuracy of ANN, K-NN, Naïve Baye’s, logistic regression, decision tree, and support vector machine. From the implementations and results obtained, it is found that ANN performs well than the other classifiers.

Precision, Recall and F1-Score

Precision is given by the value of correctly predicted positive instances divided by the total predicted positive instances as shown in Eq. 18.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{18}$$

Recall is given by the value of correctly predicted positive instances divided by the total instances in class-yes as shown in Eq. 19.

$$\text{recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{19}$$

F1-Score is found by using the values of both precision and recall and is calculated as shown in Eq. 20.

$$F1 - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{20}$$

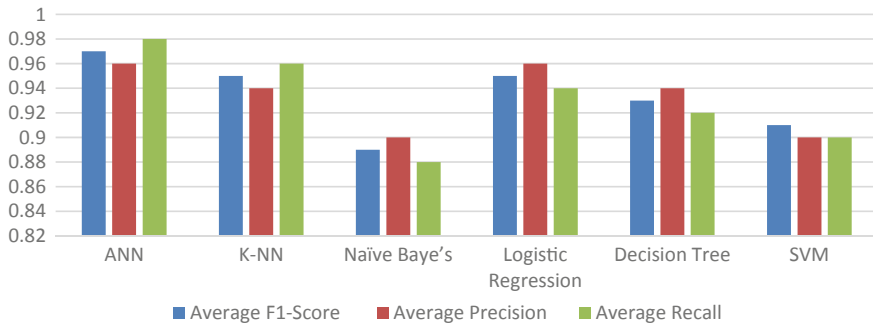


Fig. 12 Comparison of *F1*-score, precision, recall

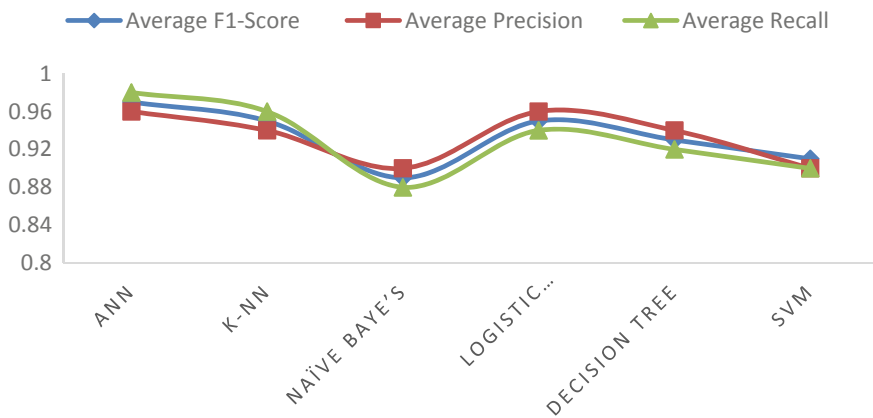


Fig. 13 Comparison of various algorithms

Figures 12 and 13 shows the comparison of Average *F1*-Score, Average Precision, and Average Recall of various algorithms. From experimental results, it can be found that ANN performs better than other classification algorithms.

5 Conclusion

Kidney takes part in the significant role of filtering the wastes that are accumulated in the blood and its one of the most essential organ of our body. Predicting the reduction in the kidney function at earlier stages is important as it prevents the risk of kidney transplantation. Nowadays, there is an increased use of machine learning and deep learning algorithms in the field of assisting medical experts. In this paper, various classification algorithms such as ANN, K-NN, naïve Baiye's, logistic regression, decision tree, and support vector machine have been used for the forecast diagnosis

of the disease. The dataset for prediction has been taken from UCI machine learning repository. The results obtained from all the above algorithms have been compared. It is perceived that ANN accomplish prediction in a better way than other classification algorithms. This is due to the fact that ANN processes data in the same way as our human brain does. Also, the performance of ANN has been analysed by differing the parameters like number of hidden layers and learning rate which has a considerable impact on the performance of ANN.

References

1. Patil PM, Kamat DK (2019) Chapter 5—Embedded healthcare system for day-to-day fitness, chronic kidney disease, and congestive heart failure. In: U-healthcare monitoring systems. Academic Press, pp 89–117
2. Online: <https://www.kidney.org/kidney-basics>
3. Suresh C, Pani BC, Swatisri C, Priya R, Rohith R (2020) A neural network based model for predicting chronic kidney diseases. In: 2020 second international conference on inventive research in computing applications (ICIRCA), pp 157–162
4. Couser WG, Remuzzi G, Mendis S, Tonelli M (2011) The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. *Kidney Int* 80(12):1258–1270
5. Alnazer I, Bourdon P, Urruty T, Falou O, Khalil M, Shahin A, Fernandez-Maloigne C (2021) Recent advances in medical image processing for the evaluation of chronic kidney disease. *Med Image Anal* 69:101960
6. James MT, Hemmelgarn BR, Tonelli M (2010) Early recognition and prevention of chronic kidney disease. *Lancet* 375(9722):1296–1309
7. Garg A, Mago V (2021) Role of machine learning in medical research: a survey. *Comput Sci Rev* 40:100370. <https://doi.org/10.1016/j.cosrev.2021.100370>
8. Alballa N, Al-Turaiki I (2021) Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: a review. *Inform Med Unlocked* 24:100564
9. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88
10. Zhang H, Hung C-L, Chu W, Chiu P-F, Tang C (2018) Chronic kidney disease survival prediction with artificial neural networks 1351–1356
11. Yildirim P (2017) Chronic kidney disease prediction on imbalanced data by multilayer perceptron: chronic kidney disease prediction 193–198
12. Islam A, Ripon S (2019) Rule induction and prediction of chronic kidney disease using boosting classifiers, Ant-Miner and J48 decision tree. <https://doi.org/10.1109/ECACE.2019.8679388>
13. Bhaskar N, Manikandan S (2020) A computationally efficient correlational neural network for automated prediction of chronic kidney disease. *IRBM* 42. <https://doi.org/10.1016/j.irbm.2020.07.002>
14. Pradeepa P, Jeyakumar MK (2020) An efficient self-tuning spectral clustering algorithm for chronic kidney disease prediction. *Mater Today Proc.* <https://doi.org/10.1016/j.matpr.2020.10.621>
15. Ma F, Sun T, Liu L, Jing H (2020) Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. *Future Gener Comput Syst* 111
16. Almansour NA, Syed HF, Khayat NR, Altheeb RK, Juri RE, Alhiyafi J, Alrashed S, Olatunji SO (2019) Neural network and support vector machine for the prediction of chronic kidney disease: a comparative study. *Comput Biol Med* 109:101–111

17. Akter S, Habib A, Islam Md, Hossen Md, Fahim WA, Sarkar PR, Ahmed M (2021) Comprehensive performance assessment of deep learning models in early prediction and risk identification of chronic kidney disease. *IEEE Access* 1–1. <https://doi.org/10.1109/ACCESS.2021.3129491>
18. Elkholy SMM, Rezk A, Saleh AAEF (2021) Early prediction of chronic kidney disease using deep belief network. *IEEE Access* 9:135542–135549. <https://doi.org/10.1109/ACCESS.2021.3114306>
19. Alloghani M, Al-Jumeily Obe D, Hussain A, Liatsis P, Aljaaf A (2019) Performance-based prediction of chronic kidney disease using machine learning for high-risk cardiovascular disease patients. https://doi.org/10.1007/978-3-030-28553-1_9
20. Revathi M, Godbin AB, Bushra SN, Anslam Sibi S (2022) Application of ANN, SVM and KNN in the prediction of diabetes mellitus. In: 2022 international conference on electronic systems and intelligent computing (ICESIC), pp 179–184. <https://doi.org/10.1109/ICESIC53714.2022.9783577>
21. Al-Shayea Q (2011) Artificial neural networks in medical diagnosis. *Int J Comput Sci* 8:150–154
22. Datasets: https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease
23. Belciug S, Gorunescu F (2014) Error-correction learning for artificial neural networks using the Bayesian paradigm. Application to automated medical diagnosis. *J Biomed Inform* 52. <https://doi.org/10.1016/j.jbi.2014.07.013>
24. Mailagaha Kumbure M, Luukka P, Collan M (2020) A new fuzzy k-nearest neighbor classifier based on the Bonferroni mean. *Pattern Recogn Lett* 140:172–178. <https://doi.org/10.1016/j.patrec.2020.10.005>
25. Wu S, Mao P, Li R, Cai Z-N, Heidari AA, Xia J, Chen H, Mafarja M, Turabieh H, Chen X (2021) Evolving fuzzy k-nearest neighbors using an enhanced sine cosine algorithm: case study of Lupus Nephritis. *Comput Biol Med* 104582. <https://doi.org/10.1016/j.combiomed.2021.104582>. <https://aliasgharheidari.com>
26. Faisal S, Tutz G (2021) Multiple imputation using nearest neighbor methods. *Inf Sci* 570:500–516. <https://doi.org/10.1016/j.ins.2021.04.009>
27. Berrar D (2018) Bayes' theorem and Naive Bayes classifier. <https://doi.org/10.1016/B978-0-12-809633-8.20473-1>
28. Blanquero R, Carrizosa E, Ramirez-Cobo P, Sillero-Denamiel MR (2021) Variable selection for Naïve Bayes classification. *Comput Oper Res* 135:105456. <https://doi.org/10.1016/j.cor.2021.105456>
29. Jamain A, Hand DJ (2005) The Naive Bayes mystery: a classification detective story. *Pattern Recogn Lett* 26:1752–1760
30. Nusinovi S, Tham Y, Chak Yan MY, Wei Ting DS, Li J, Sabanayagam C, Wong TY, Cheng C (2020) Logistic regression was as good as machine learning for predicting major chronic diseases. *J Clin Epidemiol*
31. Takada T, Hoogland J, Lieshout C, Schuit E, Collins G, Moons K, Reitsma J (2021) Accuracy of approximations to recover incompletely reported logistic regression models depended on other available information. *J Clin Epidemiol* 143. <https://doi.org/10.1016/j.jclinepi.2021.11.033>
32. Daga S, Shaikhina T, Lowe D, Briggs D, Higgins R, Khovanova N (2017) Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. *Biomed Signal Process Control*
33. Chen T, Shang C, Su P, Keravnou-Papailiou E, Zhao Y, Antoniou G, Shen Q (2021) A decision tree-initialised neuro-fuzzy approach for clinical decision support. *Artif Intell Med* 111
34. Zoubek L, Penhaker M (2004) D16: decision trees for medical records diagnosis. *IFAC Proc Vol* 37:434–437. [https://doi.org/10.1016/S1474-6670\(17\)30640-7](https://doi.org/10.1016/S1474-6670(17)30640-7)

Robotic Process Automation Powered Admission Management System



Abhishek Gupta, Pranav Soneji, and Nikhita Mangaonkar

Abstract Automation is considered among the most popular solutions to improve efficiency and growth in almost every sector. Admission management is one of the most time-consuming and repetitive tasks that all educational institutes have to carry out every year. The admission process includes a variety of processes that include pen and paper which can lead to some errors. These functions can be carried out using Robotic Process Automation (RPA). Bots can carry out all this work and reduce human errors and increase throughput thus saving money. RPA can check students' entered data and verify it from the uploaded supporting documents, accept students' admission requests, send verification emails, detect discrepancies, and brief the students regarding the same via email. Robotic Process Automation Powered Admission Management System (R.P.A.A.M.S) can carry out the specified operations competently.

Keywords Robotic process automation (RPA) · Admission · Facilitation · Bots · Pdf text extraction · Automate

1 Introduction

Robotic Process Automation can be defined as a support system which automates human work that involves routine tasks [1]. In recent times, many processes and activities have been taking the help of digitalization and are carried out online. The admission process in various professional courses is nowadays done online as well. However, some tasks are still carried out manually such as the details and document verification at the facilitation center [2–7]. The students fill in the application form

A. Gupta (✉) · P. Soneji · N. Mangaonkar
Sardar Patel Institute of Technology, Andheri, Mumbai, India
e-mail: abhishek.gupta2@spit.ac.in

P. Soneji
e-mail: pranav.soneji@spit.ac.in

N. Mangaonkar
e-mail: nikhita.mangaonkar@spit.ac.in

and upload supporting documents. The member of the facilitation center must verify each document with the uploaded documents and the details filled in the form. This process is time-consuming and tedious as there is an enormous number of students seeking admission every year [8–12]. The method often leads to errors due to the frequent human interaction involved. A larger part of the members at the facilitation center often finds it challenging for the above reasons. The process of verification can be taken over by automation. With the help of Robotic Process Automation (RPA), thus this paper aims at assisting the facilitation center members to automate the verification task on hand [13–17].

2 Objective

- (1) Automate the document and details verification process.
- (2) Check the accuracy of the details and supporting documents.
- (3) Increase speed and productivity.

To achieve the aims and objectives of this paper, the emphasis is on identifying the main features that are involved in the system. For automating the process, we will use RPA web scraping, OCR where the user can check the accuracy of the documents and details with minimum effort. The scope covers the objective of reducing human errors and increasing throughput.

R.P.A.A.M.S will use automation anywhere for supporting the automation process. IQ bots and task bots will scrape data from the forms and supporting documents. After comparing the data based on the results, the students will get verification emails or emails containing the errors that need to be updated.

3 Methodology

Admission verification process automation is done because of the following reasons:

- (1) Manually screening documents and verifying them is a time-consuming and tedious process.
- (2) To speed up the process, improve the accuracy, efficiency, and correctness in the process, and successfully perform verification.

3.1 Traditional Method

The process of admission to a professional course has several steps and one of the most important steps is to verify the information submitted by the students against their submitted documents as shown in Fig. 1. The general practice is that the students

fill out the forms online and upload the required proofs/documents. After the initial process, students are required to visit a facilitation center along with the original and duplicate copies of the submitted forms and supporting documents. The member of the facilitation center will check the forms and documents for any data discrepancies. If there are no discrepancies, then the student can progress to the next step. In a scenario where there is any inconsistency between the submitted documents and the corresponding data, the student is required to re-submit the information again, and the same steps are needed to be followed.

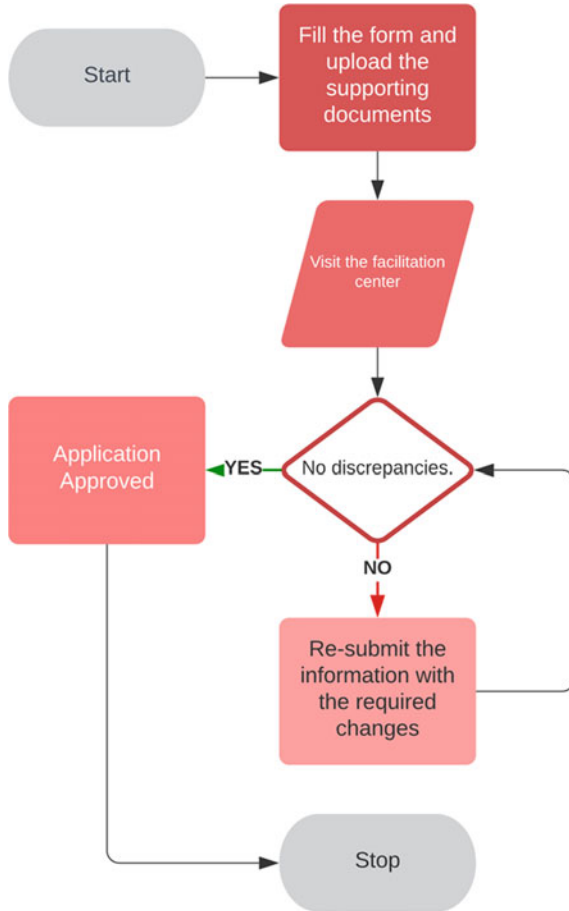


Fig. 1 Flowchart—traditional method

3.2 RPA Method

3.2.1 RPA Architecture

The architecture is shown in Fig. 2.

- (i) **Robotic Process Automation (RPA):** In layman’s terms, RPA is the process by which a software bot uses a combination of automation, computer vision, and machine learning to automate repetitive and high-volume tasks that are rule-based and trigger-driven [2].
- (ii) **Automation Anywhere:** A tool that enables you to design automation processes visually through diagrams.
- (iii) **Automation anywhere** provides us with a digital workforce of bots. The bot mimics rule-based task after recording each step of the process. A workflow is created for the process and different bots work together to execute it multiple times as required. Thus, providing automation.
- (iv) **IQ Bot:** Bots that leverage AI and machine learning from human behavior and bring structure to unstructured data. IQ Bots (Fig. 5) are the brains of the automation anywhere RPA platform. An IQ Bot relies on supervised learning, meaning that every human interaction makes an IQ bot smarter.
- (v) **Task Bot:** The task bots (Fig. 4) are the core of automation that executes repetitive rule-based tasks. Task bots execute the multistep processes with higher accuracy.
- (vi) **Text Extraction using OCR:** It is a technology used to extract text from images and documents via electronic means
 - a. **Preprocessing:** The paper documents or image files are scanned into the software for digitizing. The software works to smooth the edges of letters, remove imperfections, and extract plain text. The remaining text is then turned to black and white only, with all gray shades replaced. This makes text recognition easier, increasing accuracy.

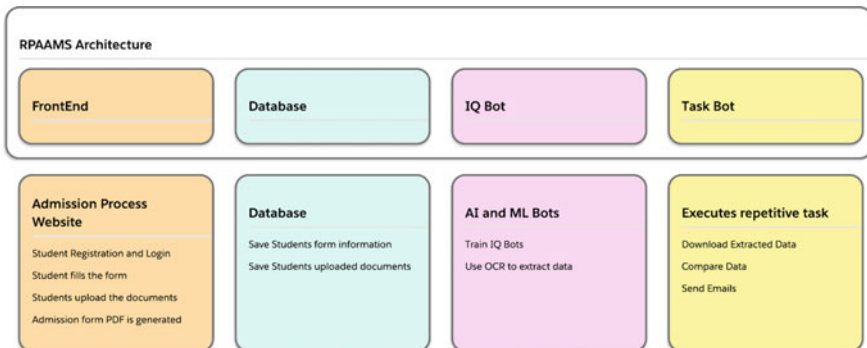


Fig. 2 RPAAMS architecture

- b. Text Recognition: OCR uses various levels of text and pattern recognition, feature detection, and feature extraction, such as the curves and corner patterns unique to each letter, to figure out what the page says.
- c. Postprocessing: Depending on how basic the OCR engine is, it will compare the text to internal dictionaries to cross-reference for context and higher accuracy. The result is a fully searchable and fully editable digital document.

3.2.2 Steps

- (1) Student fills up the form and uploads the required documents. For example, students will have to fill in a name as per the Aadhar card, enter its Aadhar number, enter marks as per the HSC mark sheet, and percentage as per the HSC mark sheet.
- (2) A PDF form is generated with all the filled data, which is used for comparing with the uploaded documents.
- (3) IQ bot will perform data extraction using the OCR recognition method on the form documents. The next process is to select important regions of document images to OCR them with adjusted parameters for these particular tasks [3]. Once the bot is provided with the file, the OCR engine performs text extraction. For IQ bot to extract the necessary data during training, it has a few sample documents where IQ bot is specified where and which data is to be extracted also, we can provide the logic to extract data in specific patterns only. Results are downloaded from the server and stored in an excel format file.
- (4) The task bot starts its process and checks if the data submitted matches the proof submitted. Here the task bot will read each field from the application form, which was extracted before, and it will match with the extracted field of the supporting document.
- (5) Auto-generated email is also sent to applicants stating the status of the verification process either success or failure (Fig. 7). The reason for rejection is also mentioned in the email automatically. Figure 3 shows the flowchart on RPA method.

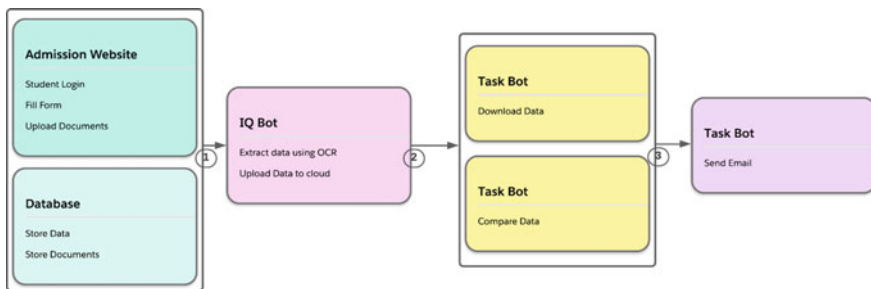


Fig. 3 Flowchart—RPA method

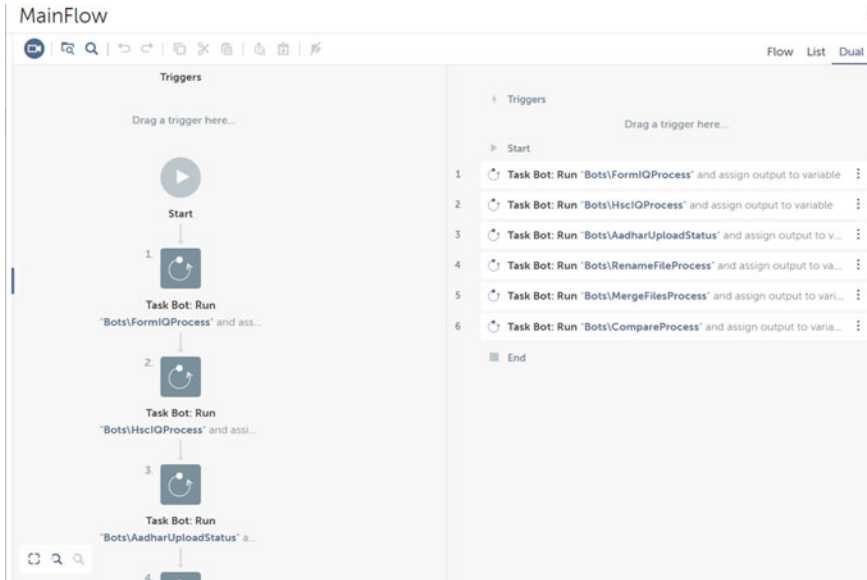


Fig. 4 Flow of task bots

Instance name	II	Provider	# of IQ Bots	# of Files	Training %	Actions
AadharExtractBot		Abbyy OCR	1	0	100%	
AadharExtractBotOne		GoogleVisionAPI OCR	1	10	0%	
FormExtractBot		Abbyy OCR	1	75	100%	
HscExtractBot		Abbyy OCR	1	20	100%	

Fig. 5 Different IQ bots

4 Results

The following survey as in Table 1 was conducted among members of the facilitation center to understand the parameters of the traditional document verification process.

5 Comparison

Robotic Process Automation is a modern approach to automating rule-based tasks. Here we will look at the comparison between the traditional approach and the modern approach of document verification during the student admission process. In the traditional approach, students must visit the facilitation center for verification of details

Table 1 Survey result

Parameter	Result	Conclusion
In the process, do you have to manually verify fields	0.9166	91.66% of participants agree
It takes more than 5 min to process a complete candidate form using the current facilitation system	0.8277	82.77% of participants agree
Do you think this manual verification process is tedious	0.9944	99.44% of participants agree
As this process requires manual intervention, would you like to have a system that automates complete processes with minimal human interaction	0.7611	76.11% of participants agree
Do you think the manual verification processes lead to errors	0.6611	66.11% of participants agree
No more than 30 candidates from processing can be done in 1 h	0.8444	84.44% of participants agree
Do you find it challenging to verify the contents of documents	0.8	80.00% of participants agree

provided by them and a member of the facilitation center will have to manually verify the details. Using RPA, this system will do verification using bots with better accuracy and speed. The survey shows that the current traditional method does not have any automation. With RPA, these processes will be automated. From the results, it's clear, that more than 76% of participants want an automated system. It is safe to say that the time taken for verification of a single student is more than two minutes as the survey indicates that not more than 30 students' verification can be done within an hour. However, in the RPA-supported system, the average time per student is about 1 min. As most participants, almost 100% feel that the traditional method is tedious our RPA-based system can help in the process using bots and OCR techniques. RPA IQ bots use machine learning to improve the efficiency and accuracy of the process since the traditional method is prone to errors the proposed system can help reduce them. In the sample data used to train and test the system, negligible errors were found. About 80% of the participants in the survey agreed that, they found the manual verification process challenging. In contrast, RPA aims to make processes like these seamless. In the proposed system, the member of the facilitation center only needs to start the process, and the rest will be taken care of automatically. The results on bots are shown in Figs. 4, 5, 6, 7 and 8.

6 Conclusion

Automation is the future. Most of the monotonous tasks are completed using automation nowadays. The presence of numerous RPA products on the market, many of which are market leaders, indicates that this technology will continue to emerge in the future years [4]. Through process automation, RPA bots can carry out the processes to help ensure that the systems run efficiently on a 24/7 basis. Robotic

<input type="checkbox"/>	Type ↑	Name	Status
<input type="checkbox"/>	Task Bot	DemoBot	New
<input type="checkbox"/>	Task Bot	AadharExtract	New
<input type="checkbox"/>	Task Bot	HSCImageExtract	New
<input type="checkbox"/>	Task Bot	FormIQProcess	New
<input type="checkbox"/>	Task Bot	AadharUploadStatus	New
<input type="checkbox"/>	Task Bot	HscIQProcess	New
<input type="checkbox"/>	Task Bot	RenameFileProcess	New
<input type="checkbox"/>	Task Bot	MergeFilesProcess	New
<input type="checkbox"/>	Task Bot	CompareProcess	New
<input type="checkbox"/>	Task Bot	MainFlow	New

Fig. 6 Different task bots

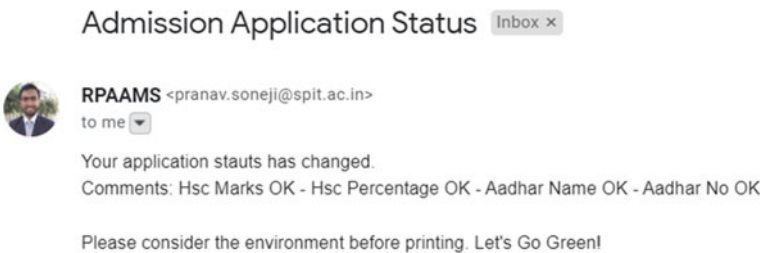


Fig. 7 Email notification after the process is complete

Process Automation or RPA gives you the ability to save time and money by having your digital workforce. Student Data Verification is a manual and paper-heavy and monotonous process that is required to be completed by universities, schools, and colleges. Students need to fill out the form and upload the required documents, and this system can take care of the verification process using IQ bots, task bots, and OCR

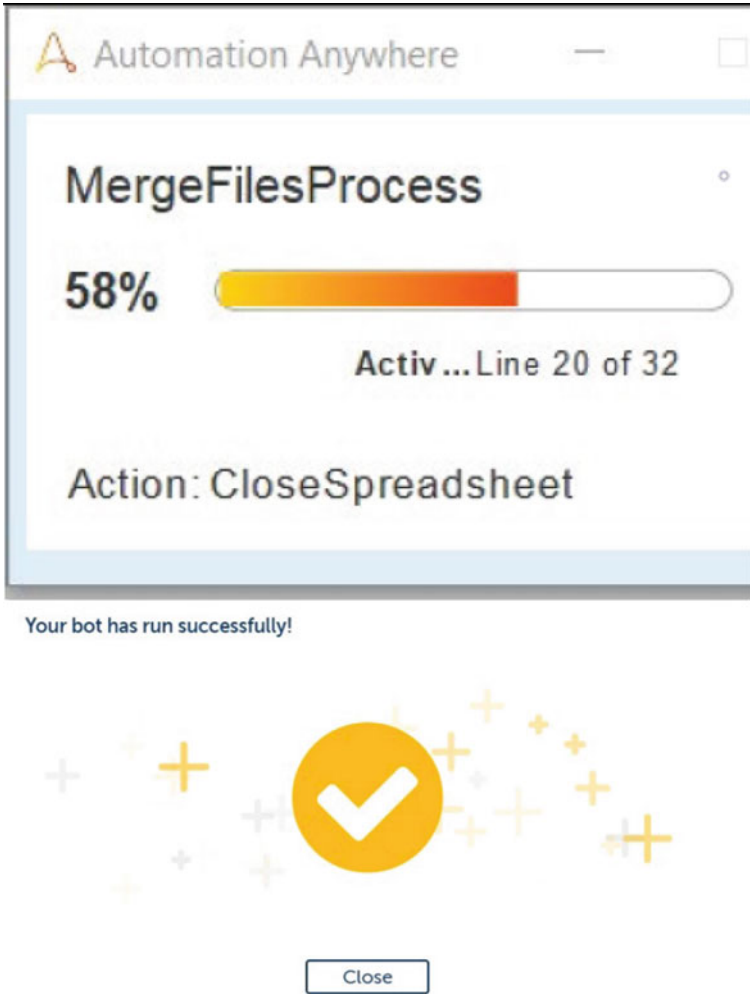


Fig. 8 Different task notification

technique. Along with this, it will also send emails to the students to inform them if their form was submitted or if any changes are required. The system can be extended to cater more validation of various types of documents with respect to the inputs given by the user. Features like advanced pattern matching for removing unwanted characters around the real data, comparing graphical data to obtain information from documents, and finding the similarity between data.

References

1. Aguirre S, Rodriguez A (2017) Automation of a business process using robotic process automation (RPA): a case study. In: Workshop on engineering applications. Springer, pp 65–71
2. <https://enterpriseproject.com/article/2019/5/rpa-robotic-process-automation-how-explain>
3. Wróblewska A, Stanisławek T, Prus-Zajączkowski B, Garncarek L (2018) Robotic process automation of unstructured data with machine learning. In: Federated conference on computer science and information systems
4. Afriliana N, Ramadhan A (2022) The trends and roles of robotic process automation technology in digital transformation: a literature review 12:51–73. <https://doi.org/10.33168/JSMS.2022.0303>
5. Wang H, Smys S (2021) A survey on digital fraud risk control management by automatic case management system. *J Electr Eng Autom* 3:1–14
6. Rajat, Verma K, Kumar V, Singh P (2020) Automated order management using robotic process automation. In: Proceedings of the international conference on innovative computing & communications (ICICC)
7. Henry LM, Pedro PS (2020) Digital transformation in higher education institutions with business process management: robotic process automation mediation model. In: 15th Iberian conference on information systems and technologies (CISTI)
8. Shidaganti G, Salil S, Anand P, Jadhav V (2021) Robotic process automation with AI and OCR to improve business process: review. In: Second international conference on electronics and sustainable communication systems (ICESC)
9. Kobayashi T, Arai K, Imai T, Tanimoto S, Sato H, Kanai A (2019) Communication robot for elderly based on robotic process automation. In: IEEE 43rd annual computer software and applications conference (COMPSAC)
10. Guleria K, Sharma A (2021) A framework for hotel inventory control system for online travel agency using robotic process automation. In: International conference on advance computing and innovative technologies in engineering (ICACITE)
11. Ketkar Y, Gawade S (2021) Effectiveness of robotic process automation for data mining using UiPath. In: International conference on artificial intelligence and smart systems (ICAIS)
12. Mohammed ST, Fatima B, Shukla VK, Chopra A (2021) Robotic process automation in banking and finance sector for loan processing and fraud detection. In: 9th international conference on reliability, Infocom technologies and optimization (trends and future directions) (ICRITO)
13. Kajrolkar A, Pawar S, Paralikar P, Bhagat N (2021) Customer order processing using robotic process automation. In: International conference on communication information and computing technology (ICCICT)
14. Domínguez-Mayo FJ, Enríquez JG, García-García JA, Enríquez JG (2014) Robotic process automation: a scientific and industrial systematic mapping study. *IEEE Access* 839113–39129
15. Rodriguez A, Aguirre S (2017) Automation of a business process using robotic process automation (RPA): a case study. In: Figueroa-García JC, López-Santana ER, Villa-Ramírez JL, Ferro-Escobar R (eds) Communications in computer and information science. Applied computer sciences in engineering. Springer International, Cham, pp 65–71
16. Van der Aalst WMP, Bichler M, Heinzl A (2018) Robotic process automation. *Bus Inf Syst Eng* 60:269–272
17. Choudhary R, Karmel A (2022) Robotic process automation. *Artif Intell Technol* 29–36

Implementation of Advanced High Performance Bus to Advanced Peripheral Bus Bridge



Chinta Sai Manasa, Navya Mohan, and J. P. Anita

Abstract The Advanced Microcontroller Bus Architecture (AMBA) is a standard interconnect protocol that allows all the components in system on chip (SOC) designs to be connected and managed. The SOC has Advanced High Performance Bus (AHB) which is for high performance system modules and Advanced Peripheral Bus (APB) which is for low performance system modules. In order to develop the internal communication between high performance peripherals and low performance peripherals, AHB to APB bridge has been implemented for converting AHB transactions into APB transactions. The AHB to APB bridge is made up of four modules: an AHB master, an AHB slave interface, an APB controller, and an APB interface. These modules were written in Verilog Hardware Description Language (VHDL) and tested on a Questa Sim-64 10.7c and also code coverage report has been generated.

Keywords Advanced microcontroller bus architecture · System on chip · Advanced high performance bus · Advanced peripheral bus · Code coverage

1 Introduction

AMBA is a bus architecture used on chip buses that is commonly used in SOC architectures. For creating high level embedded microcontrollers, the AMBA specification standard is employed. The primary goal of AMBA is to guarantee technological independence and promote modular system architecture. It also creates reusable devices while reducing silicon.

AMBA is an open standard that defines a technique for managing the components that build the SOC architecture. To increase the device performance, the high speed bus uses multiple masters. During down time, AMBA requirements can make advantage of system bus bandwidth.

AHB is a new version of AMBA that is optimized for high speed designs. It has large bandwidth and the ability to support many bus masters. APB is a bus that is part

C. S. Manasa · N. Mohan · J. P. Anita (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: jp_anita@cb.amrita.edu

of the AMBA. It is designed to consume less power and have a simpler interface. APB has the advantages of achieving high frequency, obtaining simplified static timing analysis using single clock edge, and automatic test insertion is easily integrated with cycle-based simulators.

The SOC is composed of many different components, some are high speed and some are low speed. The AHB is a high speed bus that can communicate with the high speed components. The APB bus, which is a low speed bus, is used to communicate with the low speed components. If high performance components need to communicate with low performance components, they require one data converter that is to convert AHB data to APB data so that high performance components and low performance components can communicate. As a result, an AHB to APB bridge is developed in the SOC for internal communication, converting AHB transactions to APB transactions. The design is primarily made up of four blocks: AHB master, which initiates read and write operation; AHB slave interface, which responds to read and write operation; APB controller, which converts AHB transactions into APB transactions; APB interface produces APB transactions as output.

Further sections in the paper are organized as follows: Section 2 consists of literature survey, followed by proposed work is in Sect. 3, then simulation results in Sect. 4, and followed by conclusion in Sect. 5.

2 Literature Survey

The AMBA is a system on chip interconnection standard for high and low performance peripherals. If peripherals on the AHB side want to communicate with peripherals on the APB side, a bridge is essential between them. A bridge has been implemented to provide communication between AHB and APB in VLSI Design of AMBA Based AHB2APB Bridge [1]. The bridge is designed by using handshaking signals for reducing data loss. In [2], timer concept has been added to overcome the data loss during the transfer in implementation of AMBA based AHB2APB bridge. To achieve the high performance in the design, FSM-based pipelined APB to APB bridge is developed in [3]. In [4], a verification code is written in Universal Verification Language (UVM) and using various test cases, all functions of the Bridge protocol are verified. The code coverage and functional coverage are also analyzed. Clock is a major concern in the design of any digital sequential system. Hence in [5], a bridge with clock skew minimization technique is developed by using a ripple counter. A design technique for testing the SOC is developed by an on and off chip bus bridge in [6]. A flexbus is implemented to adjust the logical connectivity of communication architectures and its components in [7]. AHB to APB has been implemented on FPGA in [8]. For establishing communication between APB and I2C, a bridge is developed in [9]. An interface is implemented between AXI and APB in [10]. In [7], different bus topologies are defined for on chip communication. The on chip communication architecture is a key factor of the overall performance in complex SOC designs. The IIC protocol is verified using UVM and IIC controller

design is written in Verilog [11]. Different functional blocks in the 80,186 processor are designed using Verilog and its functionality is verified using UVM in [12]. Bus bridge is implemented using Open Core Protocol (OCP) and I2C protocol in [13]. A testbench is built for APB-based AHB interconnect, using `uvm_config_db` in [14–16]. For the bus-based Vehicular ad hoc Network (VANET), a street-centric routing algorithm is employed to minimize relay bus and route selection problems [17].

3 Proposed Work

In the SOC, for internal communication, the proposed bridge is used. The SOC contain high and low performance devices. High performance components include the ARM Processor, On Chip RAM, Direct Memory Access, and Memory Interface, which are all connected by the AHB. Timer, UART, PIO, and keypad are low performance devices that are connected by APB. AHB to APB bridge is required when high performance devices need to communicate with low performance devices. The bridge converts AHB transactions into APB transactions using pipelining concept, where extra address register and data registers where added, through which address and data pass from AHB to APB. Figure 1 shows the typical AHB to APB bridge.

3.1 AHB to APB Bridge Architecture

The architecture mainly constitutes of four modules, they are AHB master, AHB slave interface, APB controller, and APB interface. The AHB master will communicate with one slave at a time out of three slaves based on address. Figure 2 represents the AHB to APB bridge architecture.

AHB Master: This module will initiate the read and write transfers by giving address and control information.

AHB Slave interface: Within a particular address range, this module will respond to a read or write operation.

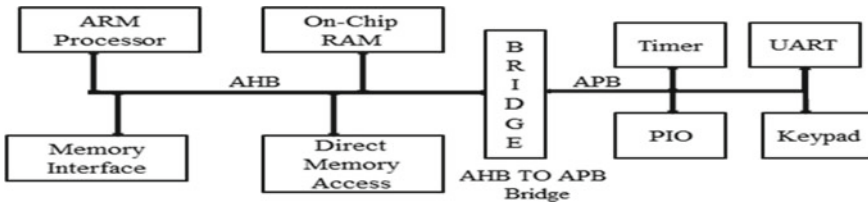


Fig. 1 Typical AHB to APB bridge [2]

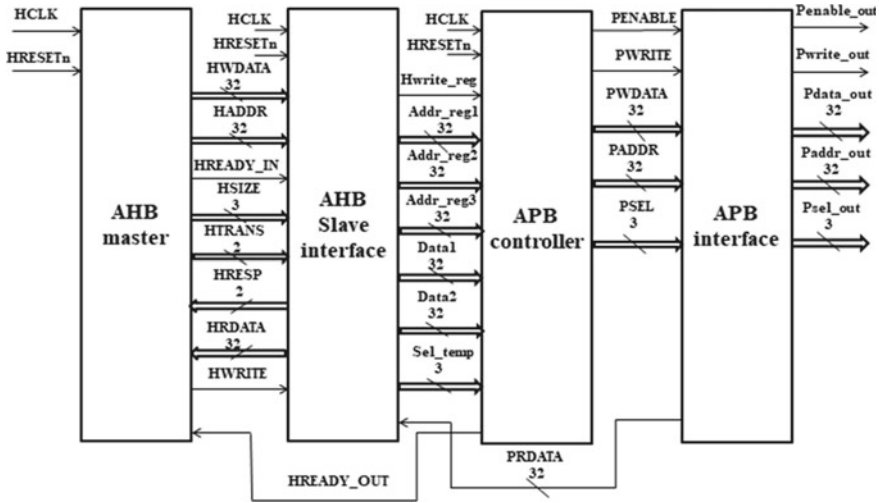


Fig. 2 AHB to APB bridge architecture

APB Controller: In this module, AHB transactions or signals are converted to APB transactions or signals. Refer Fig. 3 for finite state machine of AHB to APB interface.

APB Interface: The APB interface receives APB transactions from the APB controller and provides them as output.

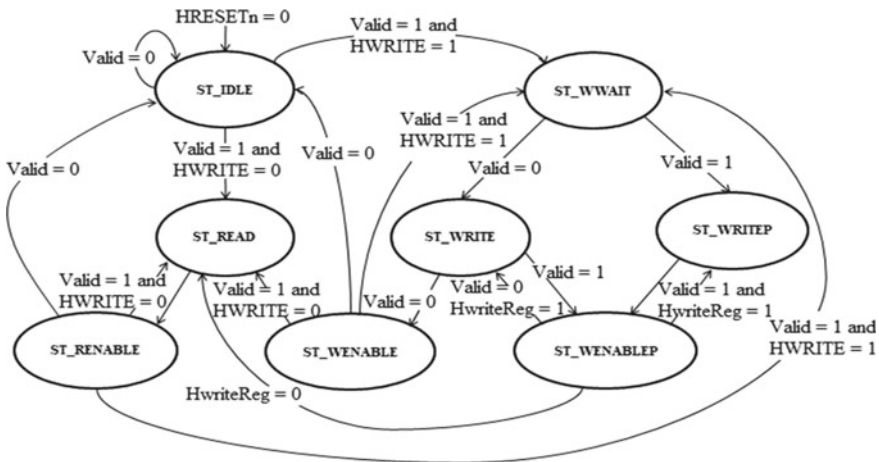


Fig. 3 FSM for AHB to APB interface [2]

Table 1 Size encoding

HSIZE (3-bit)	Type
000	BYTE
001	HALF_WORD
010	WORD

Table 2 Transfer type encoding

HTRANS (2-bit)	Type	Description
00	IDLE	No data transfer
01	BUSY	Master is continuing with transfer
10	NONSEQ	For single transfer or first transfer
11	SEQ	For multiple transfer

3.2 AHB Signal Description

HCLK: It's a clock signal; this clock keeps track of all bus transfers.

HRESETn: This signal reset the system.

HWDATA: This is a 32-bit signal used to transfer data from master to slave during write operation.

HADDR: It holds the 32-bit of address value.

HREADY_IN: It denotes that the transfer has been activated.

HREADY_OUT: It indicates that the transfer is finished.

HRESP: The OKAY response is always generated by the transfer response, which offers further information about the status of a transfer.

HSIZE: It specifies the size of the data to be transferred. Refer Table 1 for size encoding.

HRDATA: This is a 32-bit signal that transfer read data from slave to master during read operation.

HWRITE: It represents write and read transfers; when high, write transfer occur; when low, read transfer occur.

HTRANS: Specifies current transfer's type. Refer Table 2 for transfer type encoding.

3.3 APB Signal Description

PENABLE: This enable signal is used to time all peripheral bus accesses.

PWRITE: It represents write and read transfers; when high, it represents write transfer; when low, it represents read transfer.

PWDATA: It stores 32-bit write data.

PADDR: This signal stores 32-bit address value.

Table 3 Slave selection

PSEL (3-bit)	Address	Slave
001	$Haddr \geq 32'h80000000$ to $Haddr < 32'h84000000$	Slave1
010	$Haddr \geq 32'h84000000$ to $Haddr < 32'h88000000$	Slave2
100	$Haddr \geq 32'h88000000$ to $Haddr < 32'h8c000000$	Slave3

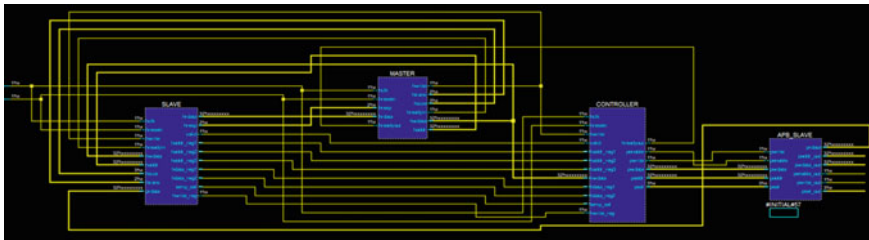


Fig. 4 AHB to APB bridge RTL schematic

PRDATA: This signal stores 32-bit read data.

PSEL: Based on the address, slave get selected. Refer Table 3 for slave selection.

4 Simulation Results

4.1 RTL Schematic of AHB to APB Bridge Design

The Register Transfer Language (RTL) diagram depicts the logic implementation and data flows in and out of a circuit. Figure 4 represents the RTL schematic of AHB to APB bridge design.

4.2 Single Write Transfer

During write operation, the data is transferred from master to slave. A data bus width of 32 bits is recommended as a minimum. On the other side, this may easily be expanded to accommodate larger bandwidth. Refer Fig. 5 to check the waveform of single write transfer.

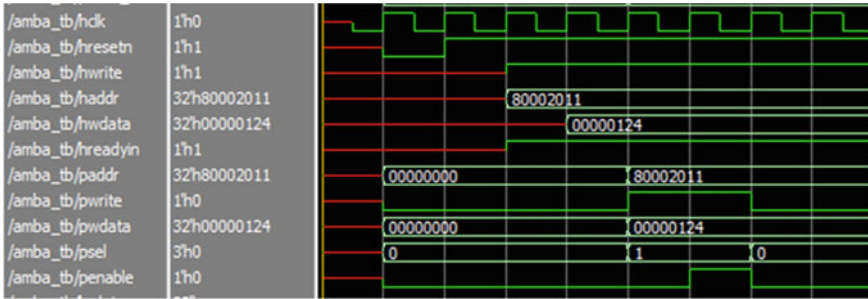


Fig. 5 Waveform of single write transfer

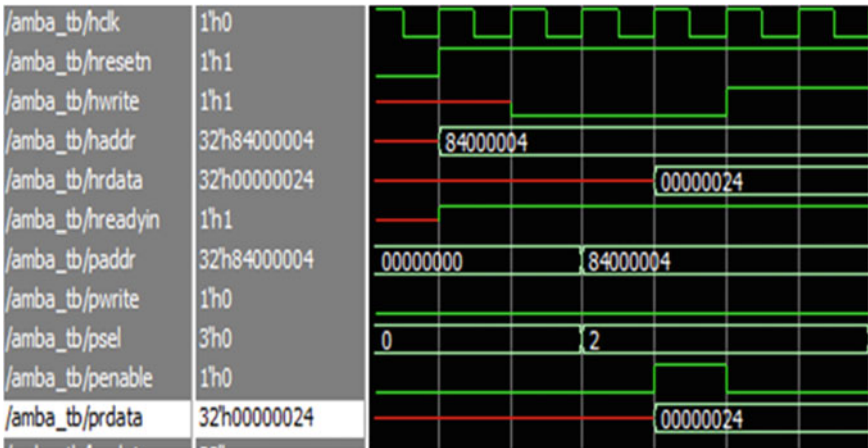


Fig. 6 Waveform of single read transfer

4.3 Single Read Transfer

During read operation, the data is transferred from slave to master. Width recommended is 32 bits. For higher bandwidth operation, data width can be easily extended. Refer Fig. 6 to check the waveform of single read transfer.

4.4 Burst Write Transfer

In burst transfer, HTRANS indicate NONSEQ for first transfer, it indicate SEQ, in SEQ type, the address depends on previous address and control information is similar to previous transfer. Here, multiple write data is transferred from master to slave. Refer Fig. 7 to check the waveform of burst write transfer.

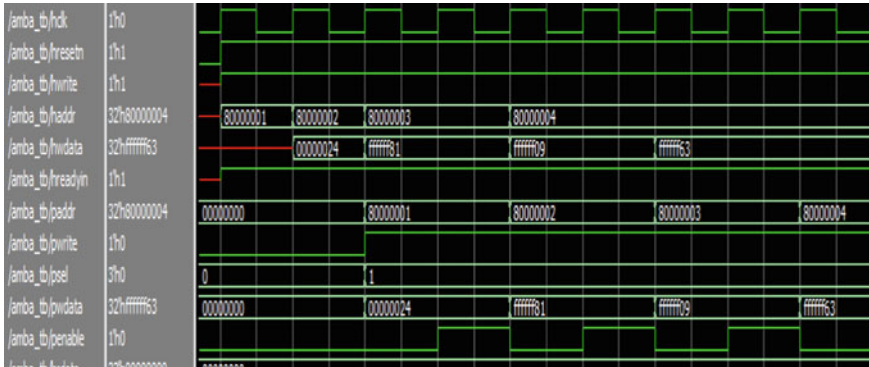


Fig. 7 Waveform of burst write transfer

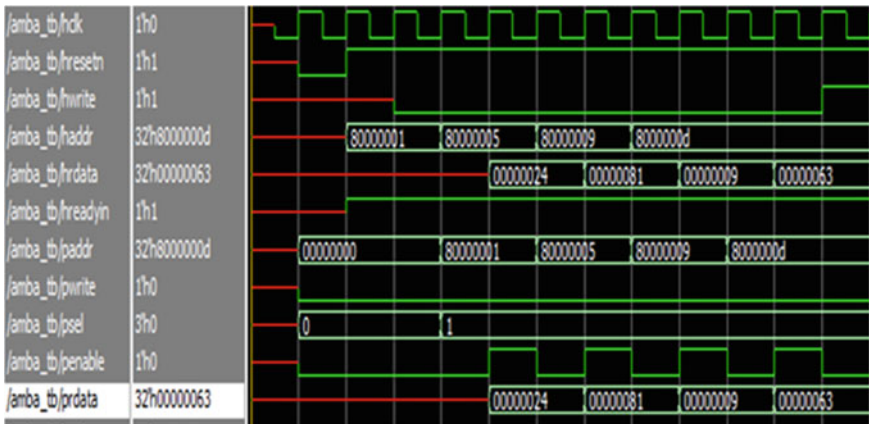


Fig. 8 Waveform of burst read transfer

4.5 Burst Read Transfer

For initial transfer, HTRANS represents NONSEQ and for remaining transfers, it represents SEQ, where address depends on previous address and the control information is similar to previous transfer. Here multiple read data is transferred from slave to master. Refer Fig. 8 to check the waveform of burst read transfer.

4.6 Write-Read Transfer

Write-read transfer is nothing but a back to back transfer where one write transfer and one read transfer occur alternatively. The write data is transferred from master to

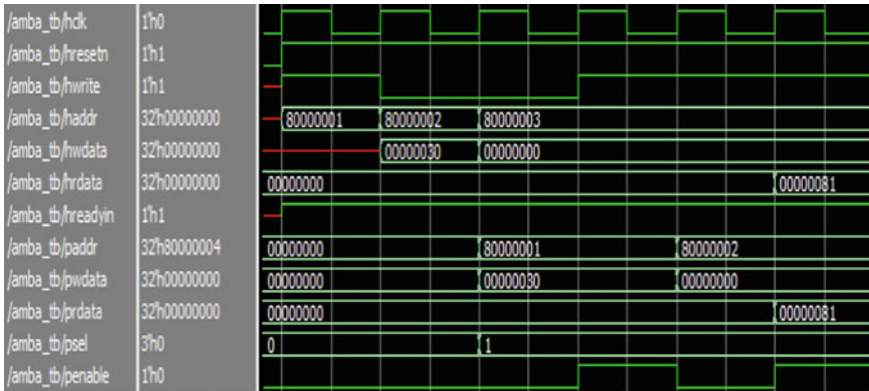


Fig. 9 Waveform of write–read transfer

slave, when **hwrite** is high. Read data is transferred from slave to master in the design, when **hwrite** is low. Refer Fig. 9 to check the waveform of write–read transfer, where one write transfer and one read transfer is shown.

4.7 Code Coverage Report

Code coverage is a measure of design code that is how many lines of code is getting executed successfully. Refer Fig. 10 for code coverage report of AHB to APB bridge design.

Total Coverage:						62.03%	75.18%
Coverage Type	Bins	Hits	Misses	Weight	% Hit	Coverage	
Statements	100	93	7	1	93.00%	93.00%	
Branches	72	61	11	1	84.72%	84.72%	
FEC Conditions	35	24	11	1	68.57%	68.57%	
Toggles	1454	852	602	1	58.59%	58.59%	
FSMs	30	19	11	1	63.33%	71.02%	
States	8	7	1	1	87.50%	87.50%	
Transitions	22	12	10	1	54.54%	54.54%	

Fig. 10 Total code coverage of AHB to APB bridge design

5 Conclusion

A system on chip integrates a variety of components, such as high and low performance components. A bridge is required to provide communication between high and low performance components. Hence, AHB to APB bridge is designed, where AHB transactions are converted to APB transactions. The AHB to APB bridge is designed for all types of transfers using Verilog. The bridge design is verified in Questa Sim and code coverage of the AHB to APB bridge design is generated, that is to know how much percentage of code is getting executed successfully. In the future, this design is verified using System Verilog or UVM for getting accurate results.

References

1. Kharade A, Jayashree V (2018) VLSI design of AMBA based AHB2APB bridge. *Int J VLSI Des Commun Syst* 9:19–30. <https://doi.org/10.5121/vlsic.2018.9302>
2. Panchal B, Parmar Y, Suthar H (2020) Implementation of AMBA based AHB2APB bridge. *Int J Recent Technol Eng (IJRTE)* 8(6)
3. Sairam RM, Sharma S, Pal G (2019) FSM ABD handshaking based AHB2APB bridge for high speed system. *IJERT* 2
4. Prasad NGN (2017) Development and verification of AHB2APB bridge protocol using UVM technique. *IJSRR* 6
5. Kiran Kumar M, Sajja A, Noorbasha F (2017) Design and FPGA implementation of AMBA APB bridge with clock skew minimization technique. *IOSR JVSP* 7(3)
6. Jaehoon Song Y et al (2009) An efficient SOC test technique by reusing on/off-chip bus bridge. *IEEE Trans Circuits Syst I Regul Pap* 56(3)
7. Sekar K et al (2008) Dynamically configurable bus topologies for high performance on-chip communication. *IEEE Trans Very Large-Scale Integr (VLSI) Syst* 16(10)
8. Aithal S, Baligar JS, Guruprasad SP (2016) FPGA implementation of AHB to APB protocol. *IJSR* 5(5)
9. Chajagauda SM, Gubbi A (2016) Using Verilog and system Verilog design and verify communication bridge between APB and I2C protocol. *IJSTE* 3(1)
10. Usha K, Ashok Kumar T (2013) AXI to APB interface design using Verilog. *IOSR J Electron Commun Eng* 8(5):1–9. p-ISSN: 2278-8735
11. Jaideep Varier EV, Prabakar V, Balamurugan K (2019) Design of generic verification procedure for IIC protocol in UVM. In: 3rd international conference on electronics, communication and aerospace technology (ICECA), pp 1146–1150
12. Rajthilak S, Mohankumar N (2021) UVM based verification of iAPX 186 processor modules. *J Phys Conf Ser* 1921:012053. <https://doi.org/10.1088/1742-6596/1921/1/012053>
13. Bhakthavathchalur R, Deepthy GR, Vidhya S, Nisha V (2011) Analysis of low power open core protocol bridge interface using VHDL. In: *IEEE recent advances in intelligent computational systems*, pp 357–362
14. Dohare N, Agrawal S (2016) APB based AHB interconnect testbench architecture using `uvm_config_db`. *Int J Control Theor Appl* 9:4377–4392
15. Palnitkar S (2008) *Verilog HDL: a guide to digital design and synthesis*, 2nd edn. Pearson
16. Bergeron (2009) *Writing testbench using system Verilog*. Springer
17. Dhaya R, Kanthavel R (2021) Bus-based VANET using ACO multipath routing algorithm. *J Trends Comput Sci Smart Technol (TCSST)* 3(1)

The Influence of Digital Financial Literacy on Interest in Investing in the Capital Market of Millennial Generation Students



Ford Lumban Gaol, Christopher, Gerry Dermawan,
Muhammad Resnandya Raffi, Rully Rubianto, and Tokuro Matsuo

Abstract In today's environment, financial knowledge is vital and should be possessed by everyone. Numerous people are financially disadvantaged as a result of the COVID-19 pandemic. This serves as a reminder of the critical nature of financial investment in the future. However, many Indonesians, particularly students, are still unfamiliar with capital market investing. In reality, investing in the stock market can provide a variety of benefits, including financial security, capital appreciation, and passive income. Financial literacy concerns can result in a lack of financial awareness among pupils. Indeed, financial literacy is widely available online and can be accessed at any time and from any location. We employed a quantitative method with a non-experimental research design for this study. A questionnaire survey is used to obtain data. The goal of this study is to determine whether there is a correlation between student interest in capital market investment and their digital financial literacy. As a result, extensive action is required to increase student interest in the capital market, which is precisely what this study wants to do. According to our findings, students who have access to digital financial and investment information have increased their financial literacy and developed a greater interest in investing in the capital market. Individuals are also encouraged by social media trends to try

F. L. Gaol (✉)

Doctor of Computer Science, Binus Graduate Program, Jakarta, Indonesia

e-mail: fgaol@binus.edu

Christopher · G. Dermawan · M. R. Raffi · R. Rubianto

School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

e-mail: christopher020@binus.ac.id

G. Dermawan

e-mail: gerry.dermawan@binus.ac.id

M. R. Raffi

e-mail: muhhammad.rafi004@binus.ac.id

R. Rubianto

e-mail: rully.rubianto@binus.ac.id

T. Matsuo

Advanced Institute of Industrial Technology, Tokyo, Japan

e-mail: matsuo@aait.ac.jp

their hand at investing in the stock market. Additionally, students prefer to invest in stocks and mutual funds over other investing options.

Keywords Digital financial literacy · Capital market investment · Investing · Interest

1 Introduction

All information is easily accessible in the present era, whether online or offline. We can now obtain information from all around the world in a matter of seconds thanks to the Internet. Digital technology development is accelerating in several countries. In Indonesia, the digital era is defined by the growing use of the Internet by society [1]. Apart from the Internet, there are numerous sources of information available through a variety of educational programs, social networking platforms, and online streaming applications such as YouTube. Social media is one of the primary sources of knowledge for the younger generation, particularly during this COVID-19 time, with many of these sources being social media influencers talking “investment.” [2] Investments are currently being discussed on social media platforms such as Instagram, YouTube, and TikTok [3] ranging from traditional capital market investments such as stocks, mutual funds, and bonds to alternative investments such as cryptocurrency or even those that are gaining popularity on social media, such as NFT. Several of these influencers educate their followers on safe investments, particularly in the capital market, and provide an excellent introduction to capital market investing, piquing the desire of a large number of other social media users to invest [2]. However, some influencers provide inaccurate information about investing, such as providing only information about returns and predictions of which stocks will increase in value without conducting an analysis first and failing to inform them about the risks associated with investing and how to determine which stocks are worth investing in. It results in the loss of a large number of people.

There is a wealth of information available on the Internet about investing, particularly in the capital market. This information ranges from an introduction to investing in the capital market, such as what is a stock, to how to choose good stocks and perform stock analysis using expert-developed formulas and ratios [4]. It is documented in a journal that a small amount of capital has a significant impact on students’ investment interests. As a result, as the required capital for investing decreases, the more interested investors get [5]. Thus, we will investigate if the presence of this digital literacy can drive students to invest, particularly in the capital market, through this research.

There are several research questions as to the problem definition that we will find out the answer.

1. Does the influence of their increased financial literacy make them interested in investing in the capital market?

2. What are the most widely used investment instruments for millennial generation students in investing in the capital market?

1.1 Survey on Financial Literacy and Investment Activities in Indonesia

Public financial literacy has grown by 8.33% in the last three years, according to OJK's third national financial literacy research released in 2019. In 2019, Indonesia's financial literacy index was at 38.03%. Meanwhile, in 2016, the financial literacy score increased to 29.7%. Additionally, Indonesia's capital market sector had a growth of 232,596 capital market investors in 2019. There were 1,084,836 investors in Indonesia's stock exchanges. This is a significant increase from the 852,240 capital market investors in Indonesia in 2018 [6]. According to the poll, financial literacy and investing activity have increased significantly among Indonesians. This suggests that Indonesians are developing a greater sense of financial responsibility. A yearly rise would benefit the Indonesian people's economy and could contribute in the development of economic progress.

1.2 Financial Technology

Financial technology, or FinTech, has been known to trend on social media and the Internet [7]. Global investment in FinTech startups reached a record high of 24.1 billion dollars in 2016 [8]. FinTech can be defined as a cross-disciplinary issue that integrates technology usage management, finance, and innovation management in order to optimize financial service operations through the use of technological solutions adapted to the unique demands of each firm. FinTech, as defined by Dahlberg in Razzaque and Hamdan [9], is a term that refers to a combination of finance and technology that is given automatically and effectively while also demonstrating cost effectiveness. Additionally, such financial services leverage contemporary technology, particularly FinTech via mobile, which enables entrepreneurs and end users to manage their financial services through the use of pre-developed algorithms saved in the cloud.

1.3 Digital Financial Literacy

Financial literacy refers to the capacity to grasp, analyze, manage, and explain personal financial matters. It is a collection of skills and knowledge that enables an individual to make sound and efficient financial decisions. Additionally, it relates to the capacity to make sound financial choices and manage money prudently [10].

Financial literacy affects everyone; people who lack financial literacy will be unable to plan for retirement and will have a reduced ability to invest in equities [11]. Chen and Volpe [12] conducted a survey of 924 college students to assess their personal financial literacy in making financial decisions. The research found that people accurately scored 53% of the time on average. According to research, individuals under the age of 30, those with limited work experience, those with non-business degrees, and women have a lower level of financial literacy. Individuals with a lesser level of financial literacy and investment acumen are more likely to make poor investment decisions. Digital Financial Literacy includes all of these facets. Morgan et al. [13] define digital financial literacy as the understanding of digital financial products and services, awareness of digital financial risks, knowledge of digital financial risk management, and knowledge of consumer rights, and redress procedures. The first dimension is familiarity with digital financial products and services; this dimension incorporates a fundamental understanding of these products and services. Individuals should be cautious of novel financial services and products offered via digital channels such as the Internet and mobile phones. The second part of digital financial literacy is an awareness of digital financial threats. Individuals and companies must be aware of these possible risks when employing Digital Financial Services which may be more diverse than traditional financial products yet are becoming increasingly difficult to detect. The third component of digital financial literacy is digital financial risk management, which refers to customers' awareness of how to safeguard themselves against any dangers associated with it. They should be able to prevent spamming, scamming, and other forms of fraud through the use of computer software and phone applications. They should probably learn how to preserve their personal identification number (PIN) and other sensitive information when employing digital banking services. When DFS customers become victims of the aforementioned dangers, the final aspect to consider is consumer rights and redress procedures understanding. DFS users who become victims of cybercrime or other losses must be aware of their rights, as well as where to go and how to seek restitution. Additionally, they should be aware of their rights regarding personal information and how to seek redress if it is misused. Financial knowledge is crucial, even more so in the digital age. These four criteria must be completed in order for an individual to possess a high level of digital financial literacy.

1.4 Capital Market Investment

A capital market is an exchange where individuals and institutions can buy and sell financial assets such as bonds, shares, and mutual funds. Stock is a sort of financial instrument that denotes ownership in the issuing company. A corporation sells shares to raise capital for the purpose of expanding or operating their business. The owner of the stock receives the corporation's assets and profits in proportion to the number of shares they own [14].

Mutual funds are a type of financial instrument in which investment managers collect funds from the public and invest them in a variety of other investment products in order to generate capital gains or other benefits for the public as investors, with profits distributed to investment managers as well [15].

Bonds are a type of financial instrument in which investors (both public and private) lend money to a business or government for a certain term in exchange for monthly interest payments. When the bond matures, the investor receives a reimbursement from the bond issuer. Bonds are categorized as fixed income investments since they make regular payments over the course of the bond's existence [16].

The capital market connects investors, businesses, and government agencies through the trade of long-term products such as bonds, equities, and other securities. Banks, firms, and issuers are examples of suppliers, while businesses, government, and individuals are examples of purchasers. There are numerous advantages for investors and issuers alike. Our research is geared toward investors, as they stand to gain benefits such as increased investment value as a result of economic expansion. This was reflected in growing stock prices, which resulted in capital gains and dividend payments to stockholders, as well as floating interest payments to bondholders [4].

2 Research Technique

Because the subject of the research is students, particularly in Bina Nusantara, and the majority of these students are still engaged in online learning, we believe that an online survey is the most appropriate tool to utilize. The questionnaire will be distributed to multiple students at Bina Nusantara in order to ensure that the questionnaire results are relevant to the subject of the research. The questionnaire will be delivered via social media and a Bina Nusantara line group. The questionnaire was prepared using a Google form, which requires students to have a Google account in order to complete it.

Flowchart of research technique (Fig. 1).

3 Research Method

The purpose of this study is to ascertain the influence of digital financial literacy on student interest in capital market investments. The research's objective is exploratory in nature. The researchers conducted quantitative research using a non-experimental design. According to Creswell [17], the quantitative method is a type of study that collects data in the form of numerical data, which is then processed statistically and used as raw material for analysis. The research approach is to collect data on past and current correlations between beliefs, views, behavioral traits, and variables,

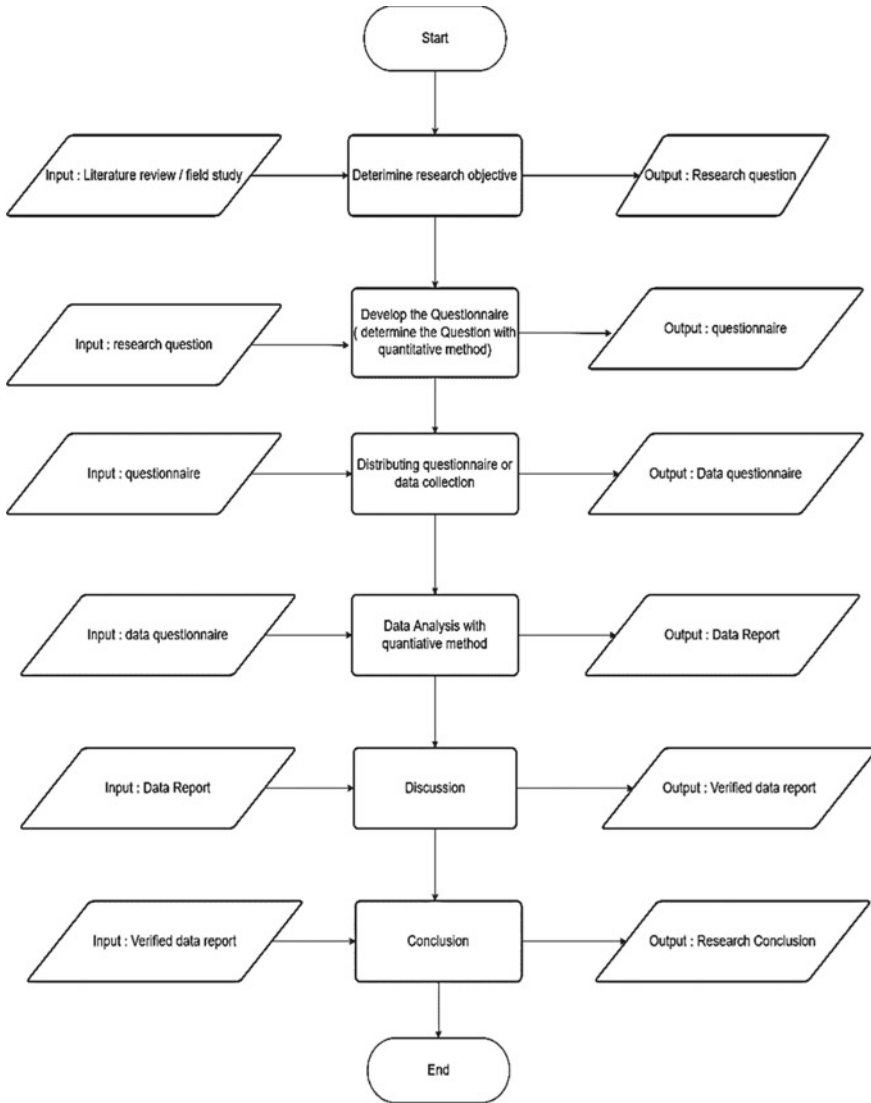


Fig. 1 Research technique flowchart

as well as to test various hypotheses about sociological and psychological variables using a sample of specified population groups. The researchers utilized a non-probability sampling technique combined with judgment sampling, which entails selecting samples from populations that meet specific criteria [18].

Students were used as the population, with a sample drawn from the BINUS class of 2024 majoring in information systems. The researchers will analyze a total of 96 samples. This is consistent with the standard rule of thumb for the number of

research-worthy samples, which is between 20 and 500 [19]. This idea is further supported by Frankel and Wallen's statement in Amiyani [20], who believe that the minimum sample size for descriptive research is 100, for correlational research is 50, for causal-comparative research is 30/group, and for experimental research is 30/15 per group. Additionally, the researchers employed the Slovin approach to establish the population's minimal sample size. Ryan [21] states that the Slovin approach follows the following formula:

$$n = \frac{N}{1 + N \cdot e^2}$$

where

n is the minimum number of samples

N is the number of populations

e is the sampling error margin [21].

Based on the sample Slovin formula above, it can be calculated if the population of the BINUS class 2024 information systems department is 373 students. By sampling, the margin of error is set at 10%. The calculation will be as follows:

$$n = 373 / (1 + (373 \times 0.1^2))$$

$$n = 373 / (1 + (373 \times 0.01))$$

$$n = 373 / (1 + 3.73)$$

$$n = 373 / 4.73$$

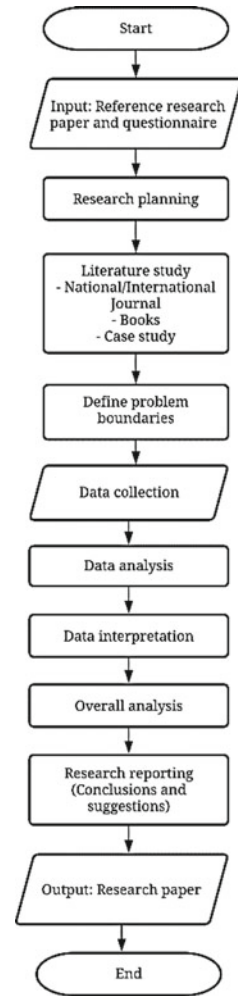
$$n = 78.858 / 79.$$

According to the preceding estimate, the minimal number of samples that can be taken with a 10% margin of error is 79. As a result, 96 samples is a reasonable sample size for this study.

The data collecting approach will be a questionnaire survey, and survey results are frequently generalized. Questionnaires will be issued to BINUS information systems majors in the class of 2024. We utilize this strategy because we believe it is more effective to study financial behavior using research methods that collect data from the past to the present. The researchers will first evaluate and process the questionnaire's data in order to determine its reliability and validity. If there is any invalid data, a word rearrangement will occur. If the questionnaire respondents reach the aim, there could be as many as 96 samples. As a result, data will be examined and gathered for subsequent analysis.

Research flow (Flowchart) (Fig. 2).

Fig. 2 Research method flowchart



4 Research Result

Profile questions

1. Question 1 (Fig. 3).
2. Question 2 (Fig. 4).
3. Question 3 (Fig. 5).
4. Question 4 (Fig. 6).
5. Question 5 (Fig. 7).
6. Question 6 (Fig. 8).
7. Question 7 (Fig. 9).
8. Question 8 (Fig. 10).

Age
96 responses

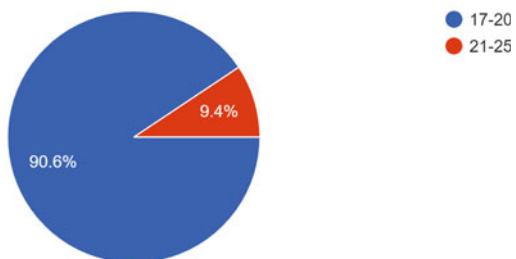


Fig. 3 Age responses

Are you Binusian?
96 responses

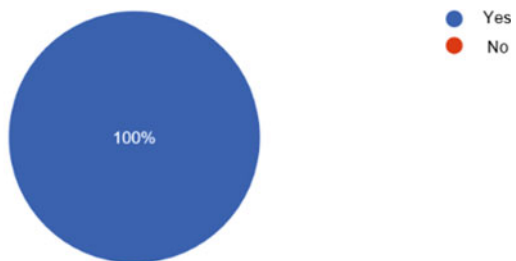


Fig. 4 Binusian responses

Your major
96 responses



Fig. 5 Major responses

Are you interested in investing in the capital market?

96 responses

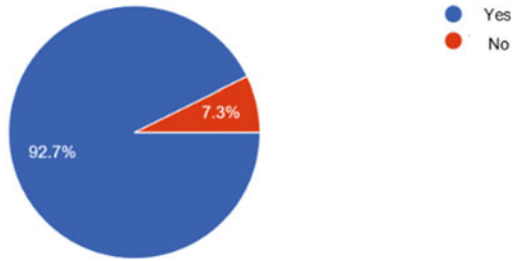


Fig. 6 Question 4 responses

I use stocks as an instrument to invest in the capital market

96 responses

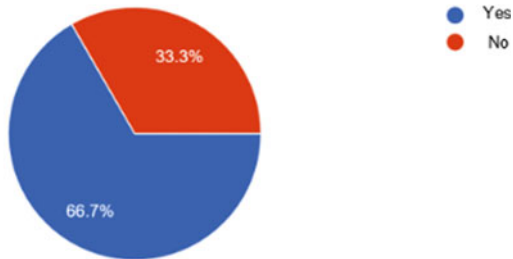


Fig. 7 Question 5 responses

I use mutual funds as an instrument to invest in the capital market

96 responses

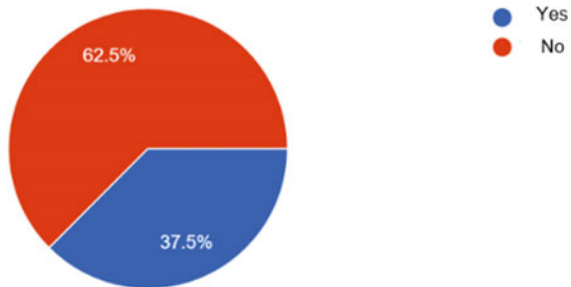


Fig. 8 Question 6 responses

I use debt securities/bonds as an instrument to invest in the capital market

96 responses

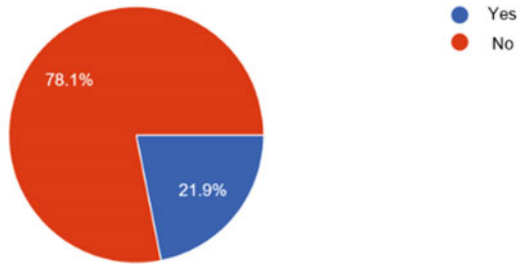


Fig. 9 Question 7 responses

I prefer to invest in financial assets (bonds, mutual funds, stocks) than real assets (houses, land, gold)

96 responses

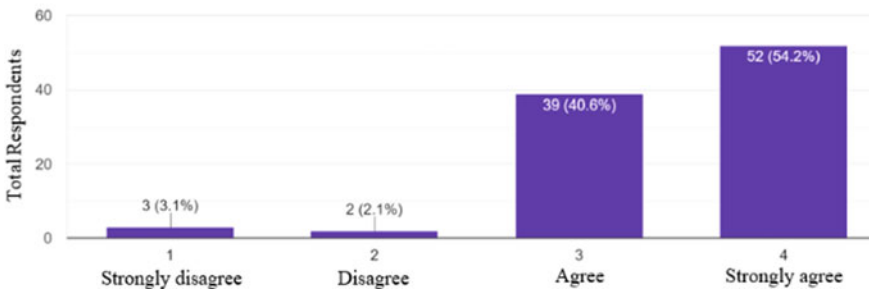


Fig. 10 Comparison of interest between financial assets and real assets responses

- 9. Question 9 (Fig. 11).
- 10. Question 10 (Fig. 12).

Notes: range 0–100 means the number of the person that chooses the answer of the options.

- 11. Question 11 (Fig. 13).
- 12. Question 12 (Fig. 14).
- 13. Question 13 (Fig. 15).
- 14. Question 14 (Fig. 16).
- 15. Question 15 (Fig. 17).
- 16. Question 16 (Fig. 18).

According to the results of a survey administered to students enrolled in the information systems class B24 at Bina Nusantara University, the majority of students are interested in capital market investment. This is demonstrated by the fact that up

How interested are you in learning more about investing in the capital market?
96 responses

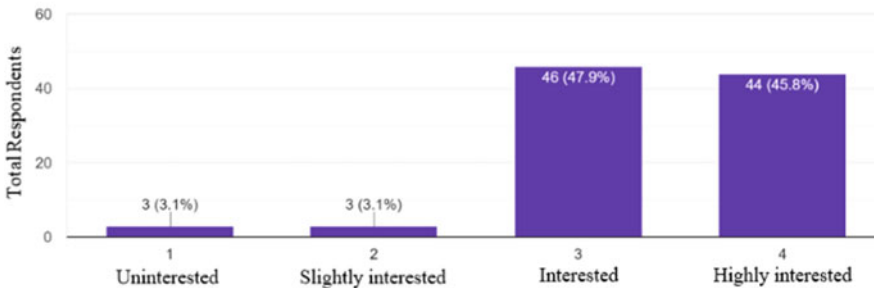


Fig. 11 Question 9 responses

What learning media do you use to learn how to invest in the capital market?
96 responses

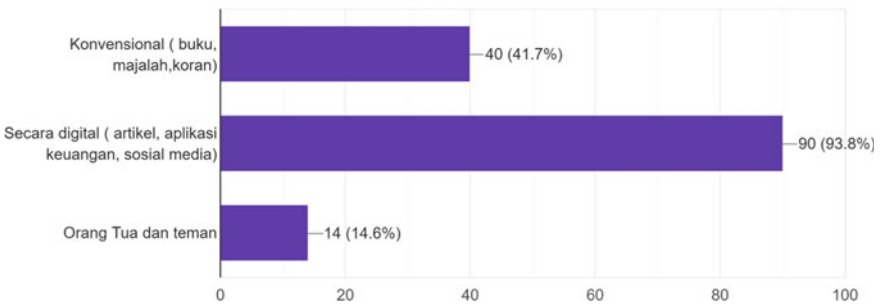


Fig. 12 Learning media use to invest in the capital market responses

Information about digital finance makes me INTERESTED to invest in the capital market
96 responses

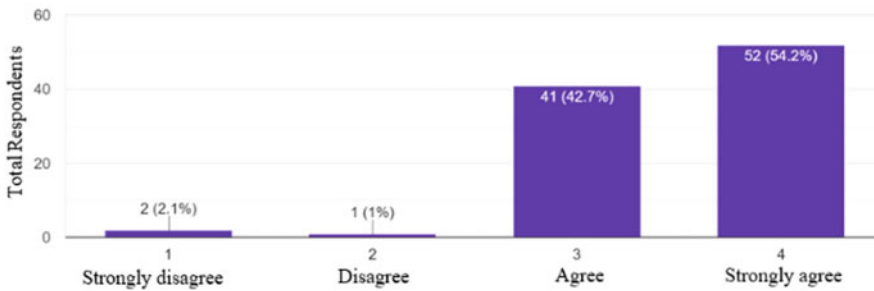


Fig. 13 Question 11 responses

Ease of accessing financial information is the reason why my financial literacy continues to increase and I am interested in investing in the capital market

96 responses

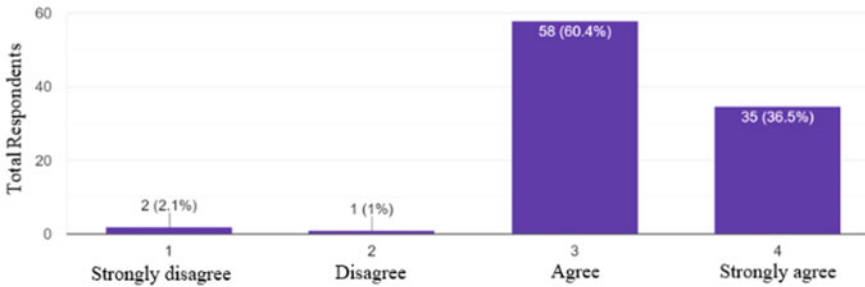


Fig. 14 Question 12 responses

The trend of investing in social media makes me interested in investing in the capital market

96 responses

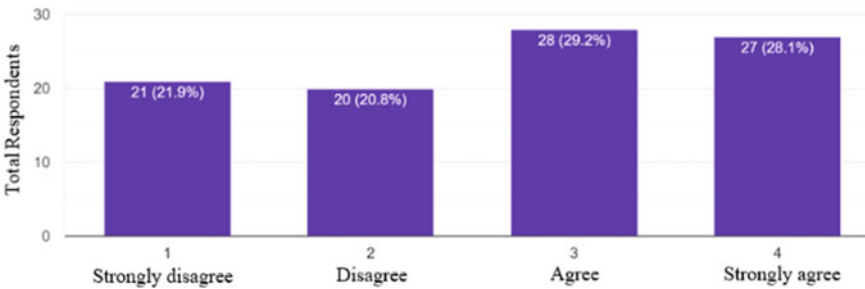


Fig. 15 Question 13 responses

How often do you read topics about finance digitally?

96 responses

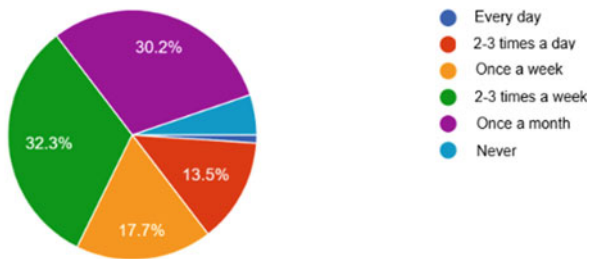


Fig. 16 Frequency of people reading financial topics digitally responses

conventional financial information (books, magazines, newspapers) does not make me interested in investing in the capital market

96 responses

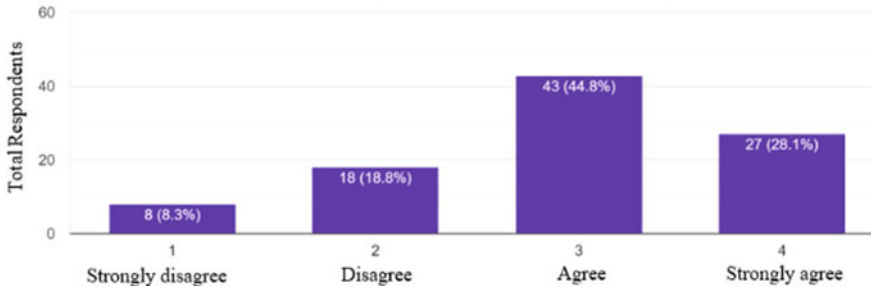


Fig. 17 Question 15 responses

Information about finance with the media (books, magazines, newspapers) is boring for me to study and reduces my interest in investing in the capital market

96 responses

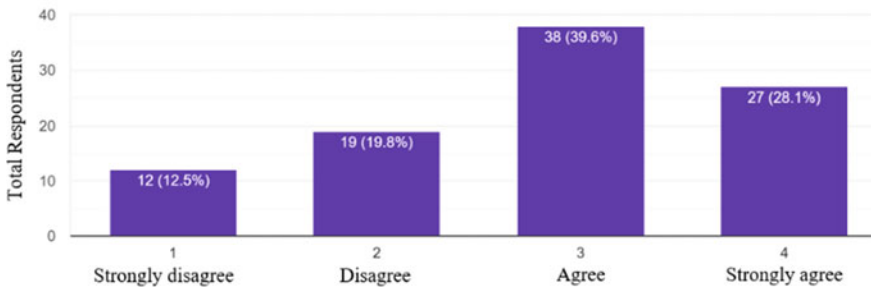


Fig. 18 Question 16 responses

to 92.7% of respondents responded affirmatively to the fourth question. Then, 66.7% of students invest in stocks, 62.5% in mutual funds, and 78.1% in bonds/debts. This demonstrates that students prefer debt securities/bonds above stocks, while mutual funds are the least desired. The majority of students favor financial assets (bonds, mutual funds, and stocks) above tangible goods (houses, land, and gold). This is demonstrated by respondents who responded “strongly agree” (with a value of 5) in a proportion of up to 54.2% and respondents who responded “very agree” (with a value of 4) in a proportion of up to 40.6%. Meanwhile, 3.1% of respondents responded “strongly disagree” (with a value of 1), while 2.1% responded “disagreed” (with a score of 2). Then, a sizable proportion of students expresses an intense interest in participating in the stock market. This is demonstrated by 47.9% of respondents who indicated they were intrigued (with a score of 3) and 45.8% who indicated they were extremely interested (with a score of 4). Meanwhile, respondents who

responded indifferent (with a score of 1) and slightly interested (with a score of 2) were both only 3.1%. The majority of students' learning media for investment, specifically in the capital market, is digital (articles, applications, finance, social media). This is demonstrated by the fact that 93.8% of respondents selected digital media. Meanwhile, 41.7% responded via traditional media and 14.6% via parents and friends. The majority of students believe that learning about digital finance piques their interest in investing, particularly in the capital market. This is demonstrated by the fact that 54.2% of respondents agreed "strongly" (with a value of 4) and 42.7% agreed (with a value of 3) on the statement "Information about digital finance piques my interest in investing in the capital market." Meanwhile, 2.1% said "strongly disagree" (with a score of 1), while 1% responded "disagree" (with a score of 2). Then, the majority of students believe that simple access to financial information contributes to their increased financial literacy and their interest in engaging in the capital market. This is demonstrated by the fact that respondents who agreed (with a score of 3) agreed with 60.4% and those who strongly agreed (with a score of 4) agreed with 36.5% on the statement "Ease of accessing financial information is one of the reasons my financial literacy continues to grow, and I am interested in investing in the capital market." Meanwhile, those who responded "strongly disagree" (with a score of 1) accounted for only 2.1% of respondents, while those who responded "disagree" (with a score of 2) accounted for only 1%.

According to several students, the practice of investing in social media has piqued their interest in capital market investing. This is demonstrated by the fact that 29.2% of respondents agreed (with a score of 3) and 28.1% strongly agreed (with a score of 4) with the statement "The trend of investing on social media has piqued my interest in the capital market." Meanwhile, 21.9% said "strongly disagree" (with a score of 1), while 20.8% responded "disagree" (with a score of 2). Each score displays a statistical percentage that is not significantly different. Then, every 2–3 weeks, the majority of pupils read about financial matters digitally. This is demonstrated by the fact that 32.3% of respondents vote every 2–3 weeks on the question "How frequently do you read financial matters digitally." Following that, up to 30.2% of respondents responded once a month. Then, 17.7% of respondents responded once a week, while 13.5% responded once every 2–3 days. And only a few respondents indicated that they never or daily. Additionally, the majority of students think that traditional financial material (books, periodicals, and newspapers) does not inspire them to invest, particularly in the capital market. This is seen by 44.8% of respondents agreeing (with a score of 3) and up to 28.1% strongly agreeing (with a score of 4) on the statement "Traditional financial material (books, periodicals, newspapers) does not pique my interest in investing in the capital market." Meanwhile, respondents who responded "disagree" (with a score of 2) accounted for only 18.8% of respondents, while those who responded "strongly disagree" (with a score of 1) accounted for just 8.3%. Additionally, the majority of students say that receiving financial knowledge via the media (books, periodicals, and newspapers) makes them tired with learning about investing in the capital market and diminishes their interest in it. This is demonstrated by the fact that up to 39.6% of respondents responded "agree" (with a score of 3) and up to 28.1% responded "strongly agree" to the statement in question 16.

Meanwhile, respondents who responded “disagree” (with a score of 2) accounted for just 19.8% of respondents, while those who responded “strongly disagree” (with a score of 1) accounted for just 12.5%. To summarize, many people are interested in and enthusiastic about investing in the financial market. This can be accomplished through the frequent practice of financial literacy.

5 Discussion

The previous survey conducted by OJK in 2019, public financial literacy increased 8.33% in the last three years. Indonesia’s financial literacy index reached 38.03% in 2019. Meanwhile, the financial literacy score rose to 29.7% in 2016. So, what about the results of the survey given to the current millennial generation? Do they just follow trends or do they really understand literacy related to financial investment?

From the results of the questionnaire survey that has been filled out by various millennial generation respondents today, we discuss whether the results obtained are in accordance with what is the problem formulation. The results of our discussion of each result obtained from the answers of respondents who filled out the questionnaire are, our respondents know that in the current era, digital financial literacy may be sufficient because of current technology that is already capable and also with education and various kinds knowledge of financial literacy which have been spread on the Internet, enabling students to understand what digital financial literacy is. Of the many respondents, 90% of the 96 respondents are 17 years old and 92% of them know a lot about various literacy from digital investment from various media on the Internet. With these results, it can be said that the average ratio of students who do not recognize digital investments well is just under 8%.

Another thing is also proven by the results of a survey that we conducted, with the ease of the current millennial generation in accessing financial information, which is the reason why my literacy continues to increase and is interested in investing in the capital market. Meanwhile, respondents who answered strongly disagree (with a score of 1) were only 2.1%, and those who answered disagreed (with a score of 2) were only 1%. Some students agreed that the trend of investing in social media made them interested in investing in the capital market. This is evidenced by as many as 29.2% of respondents answered agree (with a score of 3) and as many as 28.1% answered strongly agree (with a score of 4) on the statement “The trend of investing in social media makes me interested in investing in the capital market.”

6 Conclusion

Investing in the capital market is nothing new. However, lately investing in the capital market has become a trend on social media in Indonesia because of the large number of influencers who educate the public and the increasing number of applications

or information media about finance and investment that provide ways to invest in the capital market. Based on the findings of our research, it turns out that digital information about finance and investment makes students improve their financial literacy and makes them interested in investing in the capital market because of easy access. However, trends in social media also encourage them to try to invest in the capital market. In addition, investment instruments in the form of stocks and mutual funds are the most preferred choice by the current millennial generation to invest.

Another conclusion that we can see is that the survey conducted by OJK in 2019 experienced a significant level compared to previous years. This matter related to financial literacy is also reinforced by a survey we conducted for the current millennial generation, that is 96.9%. Then, most students think that the ease of access to financial information is the reason for increasing their financial literacy. So, that is what makes them interested in investing in the capital market.

On average, 66.7% of students use stocks as an instrument to invest in the capital market, 62.5% of students use mutual funds as an instrument to invest in the capital market, and 78.1% of students use mutual funds bonds/debt as an instrument to invest in the capital market. This shows that debt securities/bonds are the most preferred investment instrument by the millennial generation, followed by stocks and mutual funds which are the least preferred.

We hope that in the future, there will be further research on this subject with sufficient time and resources to find out further whether this increase in financial literacy continues to increase steadily or just because of the pandemic or trends and research results can also be explained better.

Acknowledgements Our team would like to thank all respondents, lecturers, and everyone who contributed to this research project. We hope that this research can be useful for all parties.

References

1. Rachman MA, Salam AN (2018) The reinforcement of zakat management through financial technology systems. *Int J Zakat* 3(1):57–69
2. Yusuf M (2019) The effect of advances in technology and knowledge on millennial generation's interest in investing in the capital market. *J Manag Bus Dyn* 2(2):86–94. <https://doi.org/10.21009/jdmb.02.2.3>
3. Ghozali I (2011) Analisis Multivariater dengan Menggunakan Program SPSS. Badan Penerbit Universitas Diponegoro, Semarang
4. Hayes A (2022) Capital markets: what you should know. Investopedia. <https://www.investopedia.com/terms/c/capitalmarkets.asp>
5. Nisa A (2017) Pengaruh pemahaman investasi, modal minimal investasi dan motivasi terhadap minat mahasiswa berinvestasi di pasar modal (Studi pada Mahasiswa Sekolah Tinggi Kesuma Negara). *Jurnal Penelitian Teori & Terapan Akuntansi (PETA)* 2(2):22–35. <https://doi.org/10.51289/peta.v2i2.309>
6. OJK (2020, December 1) Survei Nasional Literasi dan Inklusi Keuangan 2019. <https://www.ojk.go.id/id/berita-dan-kegiatan/publikasi/Pages/Survei-Nasional-Literasi-dan-Inklusi-Kuangan-2019.aspx>. Retrieved on 13 Jan 2022

7. Leong K (2018) Fintech (financial technology): what is it and how to use technologies to create business value in fintech way? *Int J Innov Manage Technol* 74–78. <https://doi.org/10.18178/ijimt.2018.9.2.791>
8. KPMG (n.d.) (2016) The pulse of fintech Q4 2016: global analysis of investment in fintech. Retrieved 4 Feb 2022, from <https://assets.kpmg/content/dam/kpmg/xx/pdf/2017/02/pulse-of-fintech-q4-2016.pdf>
9. Razzaque A, Hamdan A (2020) Role of financial technology FinTech: a survey. *Adv Intell Syst Comput* 112–117. https://doi.org/10.1007/978-3-030-44289-7_11
10. Prasad H, Meghwal D, Dayama V (2018) Digital financial literacy: a study of households of Udaipur. *J Bus Manag* 5:23–32. <https://doi.org/10.3126/jbm.v5i0.27385>
11. Awais M, Laber MF, Rasheed N, Khursheed A (2016) Impact of financial literacy and investment experience on risk tolerance and investment decisions: empirical evidence from Pakistan. *Int J Econ Financ Iss* 6(1):73–79
12. Chen H, Volpe RP (1998) An analysis of personal financial literacy among college students. *Finan Serv Rev* 72(2):107–128
13. Morgan PJ, Huang B, Trinh LQ (2019) The need to promote digital financial literacy for the digital age. *The Digital Age*. https://www.researchgate.net/publication/343682203_The_Need_to_Promote_Digital_Financial_Literacy_for_the_Digital_Age
14. Hayes A (2021) What is a stock? (2022). Investopedia. <https://www.investopedia.com/terms/s/stock.asp>
15. Masruroh A (2014) Konsep Dasar Investasi Reksadana. *SALAM: Jurnal Sosial Dan Budaya Syar-i* 1(1). <https://doi.org/10.15408/sjsbs.v1i1.1526>
16. Napoletano E (2021) Fixed-income basics: what is a bond? *Forbes*. Retrieved from <https://www.forbes.com/advisor/investing/what-is-a-bond/>
17. Creswell JW (2012) *Educational research: planning, conducting and evaluating quantitative and qualitative research*, 4th edn. Pearson Education, Boston. <http://repository.unmas.ac.id/medias/journal/EBK-00121.pdf>
18. Jogyanto (2008) *Metodologi Penelitian Sistem Informasi*. Andi, Yogyakarta
19. Sandelowski M, Barroso J (2003) Writing the proposal for a qualitative research methodology project. *Qual Health Res* 13:781–820
20. Amiyani R (2016) TEKNIK Sampling. Rofi Amiyani, Oleh. Download Gratis. Docplayer.info. <https://docplayer.info/31778225-Teknik-sampling-oleh-rofi-amiyani.html>
21. Ryan TP (2013) Sample size determination and power. *Wiley Ser Probab Stat*. <https://doi.org/10.1002/9781118439241>

Security Enhancement in IoT with an Application of Water Monitoring



Jui Sanjay Berde and Shirish S. Kulkarni

Abstract In recent years, numerous innovations have been observed in IOT-based process management systems. There are smart management technologies and architecture available that enable the user to fetch real-time data from remote areas and display/log them accordingly. The controller, cloud, and database usually used in this architecture is of original equipment manufacturer (OEM) IOT system or a third-party service provider. Thus, it will expose your data to unauthorized entities. In this project, we will explain security measures that can be taken to create a secure IOT architecture that will ensure the safety of user data and give assurity to the user that the data is not accessed without their consent. As a proof of concept, we will develop an Internet-based water flow monitoring system that will avoid water theft and will enable users to access data and create statistical inferences.

Keywords Raspberry Pi 4 · Security level · Water management · Water flow monitoring · Internet of things

1 Introduction

Internet of Things (IoT) encompasses a wide range of technologies and applications that bring both commercial benefits and threats. It is redefining and transforming the security industry's dynamics. It enables data to be smoothly transported from physical devices to the Internet. The proliferation of intelligent gadgets will generate a data-rich network that will enable supply chains to assemble and interact in novel ways. Gartner projects that by 2030, the installed base of Internet of Things (IoT) devices would have grown from 1.6 to 3.2 billion worldwide [1–3]. “Because of a

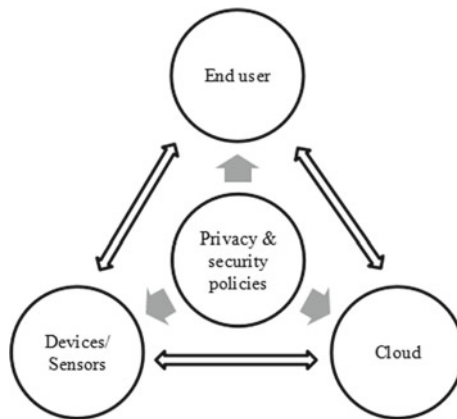
J. S. Berde (✉) · S. S. Kulkarni
Department of Instrumentation Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, India
e-mail: juioct18@gmail.com

S. S. Kulkarni
e-mail: shirish.kulkarni@rait.ac.in

lack of rules and standards, high data privacy, and security issues, all of the system's advantages are eclipsed".

We've demonstrated the secured IoT model with the features mentioned below

1. IoT device: Capable of receiving analog signal, discreet signal, digital signal, device-level protocols, having a unique identity, and storing historic data. Encrypting and decrypting stored data and received data using public and private key philosophy, traceability, secure communication, and device protection.
2. Cloud: As a supervisory controller and data acquisition unit, Data visualization, Database, Multiple level access.



IoT's generic model with privacy & securities

2 Literature Review

Security has always been an important paradigm in the very sector, due to the rapid increase of technology in the past years and still the rise is exponentially growing and due to this only the information security has always been a priority. Security, like everything else, has a history that dates back to the 1960s, when businesses first began using passwords to safeguard their computers. Because the Internet idea was not formed in the 1960s, security was primarily focused on a physical level, such as ensuring that no unknown individual had access to the computer [4].

The phrase "Internet of Things" was invented for the first time by a Briton, Kevin Ashton, three decades ago. Ashton, a British technologist who worked at the Procter & Gamble Company and was working on manufacturing process optimization, discovered during a presentation that if people collect data, it will take a long time. As a result, the usage of RFID will accelerate the process of data transfer directly between devices, and data will be gathered, processed, stored, and communicated without the need for human intervention [5-8]. He dubbed it as the "Internet

of Things” from then on, and by the twenty-first century, IoT had spread like wildfire and evolved significantly. When it comes to data transport, collecting, and processing through the Internet, it opens the door to undesired users who have the capacity to tamper with that data and abuse sensitive data of a certain corporation or organization [9–11].

According to Gartner Hype, the IoT industry will reach 5.8 billion endpoints in 2020, a 21% growth over 2019. Everything that can be linked will be linked, resulting in a fully digital system in which all gadgets communicate with people and each other [12]. Even with all of this progress in the IoT area, security remains a blind spot. Despite the numerous efforts made to safeguard IoT privacy and laws, as well as improve protective mechanisms for greater inter-connectivity, data security and privacy issues will never be completely eliminated. Cyber criminals will utilize increasingly advanced technology to exploit flaws in connected devices, giving them access to the user’s personal information. As a result, consumers and businesses are increasingly concerned about IoT security, which they regard as the most significant hurdle to mainstream IoT adoption.

To overcome such types of problems different researchers have proposed journal papers. Shipley [13] and Jing et al. [14] provide a list of security needs that should be evaluated phase-wise to avoid such events. Following are the prerequisites for security enhancement.

- Secured routing algorithms—When communicating via a wireless channel, the routing protocol must assure the validity of the routing information and eavesdropping must be avoided. Routing protocols should be protected to avoid Dos, Wormhole, black hole, and selective forwarding attacks.
- Secure operating system—It should have limited access permissions. Operating system should be developed with components, packages, and libraries needed to run an IoT device. Update must be given throughout the duration of the deployed device’s life. Non-used ports, protocols, and services must be deactivated. Separate access permissions to files and directories must be granted to users and administrators.
- Encryption—It should be cost-effective, strongest and most updated an IoT network and to be chosen in accordance with the sensitivity of the data to be safeguarded. A device’s private key should never be shared. The encryption keys should be replaceable remotely.
- Network connection—The number of interfaces through which device connects to the external network should be kept to a minimum. Gadget must only be accessible to a few ports, interfaces, and services. To safeguard connections, secure protocols shall be used.
- Secured event logging—Logs shall be reviewed on a regular basis to uncover any flaws, and quick action must be done. The log files must be placed on distinct file system partitions. The log file’s access privileges are to be limited. Logs must not contain any critical credentials, such as passwords [15–19].

3 Security Issues in IoT

- **Identification:** Identification of every device, whether original or malicious. A manufacturer's reference is necessary to be accessible.
- **Authentication:** Because of the large number of devices, one of the most serious difficulties in IoT is authentication. Authenticating each and every gadget is a difficult task. Many security procedures have been proposed as a result of the properties of quick processing and energy efficiency, which are based on private key cryptography primitives.
- **Data Management** The identification and addressing of billions of devices is a major challenge in the Internet of Things. According to predictions, more than 50 billion smart and network devices will exist by 2020. Even with IPv6, controlling the devices and addresses will be difficult. There are various methods for identifying things in the Internet of Things. Few examples include bar code recognition, vision-based object recognition, and others.

4 Software Languages

4.1 Operating System

Raspbian: It is an operating system based on Debian. This operating system is used on the transmitter side. Will be in charge of collecting data and uploading it to the cloud.

Server: Apache, CPU: 19.2 Ghz (8 cores), RAM: 8 Gb, SSD: 160 Gb, Bandwidth: 8000 Gb.

4.2 Programming Language at Transmitter Side

Python: It will act as a medium to interpret the data received from ADC and IO ports. Store it and post it to the cloud. For this, we have used the following modules:

- Import board and import busio: To access IO ports of raspberry pi
- Import requests: To make HTTP requests to cloud
- Import Adafruit_ads1 × 15.ads1115: Acquire analog data it interfaces with raspi using I2C protocol
- Import AnalogIn: Acquire analog voltage from pin in terms of percentage
- Pymodbus: To make transmitter compatible with Industrial protocol.
- Csv: To store data in csv file.

4.3 Database at Cloud Side

Securing data on the cloud (using AWS platform): To upload and safeguard data created throughout the process, we've used the Amazon S3 (simple storage service) cloud platform. We've made sure that the Amazon S3 bucket follows the proper restrictions and is not accessible to the general public. Similarly, we've established Amazon S3 bucket access control lists (ACLs) that grant chosen or authenticated users read, write, or full access. The following aspects were taken into consideration while selecting the Amazon cloud service to safely upload the data:

- Discrepancy in who has read and write access to our data
- Multi-factor authentication before deleting an object or data from the server.

It is the additional layer of security used to delete or update the data on ASW. It prevents accidental deletions from the bucket by requiring the following two forms of authentication:

1. Security credentials
2. Concatenation of a valid serial number, a space and the six-digit code displayed on an approved authentication device.

5 Hardware Components

5.1 Raspberry Pi 4B

The Raspberry Pi Foundation designed an ARM-based credit card-sized single board computer (SBC) as in Fig. 1. It consists of Wi-Fi and Bluetooth module. With the help of these, we can send data to cloud-based storage through Internet.

Features

1. Vast peripheral Support
2. Multiple Sensors
3. Supports all types of Codes
4. Faster Processor
5. Can be used as a portable computer
6. Maximum power consumption: 6.0 Watts.

5.2 LTE Module

Long-term evolution (LTE) is a relatively new cellular network paradigm established by the 3GPP as part of the fourth generation (4G) standardization process. A 3GPP has created standards for LTE networks that deliver high peak data speeds and effective

Raspberry Pi | Model 4 B

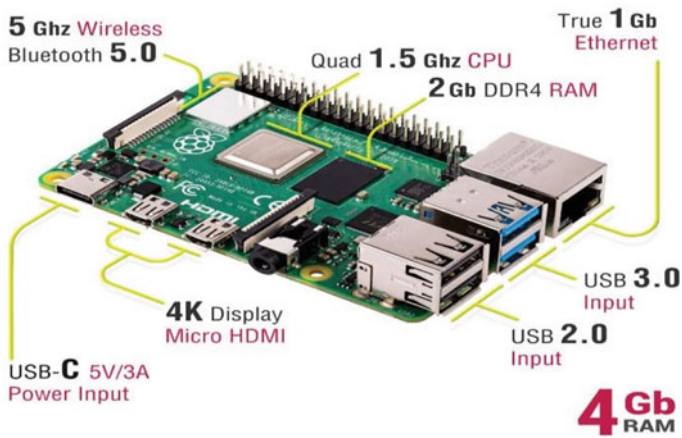


Fig. 1 Raspberry Pi model 4B

radio resource management starting with Release 8. High-speed cellular modules allow users to connect to 4G/LTE networks. They also integrate multiple-input multiple-output (MIMO) technology, which is required in current wireless communication systems like LTE to fulfill data throughput and connection dependability standards as in Fig. 2.



Fig. 2 Long-term evolution module

Fig. 3 4–20 mA signal generator



5.3 4–20 mA Signal Generator

To simulate analog signals. It takes 12 v or 24 V as input. Users can adjust the current from 4 to 20 mA. Adjust the actual fineness of up to 0.01 mA, and operate freely. Output load impedance range between 0 and 350 Ω (DC 24 V 0) Total Power: < 1 W as in Fig. 3.

5.4 ADS1115 4-Channel 4–20 mA 16-Bit Current Receiver

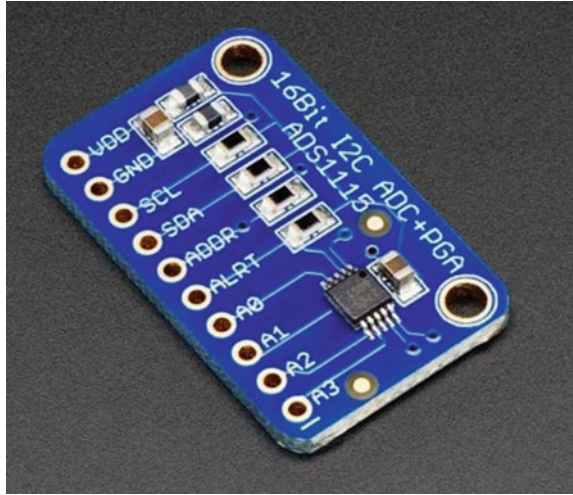
It offers 16-bit precision at 860 samples per second through I2C. The chip is set with four single-ended or two differential input channels. It also includes a programmable gain amplifier with a gain range of up to $\times 16$ to help boost smaller single/differential signals to the full range. It can operate from 2 to 5 V power/logic, and measures a wide variety of signals shown in Fig. 4.

1. Resolution: 16 Bit
2. Programmable Sample Rate: 8–860 Samples/Second
3. Low Current Consumption: Continuous Mode: Only 150 μ A Single-Shot Mode: Auto Shut-Down
4. Internal Low Drift Voltage Reference
5. Internal Oscillator, Internal PGA: up to $\times 16$
6. I2C Interface: 4-Pin-Selectable Addresses
7. Four Single-Ended or two Differential Inputs.

6 Design and Implementation

In this proposed system for security of IoT with the help of water management as an application to explain security and privacy needs, we have divided it into two parts,

Fig. 4 ADS1115
4—channel 4–20 mA 16 bit
current receiver



i.e., transmitter side code architecture and server-side code architecture. In transmitter side code architecture, I have explained the Raspberry Pi 4 overall internal connection and explained the terms in detail. In the server-side code part, it will acquire data from the transmitter and encode it and understand it with the programming part is explained in detail. Raspberry Pi 2 wiring diagram is shown in Fig. 5.

Raspberry Pi empowered with Raspbian is used for data acquisition, processing, and communication shown in Fig. 5. The controller collects all the data from the

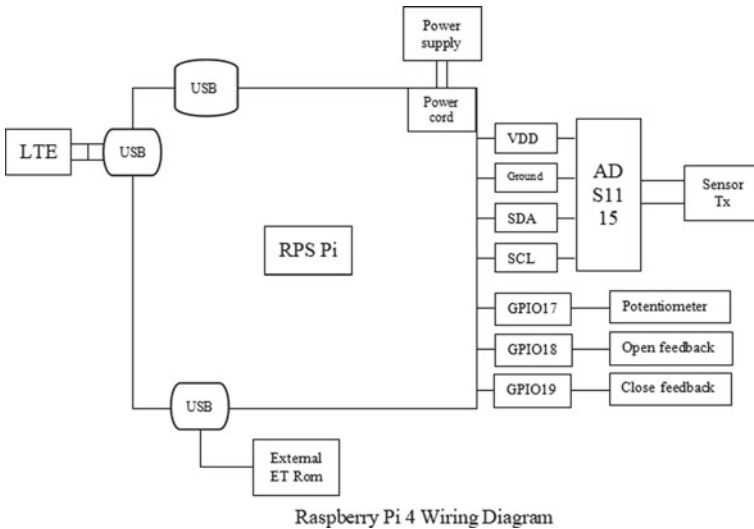


Fig. 5 Raspberry Pi 4 wiring diagram

respective instruments process it, encrypt it, and post it to the server. As you can see in the above diagram, there two USB ports one is connected to Long-Term Evolution (LTE) and another is connected to SD Card. There is one more power input port 5 V/3A, where we can connect a power supply. The Raspberry Pi 4 is connected to ADS 1115 which is connected to the sensor of the transmitter.

7 Security Enhancement in IoT

7.1 *Public and Private Key*

Secure Shell is a network protocol that allows users to remotely administer their servers, with communication between the two machines fully encrypted and shielded from eavesdropping. SSH provides a safe alternative to an unprotected network connection to the earlier technique of remotely controlling servers that were using network protocols such as telnet and rlogin (like the Internet).

SSH public key authentication and their functioning: A user connects to their server through puTTY/Terminal/Bash and initiates an SSH connection by signing in, followed by entering the password when required. SSH key authentication requires a private key and a public key. Unlike symmetric encryption, which uses the same key to encrypt and decode, SSH key authentication utilizes an asymmetric encryption technique, which uses distinct keys for encryption and decryption. A user (or an authorized person) must produce a private key as well as a public key. Secret key must be kept safe, confidential, and should not be shared with anyone else.

Secured login through Putty with the use of public and private key and VNC for UI is shown in Fig. 6.

7.2 *Securing PORT*

TCP ports make use of the transmission control protocol. It allows hosts to connect and exchange data. An IP address is a logical address that is used to identify a network device. For identification, each device connected to the internet is allocated a unique IP address. This identifying information allows devices to connect with one another via the internet. Port numbers are part of the addressing information that helps identify information senders and receivers, as well as a specific application on the devices. Port numbers are made up of 16-bit numbers.

We've changed the port number from the default value of 22 to a different number; as a result, if someone tries to log into the system using the default assigned port number of 22, the connection will fail since the port number does not match the allocated value. It will aid in the security of IoT devices (Raspberry Pi) and data transfer.

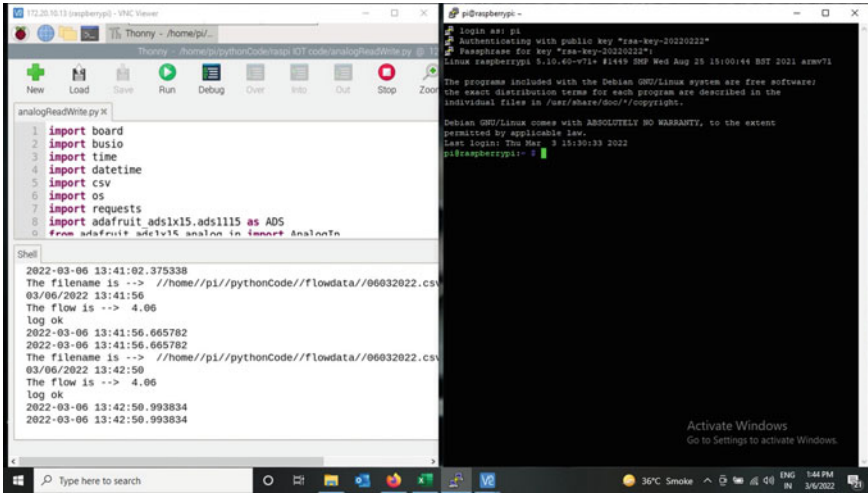


Fig. 6 Secured login through Putty with the use of public and private key and VNC for UI

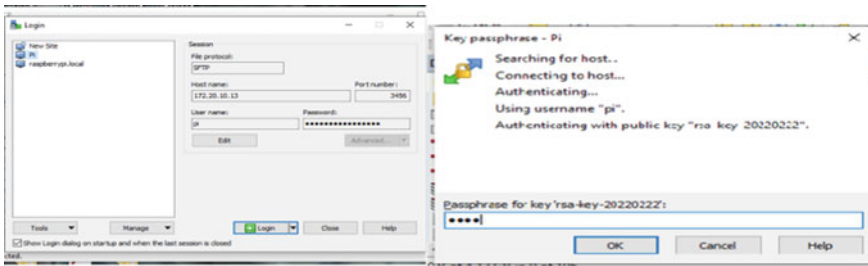


Fig. 7 Secured login (through defined port-3456)

The results on securing port are shown in Figs. 7, 8 and 9.

7.3 Fail2ban

There are several firewall options for Linux to offer packet filtering, most use the underlying iptables project. The command used to setup the firewall is: apt install fail2ban.

Fail2Ban is an intrusion preventive software framework that protects computer servers from brute force attack. Commands to verify the Fail2Ban Framework as in Fig. 10 with the help of following commands:

sudo su - & iptables -L

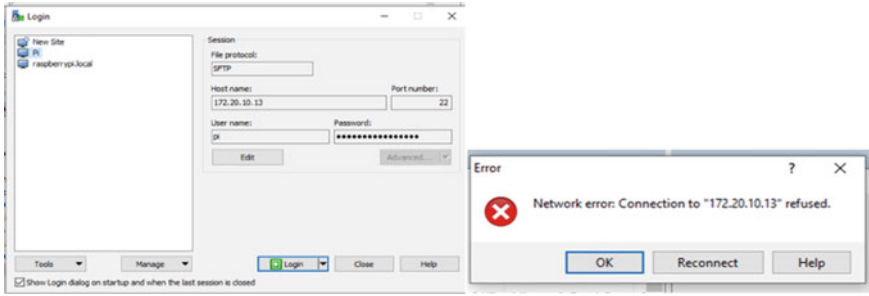


Fig. 8 Unsecured login (through default port-22)

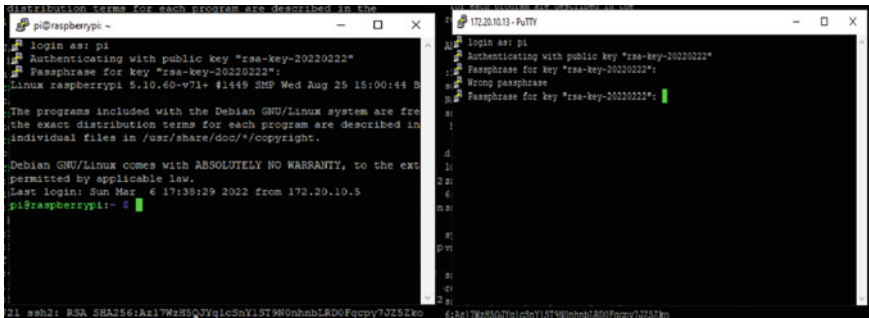


Fig. 9 Login through the correct key and Login through the wrong key

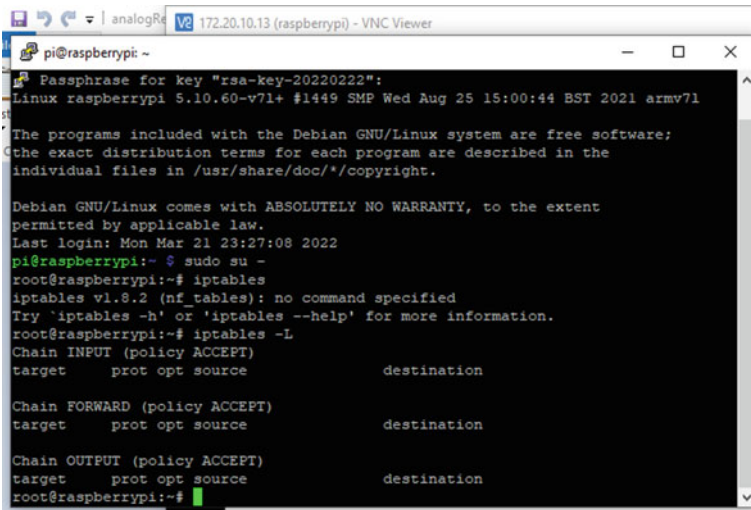


Fig. 10 Failban command execution

7.4 *Experimental Setup*

The above photograph describes the arrangement of an integrated flow totalizer in a remote location with the transfer of data using mobile network components constitutes as below:

1. Operating system for user interface
2. Raspberry pi (Details mentioned in Sect. 5, hardware components)
3. Integrated Flow Totalizer
4. Internet
5. Power Adaptor
6. Exhaust Fan (Exhaust fan provision is provided to maintain efficient cooling across the system)
7. TBs (for termination of all components internally).

Working: Once all of the connections have been validated and Raspberry Pi activated, log in to the system using the verified authentication process. After establishing a secure connection, the integrated flow transmitter receives sensor input and sends the value through the internet. The data results are consolidated in a CSV format file, which contains detailed flow values (flow per minute).

Electronic water flow meters are widely employed in all residential, commercial, and industries to manage, control, and regularize overall water requirements, and are typically positioned in isolated areas where it is simple to cut the power supply and steal water. The server will keep a record of transmitter data loss and power outages and generate a report appropriately. This may be viewed in combination with steady flow data as opposed to sudden changes in measurements, and discrepancies can be easily seen. It is a cost-effective monitoring solution for the field-installed remote electronic flow meter. The experiment setup and experiment results are shown in Figs. 11 and 12.

8 Conclusion

With regard to a water management system, we illustrated how supervisory data collecting and control architecture may be done safely in an IoT framework. This paradigm may also be applied to other contemporary processes that are low in criticality but high in relevance. It can be observed that by utilizing software languages, we can create our own IoT analytic, which has significantly aided us in safeguarding our data over other existing processes. Technology moves quickly, and data security has always been a concern, with the support of several research articles and the development of a secure IoT framework, we have attempted to put this idea into practice; which in turn will help not only to secure the data but collectively it'll help to analyze and improve the system in totality.

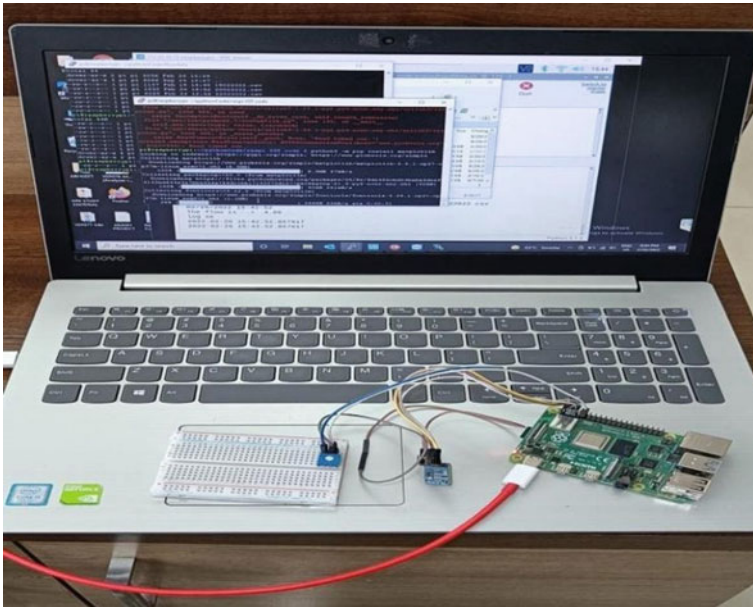


Fig. 11 Experimental setup of the project

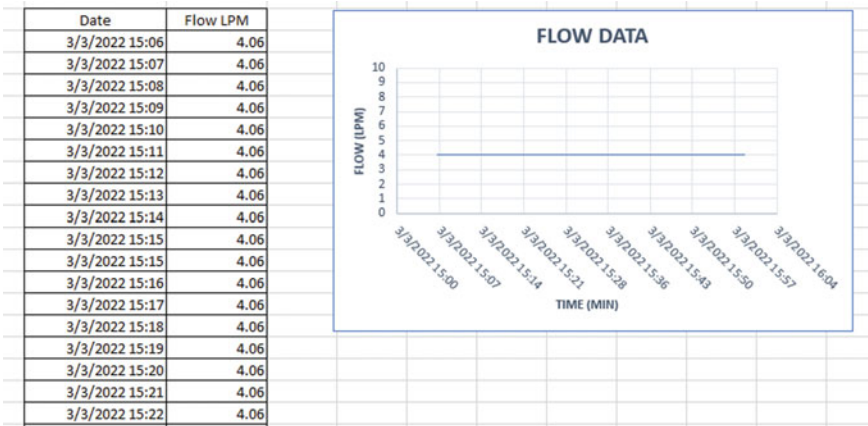


Fig. 12 Experimental result in tabular and graphical form

9 Future Scope

Replace the present hardwired control system with a cloud-based full-fledged SCADA that is just as safe. As a result, it will provide developers greater freedom by removing monopolies, allowing them to achieve a high level of flexibility in terms

of instrument and control systems, and integrate Big Data, Artificial Intelligence, and Machine Learning in a safe environment. In essence, IoT devices are integrated with these features to provide optimal automation. As a result, the Internet of Things (IoT) has broadened its range of applications in a variety of industries, and it is still being used by many multinational corporations for data management.

References

1. Fotiou N, Kotsonis T, Marias GF, Polyzos GC (2016) Access control for internet of things. In: 2016 international workshop on secure internet of things (SIoT)
2. Rehman A, Khan IU, Rehman SU, Moiz M (2016) Security and privacy issues in IoT. *Int J Commun Netw Inform Secur (IJCNIS)*
3. Pooja Yadav E, Ankur Mittal E, Yadav H (2018) IoT: challenges and issues in Indian perspective. In: 2018 3rd international conference on internet of things: smart innovation and usages (IoT-SIU), 23–24 Feb 2018
4. Meneghello F, Calore M, Zucchetto D, Polese M, Zanella A (2019) IoT: internet of threats? A survey of practical security vulnerabilities in real IoT devices. *IEEE Internet Things J* 6(5)
5. Atamli AW, Martin A (2014) Threat-based security analysis for the internet of things. In: 2014 international workshop on secure internet of things, 10–10 Sept 2014
6. Dikii D (2020) Authentication algorithm for internet of things networks based on MQTT protocol. *Serb J Electr Eng* 1(3):389–409
7. Niraja KS, Srinivasa Rao S (2021) A hybrid algorithm design for near real time detection cyber-attacks from compromised devices to enhance IoT security. *Mater Today Proc*
8. Garg H, Dave M (2019) Securing IoT devices and securely connecting the dots using REST API and middleware. In: 2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU), 18–19 Apr 2019
9. Fang Z, Fu H, Gu T, Qian Z, Jaeger T, Hu P, Mohapatra P (2021) A model checking-based security analysis framework for IoT systems. *High-Conf Comput* 1:100004
10. Iqbal W, Abbas H, Daneshmand M, Rauf B, Abbas Y (2020) An in-depth analysis of IoT security requirements, challenges and their countermeasures via software defined security. *IEEE Internet Things J* 7(10)
11. Shin S, Seto Y (2020) Development of IoT security exercise contents for cyber security exercise system. In: 2020 13th international conference on human system interaction (HSI), 6–8 June 2020
12. Mohamed AMA, Hamad YAM (2020) IoT security review and future directions for protection models. In: 2020 international conference on computing and information technology (ICCI-T-1441)
13. Shipley AJ (2013) Security in the internet of things, lessons from the past for the connected future. In: Security solutions, wind river, white paper
14. Jing Q, Vasilakos AV, Wan J, Lu J, Qiu D (2014) Security of internet of things: perspectives and challenges. *Wirel Network* 20(8):2481–2501
15. Jayapal C, Sultana P, Saroja MN, Senthil J (2019) Ubiquitous computing and computing security of IoT, pp1–28
16. Lam K-Y, Mitra S, Gondesens F, Yi X (2022) ANT-centric IoT security reference architecture—security-by-design for satellite-enabled smart cities. *IEEE Internet Things J* 9(8)

17. Jung J, Kim B, Cho J, Lee B (2022) A secure model based on ARM platform security architecture for IoT devices. *IEEE Internet Things J* 9(7)
18. Chen JIZ, Lai K-L (2020) Internet of Things (IoT) authentication and access control by hybrid deep learning method—a study. *J Soft Comput Paradig (JSCP)* 2(4):236–245
19. Kamel DK (2021) Wireless IoT with blockchain-enabled technology amidst attacks. *IRO J Sustain Wirel Syst* 2(3):133–137

Decision Support System for Weeding Out Drug Seeking Behavior from Emergency Clinics



Rugved V. Deolekar, Sunil Wankhade, and Mayur Wanve

Abstract The U.S. Food and Drug Administration has authorized many narcotic (i.e., opioid) painkillers for the treatment of mild to moderately severe acute or chronic pain in humans. In recent years, there has been a sharp increase in the number of individuals who fake symptoms in order to get prescriptions for these opiates for recreational use. The purpose of this study is to develop a multi-branched system that would aid physicians in identifying and eliminating such deceptive patients. The first branch attempts to determine whether patients are faking their symptoms, while the second branch utilizes the EHR to predict how much of the prescription opioid the patient requires. This will eventually prevent patients from becoming dependent on these medications and preserve these pharmaceuticals for their intended use.

Keywords Drug seeking behavior · Electronic health record · Electronic medical record · Opioid abuse · Decision support system · Deep learning

1 Introduction

Health care is one of the fastest-growing open industries in the world. This growth in the economic turnover of the medical industry has often resulted in patients having to shell out large sums of money. Technologies like machine learning and artificial intelligence can be used to aid doctors to make more accurate decisions

R. V. Deolekar (✉)

Vidylankar Institute of Technology, Mumbai, India

e-mail: iamrugved@gmail.com

S. Wankhade

Rajiv Gandhi Institute of Technology, Mumbai, India

e-mail: sunil.wankhade@mctrgit.ac.in

M. Wanve

University of Mumbai, Mumbai, India

through the development of expert decision support systems. The use of such technologies can tackle several problems like the low doctor to patient ratio, the high prices of basic health care, the average time spent by a general physician tending to a patient. Machine learning has many applications in health care ranging from the basic machine learning functionalities such as classification and prediction to more advanced functionalities like generating images from a sparse set of training data using Auto-encoders and Variational Auto-encoders to augment a training set. Furthermore, within the umbrella of machine learning, the concept of deep learning has great application in the medical industry. This concept of deep learning is the basis of a system that we will propose to notify physicians that the patient that they are tending to might have drug-seeking behavior. Deep learning has been used widely in processing radiological scans like CT scans, MR images to identify various anomalies, and monitor traumatic head injuries. With the advent of cloud computing and increase in digitization of records, hospitals and medical institutions have taken on the Electronic Health Records (EHR) or Electronic Medical Records (EMR) era. These health records can be used for a variety of predictive applications. One of these applications is proactive treatment, which deals with the handling of diseases before they have been contracted. This involves an involved problem which is the prediction of diseases.

2 Purpose

Across America, there has been a great increase in the war against drugs on their own land, i.e., the war against opioids or other strong drugs that are prescribed over the counter. These drugs are legally available on the advocacy of your need by a doctor; however, they can be used for recreational purposes by deceitful patients. Often patients may come into emergency clinics complaining about a set of symptoms that have been carefully crafted or selected by them so that the doctor prescribes a certain drug to them. We are proposing a system that can be used in medical clinics that prescribe strong pain relief pills to avoid enabling addicts. The system can be visualized as a module in any existing decision support system. The goal of the system, as stated earlier, is not to take the final decision but to aid in the form of a red flag, while the doctor is diagnosing the patient for such drugs. Every life is interconnected with many other lives and addiction ruins not only the lives of those who are addicted but also of the people that are associated with that life. It is impossible to expect every doctor to be able to identify such fraud with probability equaling unity, which means that there is a probability that a doctor may prescribe drugs to a patient and inadvertently enable him [1–5]. Our system tries to tackle this social evil, i.e., addiction head-on, by applying a multi-branched system that integrates deep learning and behavioral modeling using machine learning.

3 Literature Survey

Meyer et al. have highlighted the research studies based on the social and economic burden in their paper [5]. The latest economic and clinical evaluation associated with the prescription-based opioid abuse is also being reviewed. The opioid abuse resulted in an increased number of deaths from 5528 to 14,800 in a decade of 2002–2012. It was observed that the usage of medical services like outpatient visits to physicians, emergency clinics, and inpatient hospital stays were relatively more in opioid abusers as compared to non-opioid abusers. The years 1999–2007 have seen rapid increase in the growth rate of opioid overdose. The cost to society due to prescription opioid abuse was quite significant. Most individuals with a mean age of 40 were most likely the opioid abusers losing their life potential of several years.

The case study was conducted using the set of people with private insurance. Days of Supply (DoS) was used as a unit of measuring opioid use. It used to represent the number of days for which the therapy was given. These patients under study were categorized in 3 distinct groups as non-opioid group, acute opioid group, and chronic opioid group. The analysis showed that the number of emergency department (ED) visits was significantly more for the patients belonging to the chronic opioid group. Opioid addiction or dependency is significantly linked to ED usage. The Centers for Medicare and Medicaid Services (CMS) created the Overutilization Monitoring System (OMS) to track opioid consumption and potentially reduce opioid usage. OMS employ outlier measures to identify individuals with overutilization concerns, including opioid outliers. Pharmacies are also introducing procedures such as the “Good Faith Dispensing” policy, which compels pharmacists to call the prescribing physician and check that the diagnosis, billing code, prior drugs tried and failed, and projected period of therapy are accurate [6–10].

Che et al. [11] talk about classifying patients on opioid use. It also talks about the increasing amount of opioid misuse and abuse. The success of deep learning and availability of abundant data in the form of Electronic Health Records has resulted in proposing a deep and recurrent neural network model for the classification of opioid users. In this study, deep learning algorithms were used to a large-scale real-world EHR dataset to predict opioid usage groups. Deep learning models outperformed traditional classification methods and identified important feature indicators for opioid-dependent and long-term users. The research demonstrated how revolutionary deep learning models may be used to achieve cutting-edge performance in actual clinical investigations.

According to Compton et al. [12], prescription drug overdoses resulted in 70,237 deaths during 2017, and 47,600 (67.8%) of them involved opioids, according to Scholl L3. Between 2005 and 2014, OD-related hospitalizations and emergency room visits increased 64 and 99%, respectively. Hudgins et al. [13] highlighted that prescription opioid usage is widespread among teenagers and young adults, and it is frequently accompanied with other substance addiction, emphasizing the significance of drug and alcohol screening programmers in this demographic. 55.7% of those misusing opioids received it via friends or family, 25.4% from the healthcare system, and

18.9% from other sources. Cicero [14] did a systematic assessment of the literature, concentrating on peer-reviewed studies that used a qualitative approach to examining the development of an opioid use disorder from the initial exposure point. First of all, in the 1990s and early 2000s, medical circles placed a high premium on pain management, propelled in large part by a major “educational” effort launched by a pharmaceutical corporation to promote a new, extremely popular extended-release painkiller (see below). The JCAHO investigation stated that effective narcotic analgesics were accessible but were hardly utilized, and that physicians were disregarding pain care due to an illogical fear of addiction.

Oxycodone [14], an extended-release drug, was the second main factor contributing to the increase in prescription opioid abuse. Oxycodone is an opioid agonist with a high affinity for the opioid receptor, making it not only an efficient painkiller but also a powerful euphoriant. Due to the formulation’s slow release characteristics, the FDA enabled the sponsoring company to state in the package insert that abuse was anticipated to be minimal. Addicts quickly discovered, however, that they could avoid the slow release mechanism by crushing or dissolving the pills, so releasing considerable amounts of oxycodone in a form suitable for snorting or intravenous injection.

Using claims data and Electronic Health Record data from patients, Dong et al. [15] created various prediction models for assessing the probability of opioid poisoning in the future. They analyzed large-scale datasets to identify characteristics typically linked with opioid toxicity, as well as a patient’s medical records for prediction, diagnosis, test results, prescriptions, and associated clinical events. Among the different prediction models tested on the SPARCS and Health Facts datasets, such as Random Forests, Decision Trees, Logistic Regression, and deep learning, Random Forest performed the best on these datasets, with an AUC of 94.94% for SPARCS and 95.41% for Health Facts. It illustrates that an AI-based approach, when implemented at the clinical level, may achieve high-accuracy automated prediction and opens the door for AI-assisted prediction to support healthcare practitioners in combating the opioid problem.

In his study, Vunikili [16] classified patients with opioid prescriptions into eight age groups by subtracting the patient’s date of birth from the day the prescription was written. In addition, to identify individuals with side effects, he used ICD 9 codes to examine the diagnosis of every patient supplied with opioids for symptoms associated with overdose and/or dependency. Principal Component Analysis was used in order to obtain a more refined subset of characteristics [16]. The patients of interest were divided into two groups based on their proclivity to opioid abuse using two models: Logistic Regression with L2 regularization (baseline model) and Extreme Gradient Boosting (enhanced model). The baseline model has an F1 score of 76.64% (accuracy 77.16%), but the upgraded model has an F1 score of 94.45% (accuracy 94.45%) (accuracy 94.35%). Electronic Health Records (EHRs) facilitate the enhancement of patient care, the incorporation of performance measures into clinical practice, and the identification and recruitment of qualified patients and healthcare providers for clinical research [17]. EHRs may help determine, on a macroeconomic scale, whether novel medicines or advances in healthcare delivery result in improved outcomes or

cost savings. EHRs may be utilized to establish the feasibility of a study, to streamline patient recruitment, to expedite data collection, or to conduct completely EHR-based observational, embedded pragmatic, post-marketing randomized registry studies, or comparative effectiveness studies.

4 Design

4.1 Proposed System

A. Branch 1: Patient Symptom Faking

- We predict a set of plausible symptoms by processing Electronic Health Records using deep learning, generating normal space for likely symptoms, and performing novelty detection to estimate the possibility of patient faking symptoms [18].
- While extensive research is being conducted on Electronic Health Records modeled using deep learning for disease prediction, it is exceedingly difficult to understand the variation of health notes or facts recorded in EHRs and their mapping to disease prediction due to a current open research problem called “explainability of neural networks”. While predicting symptoms with a deep learning network (which is effectively a neural network with more than one hidden layer) is computationally expensive due to the scalability difficulty with neural networks, there is considerable interest and capability in disease prediction. As a result, we established a simple work around with a fact that each sickness has a prescribed set of symptoms in medical research.
- As shown in Fig. 1, we extrapolate predicted diseases to produce symptoms, which may be thought of as a pair of (key, value) values. Each disease has a value, which is comprised of a collection of symptoms. This set can be further reduced to provide a more concentrated set of symptoms through rule-based elimination of symptoms, such as those induced by age, gender, family history, or environmental factors.
- This set of symptoms can be understood as the normal space for patient symptoms (P2). The patient may complain about a set of symptoms which will be compiled as P1. The set difference may be thought of as the number of symptoms that the patient complains about (P1) that are not present in the expected set of symptoms created by the trained deep learning network using the EHR’s longitudinal data (P2). The set difference can then be transferred to a function analyzer that produces an analytical result based on the symptoms.

For example, if the symptoms are trivial and common (such as a common cold), they can be assigned a low value; yet, if a symptom is unusual and serious (such as sleeplessness), it will be allocated a higher value. Furthermore, a sum of all the

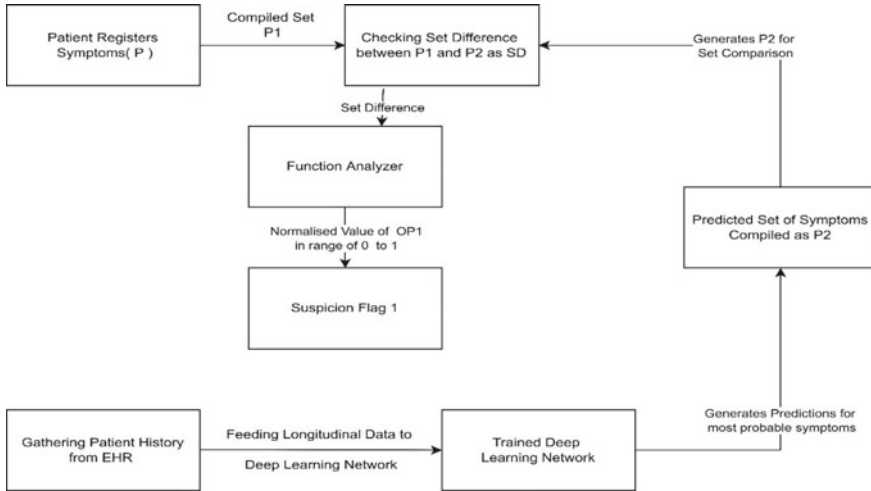


Fig. 1 State transition diagram for Branch 1

values corresponding to the symptoms will be returned by the function. This value is normalized to get the value of Op_1 in the range of 0–1.

B. Branch 2: Patients Suffering from Substance Abuse (In this Case, Opioid Abuse)

- Using the EHRs of patients as input, it is feasible to predict whether a patient is abusing substances.
- This strategy aims to identify features in the EHR that can be considered as the strongest predictors of substance abuse.
- This branch is largely concerned with the patient’s profile information, such as living arrangements, age, gender, education, employment status etc. to arrive at a suitable conclusion.
- As seen in Fig. 2, our Machine Learning Model takes the patient’s EHR as input and produces a likelihood that the patient needs the requested treatment.

C. Aggregation of Results of Branches: The Decision Function

Op_1 signifies the chance that the patient is lying, expressed as a number between 0 and 1. A greater value indicates a greater likelihood of lying.

Similarly, Op_2 shows the likelihood that the patient needs the specified treatment, given the profiling data from the EHR.

Our final decision is determined by the intermediate outputs, Op_1 and Op_2 . We can try to learn the relevance of Op_1 and Op_2 in creating the final choice by using a machine learning model to make our system more resilient.

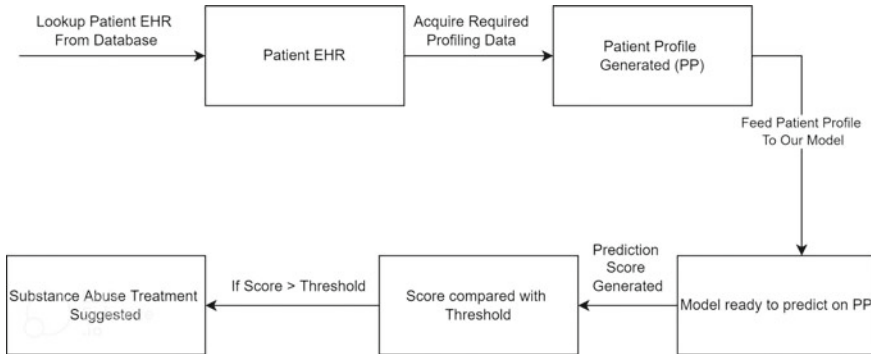


Fig. 2 State transition diagram for Branch 2

Our final output, the Alert Flag, may be parametrized as a linear correlation between Op_1 and Op_2 and can be represented as the following:

$$\text{Alert Flag} = W_0 + W_1 * OP_1 + W_2 * OP_2 \tag{1}$$

Essentially, during the initial stages of the development of this model, we may have to consider the choice made by the doctor utilizing this system to be the truth value. As a result, we would have Alert Flag as a doctor output along with OP_1 and OP_2 created by the system. This is now a supervised learning problem, using OP_1 and OP_2 as the input data and Alert Flag as the correct output data. The mapping function can thus be learnt by finding optimal values for the hyper-parameters (W_0 , W_1 and W_2).

• **Nonlinear Parameterization Approach**

- As shown above, the parameterization has been linearly modeled but this parameterization may also be represented using nonlinearly forms [19].

$$\text{Alert Flag Rep} = W_0 + W_1 * Op_1 + W_2 * Op_2 + W_1Op_1 * Op_2 \tag{2}$$

• **Modeling the Parameterization Approach**

- The parameterization approach opens doors to modeling this relationship.
- This is essentially a classification in which we will classify a patient as a drug-seeking patient or not. Hence, this is a Binary Class Problem we can think about modeling this approach using.
- Linear Regression with a threshold and output can be used for classification.
- Logistic Regression with sigmoid as nonlinearity with

$$\text{Alert Flag} = \text{IF}(\text{Sigmoid}(\text{Alert Flag Rep}) \geq 0.5) \tag{3}$$

4.2 Proposed Algorithm

The proposed design is shown in Fig. 3.

I/P: EHR of the patient.

O/P: Probability that the patient is a drug seeker.

Step 1: Take the patient’s EHR as input and feed it to multiple deep learning models, each of which extracts information from the EHR and outputs a likelihood that the patient is suffering from that disease. So, at the end of this process, we’ll have a list of diseases that the patient has had in the past or may have in the future based on the symptoms identified by the models. $D = D1 \dots Di$;

Step 2: Every disease can be thought of like a dictionary of symptoms that are common to that disease. $S_1 = \{s_1, s_2, \dots, s_n\}$; So, for ‘i’ diseases that got recognized,

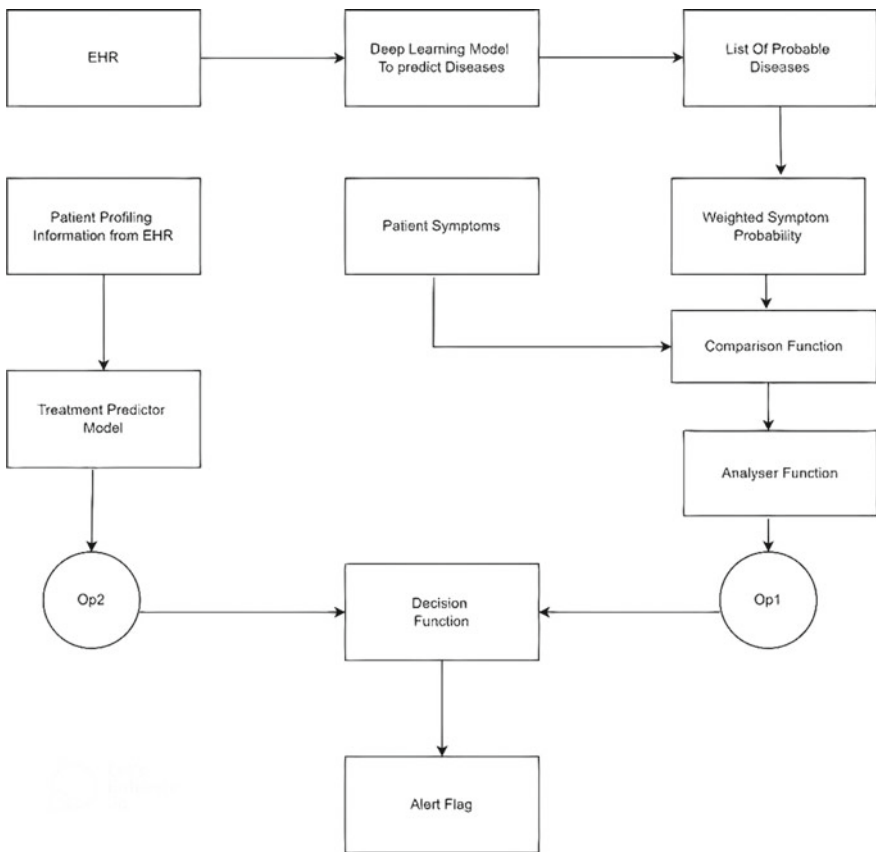


Fig. 3 Proposed decision support system

we have ‘ i ’ dictionaries available and a set S can be formed by taking a union of all these $S = S_1 \cup S_2 \cup \dots \cup S_n$;

Step 3: Now that we have S , there may be certain symptoms that the patient has that are present in S' (based on EHR). To narrow our search space, we build S' a collection of reduced symptoms’ space. $S' = \text{EHR Reduction Function}(S)$;

Step 4: So now we have a more limited range of symptoms that the patient may have had previously to cross-check whether the patient is faking his symptoms or not. Let P represents the patient’s list of symptoms.

$$OP_1 = \text{Comparison_function}(S', P) \quad (4)$$

OP_1 attempts to map the degree of similarity between S' and P in this case. If the degree of similarity exceeds a Suspension Threshold, a numerical output V_1 is produced. Suspension Threshold is a hyper-parameter that may be improved later in the development process.

$$V_1 = \text{Calculate_Relativity}(OP_1, \text{Suspension_Threshold}) \quad (5)$$

Step 5: We have a behavioral model that takes the EHR profiling information (P) as input and outputs the probability that the patient needs the requested treatment.

$$OP_2 = \text{Treatment_Predictor_model}(P)$$

Step 6: Based on the results of Ops 1 and 2, an Alert Flag is raised, assisting doctors in detecting and weeding out individuals who are faking symptoms.

5 Conclusion and Future Scope

We have proposed a theoretical pipeline that integrates the use of deep learning techniques to learn and extract information from EHRs and machine learning techniques to efficiently and accurately provide insights based on patient profiling information. The relatively new breakthrough in learning from longitudinal, time-sequential data enables the holistic interpretation of data and the inclusion of insights not only from the dimensions of the dataset but also from their continuity across time, allowing for the inclusion of the dimension of time. This idea of using some recurrent neural network architectures like Gated Recurrent Units and Long Short-Term Memory Units in the decision support system should be able to provide value to the medical professional by instantly delivering a birds-eye perspective analysis of the patient’s history. The implementation of this proposed system can be considered as a future scope.

References

1. Wu QX, Bell DA, McGinnity TM, Prasad G, Qi G, Huang X (2006) Improvement of decision accuracy using discretization of continuous attributes. Springer, Berlin. https://doi.org/10.1007/11881599_81
2. Veitch DP et al (2019) Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia* 15(1):106–152
3. Choi E et al (2016) Doctor AI: predicting clinical events via recurrent neural networks. In: Machine learning for healthcare conference. PMLR
4. Solares JRA et al (2020) Deep learning for electronic health records: a comparative review of multiple deep neural architectures. *J Biomed Inform* 101:103337
5. Meyer R, Patel Anisha M, Rattana SK, Quock TP and, Mody SH (2014) Prescription opioid abuse: a literature review of the clinical and economic burden in the United States. *Popul Health Manage.* <https://doi.org/10.1089/pop.2013.009>
6. Jing X (2012) A Bayesian network based intelligent plant classification system. Fourth Int Symp Inform Sci Eng 2012:263–265. <https://doi.org/10.1109/ISISE.2012.65>
7. Roveri M (2019) Learning discrete-time markov chains under concept drift. *IEEE Trans Neural Netw Learn Syst* 30(9):2570–2582. <https://doi.org/10.1109/TNNLS.2018.2886956>
8. Mbarki A, Naouai M (2020) A marked point process model for visual perceptual groups extraction. *IEEE Int Conf Vis Commun Image Proces (VCIP) 2020*:511–514. <https://doi.org/10.1109/VCIP49819.2020.9301776>
9. Van Houdt G, Mosquera C, Nápoles G (2020) A review on the long short-term memory model. *Artif Intell Rev* 53(8):5929–5955
10. Chung J, Gulcehre C, Hyun Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 deep learning and representation learning workshop, <https://arxiv.org/abs/1412.3555>
11. Che Z, Sauver JS, Liu H, Liu Y (2017) Deep learning solutions for classifying patients on opioid use. *AMIA Annu Symp Proc.* 16 Apr 2018
12. Compton WM, Han B, Blanco C, Crane E, Lee J, Jones CM (2017) Prescription opioid use, misuse, and use disorders in US adults: 2015 national survey on drug use and health. *Ann Intern Med* 167(5):293–301
13. Hudgins JD, Porter JJ, Monuteaux MC, Bourgeois FT (2019) Prescription opioid use and misuse among adolescents and young adults in the United States: a national survey study. *PLoS Med* 16(11):e1002922. <https://doi.org/10.1371/journal.pmed.1002922>
14. Cicero TJ, Ellis MS, The prescription opioid epidemic: a review of qualitative studies on the progression from initial use to abuse. *Dialog Clin Neurosci* 19(3):259–269. <https://doi.org/10.31887/DCNS.2017.19.3/tcicero>
15. Dong X, Rashidian S, Wang Y, Hajagos J, Zhao X, Rosenthal RN, Kong J, Saltz M, Saltz J, Wang F (2019) Machine learning based opioid overdose prediction using electronic health records. In: *AMIA annual symposium proceedings*, pp 389–398. 4 Mar 2020. PMID: 32308832; PMCID: PMC7153049
16. Vunikili R et al (2021) Predictive modelling of susceptibility to substance abuse, mortality and drug-drug interactions in opioid patients. *Front Artif Intel* 4:742723. 10 Dec 2021. <https://doi.org/10.3389/frai.2021.742723>
17. Cowie MR, Blomster JI, Curtis LH et al (2017) Electronic health records to facilitate clinical research. *Clin Res Cardiol* 106:1–9. <https://doi.org/10.1007/s00392-016-1025-6>
18. Hamdan YB (2020) Faultless decision making for false information in online: a systematic approach. *J Soft Comput Paradig (JSCP)* 2(4):226–235
19. Sungeetha A (2021) COVID-19 risk minimization decision making strategy using data-driven model. *J Inf Technol* 3(1):57–66

Analysis of Machine Learning Algorithm to Predict Symptoms of Diabetes—A Review



M. Sivaraman and J. Sumitha

Abstract Diabetes is the major disease increasing ratios for many peoples. This issue causes loss of sight, kidney failure, amputations, heart problems, lower and stroke limb amputation. Diabetes can be caused by various reasons such as an unhealthy lifestyle, lack of exercise, poor dietary habits, genetics, obesity, etc. Our body turns foodstuff into sugars or glucose. It is a set of disorders characterized by elevated blood glucose levels. For diabetes, the pancreas is not supposed to release insulin. Different ML technologies can be used to diagnose and forecast diabetic individuals. This paper discusses current enhancements in ML techniques that have significantly impacted the prediction of diabetes.

Keywords Diabetic prediction · Data mining · DBN · HRRFNN · Machine learning · SVM

1 Introduction

Machine Learning: Machine Learning techniques are work of excerpting enormous datasets to invent models and create relationships to determine the problems. It is defined as hauling out fascinating and nifty details against a populated dataset. Machine Learning techniques are used different fields such scientists, governments, airline passenger records, census data, and businesses. Data mining is intermittently studied as knowledge discovery in databases-(KDD). Data mining is also a knowledge discovery process. In recent days, data mining techniques have been used the healthcare field to discover useful information from the healthcare dataset.

M. Sivaraman (✉) · J. Sumitha
Department of Computer Science, Dr. SNS Rajalakshmi College of Arts and Science,
Coimbatore, Tamil Nadu, India
e-mail: sivaramanranjith@gmail.com

J. Sumitha
e-mail: sumivenkat2006@gmail.com

Diabetes: In 2022, almost 400 million people will have diabetes mellitus. These numbers are predicted values to increase to 700 million by 2045. Diabetes is a chronic metabolic disease defined by abnormal blood sugar levels caused by either ineffective use or inadequate production of insulin [1]. Diabetes mellitus is commonly referred to as diabetes. Sugar levels may increase in the food which has especially carbohydrate foods. Those foods contain carbohydrates, which give everyone a primary energy source, including diabetes. Blood sugar levels rise in diabetes when human pancreas does not make needed insulin or the human body cannot utilize it. Diabetes means that the bloodstream and urine contain high sugar levels [2, 3]. Diabetes is classified as:

Type-1, also known as Insulin Dependent Diabetes Mellitus, requires the patient to be injected with insulin because the human body cannot produce enough insulin.

Type-2, also known as Non-Insulin Dependent Diabetes Mellitus, arises when cells in the body cannot use insulin efficiently.

Type-3 diabetes, commonly known as Gestational Diabetes, is characterized by increased blood insulin levels in pregnant women who have not been diagnosed with diabetes early in their pregnancy.

Symptoms of Diabetes: Frequent urination, dehydration, hungry, malaise, drowsiness, weight loss, unclear vision, confusion, poor concentration, increased infections, and low level of healing [3].

Diet for Diabetes Person: Diabetes patients need to follow some habits in real life, such lose extra weight, stopping smoking, doing regular physical activity, drinking more water, eating healthy plant foods, and stopping taking sugar products. This paper is split into some main parts. Section 2 [4, 5] covers details about the flow diagram and measurement parameters. Section 3 shows the related work of this research work. Section 4 shows the details of data mining techniques. In Sect. 5, the comparative analysis research work. Finally, Sect. 6 concludes this research work.

2 Flow Diagram and Measurement Parameters

Figure 1 explains the diabetes prediction process. The data samples are collected first. The second process is preprocessing. Then implement some proposed algorithms. Finally, will get results and output.

The results are calculated using parameters like accuracy, precision, recall, etc. The parameters formulae have displayed in Fig. 2. The Confusion matrix the table layout and demonstrates the overall performance of the algorithms. It is split into four parts. TP is considered a true positive, TN is a true negative, FP means false positive, and FN means false negative.

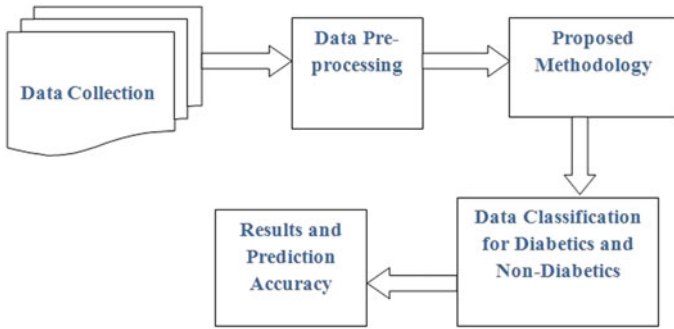


Fig. 1 Model flowchart for diabetes prediction

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Metrics	Formula
Recall	TP
	$\frac{TP}{TP + FN}$
Precision	TP
	$\frac{TP}{TP + FP}$
Accuracy	TP+TN
	$\frac{TP+TN}{TP+TN+FP+FN}$
F1_score	2 * precision * recall
	$\frac{2 * precision * recall}{precision + recall}$
Support	TN +FP , TP + FN

Fig. 2 Measurements calculated formulae

3 Background Study

Diabetes prediction systems based on various algorithms and styles have already been researched and built by research.

Roxana et al. [6] present an Ensemble Perceptron Algorithm (EPA) and Boosting Algorithm (BA). PA is a simple algorithm compared to BA, executed many times on the training dataset. The EPA has proposed to improve the PA classification of hidden data by combining it with the BA. The average AUC is obtained at 0.75 by EPA compared with 0.72 by PA.

SVM Machine Learning model was proposed by Mohan and Jain [7] for diabetes prediction. SVM is trained with four different kernels, and its prediction accuracy is calculated through the test set. The SVM is tested with four kernels: linear, polynomial, sigmoid, and RBF. The accuracy level of RBF is obtained at 0.82, which is higher than other SVM Kernels.

Vijayakumar et al. [8] present a random forest algorithm used of diabetes prediction. The projected model provides the most satisfactory result for diabetes prediction, showing that the forecast system can effectively and efficiently predict diabetes. This paper uses a different algorithm. Naive Bayes showed an accuracy of 80.37%, while the REP tree recorded up to 78.5%. Logistic regression showed 77% accuracy. The random forest algorithm analysis results were 83%, so the implemented methodology works relatively well.

Prabhu et al. [9] proposes Deep Belief Networks (DBN) for identifying diabetes. DBN forecast model is created in three phases; first, the preprocess is done for the noise removal. They were second, preprocessing Deep Belief Networks followed by fine-tuning back propagation. Finally, supervised classification provides result to identify diabetes.

Choubey et al. [10] proposed diabetic diagnosis is made in two stages. In the beginning, to find features from diabetes data, a genetic algorithm is used. Second, a multi-layer feed-forward neural network is used to predict diabetes. The tentative results are capable of providing a better result when compared with other models.

Syed et al. [11] explore the Tree Partitioning Adaptation Support Vector Machine method presented in the proposed model, which utilizes the tree approach. Find the right classified partitioning for big data. An adaptive SVM that follows the rules generated by partitioned data is also designed to handle the final classification.

Jayalakshmi et al. [12] proposed the Levenberg–Marquardt method for back propagation for the finding of diabetes. The primary approach adopted in this paper was to compare the results of various forms of lost value analysis [12]. This paper shows that higher accuracy is obtained with a minor training point. This paper demonstrates to use a “combination of missing values and preprocessing techniques” [12] to achieve improved results.

Messan et al. [13] compare various algorithms for the detection of diabetes. In this paper, the Artificial Neural Network (ANN) provides 89% of the highest accuracy compared with the Gaussian mixture model (GMM) is a probabilistic model, Extreme Learning Machine (ELM), logistic regression, and SVM.

Erdem et al. [14] proposed Genetic Programming and Majority Voting to predict diabetes. In contrast to SVM, ENN ANN, and other algorithms, which rely heavily on computational resources, GP Symbolic Regression’s simple equation classifiers have sufficient accuracy and minimal computational requirements for prediction. It may be an alternative for predicting diabetes.

4 Diabetes Prediction

Diabetes is an incurable disorder caused by deficiency or loss of insulin in body. Human cells take up glucose from people's food and can get the required energy. If left unchecked, diabetes is fatal and can cause other illnesses, either explicitly or indirectly, such as heart attack and stroke. Thus, the diagnosis of diabetes and assessment of severity are essential to avoid further problems and take preventive measures and the timely onset of the disease. Healthcare companies collect vast amounts of data, such as electronic databases, photos, weak data, and text, but gaining insight into information and data remains a significant challenge. Various papers suggest efficient Machine Learning approaches for predicting diabetes and assessing its severity.

4.1 *Unsupervised Machine Learning*

Unsupervised Machine Learning does not use labeled records. Unsupervised Machine Learning requires analyzing accurate data to discover the relationships, models, and correlations in data, investigate and learn the relationships in the data. Unsupervised machine learning is helpful for data with low-level knowledge. For example, answering a question such as "What are the gene expression patterns in a disease?" Two everyday unsupervised Machine Learning tasks are data clustering and data dimensionality reduction.

Clustering is the process of finding the relationship between data in unlabeled data to group similar data elements into a group. Several types of clustering techniques exist, and each method uses a different concept or set of rules for defining the similarity between data points. Figure 3 shows various clustering techniques. The most common use of disease prediction is clustering data into expression data. Expression levels of many variables can usually be measured with several samples of microarrays and the clustering to cluster data with similar levels of expression in all instances.

4.2 *Supervised Learning*

Supervised learning (SL) is a classification of data process for learning functions that map inputs to outputs based on exemplary input/output data. It derives a method from labeled training data containing a series of training models. Each example is a pair of inputs and desired outputs. Any supervised learning algorithms examine the trained data and generate a derived function that can map new models. The best scenario allows the algorithm to find the class label for hidden instances properly. This requires the training algorithm to be "reasonably" generalized to situations

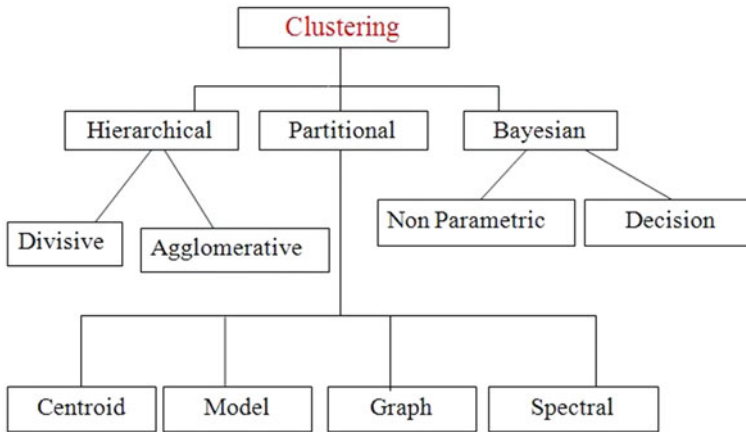


Fig. 3 Clustering techniques

that are invisible from the training data. Various supervised learning algorithms can perform multiple methods for separating test data by comparing it with the trained or labeled data.

5 Comparative Study of Survey

The comparative analysis of the survey explores the different author's views and the methodologies used for detecting diabetes. Table 1 displays comparative study of survey. The first column is the author's names, the second column is using previous algorithms, and the third column is the limitations of that works.

6 Results and Discussion

In this section, the classification algorithms for predicting diabetes are evaluated with various algorithms.

The comparison with different methods was assessed using multiple accuracy procedures such as Recall (Rec), Precision (Pree), F1-score, support, etc. Figure 4 represents the comparison between some of the algorithms, and their proposed level of accuracy is displayed. This figure shows various classification algorithms and proposed DBM methods such as, radial basis feed-forward neural network-(RBFNN), Naive Bayes, decision tree (DT), SVM, logistic regression (LR), random forest (RF), etc. From Fig. 4, DBN gives better results. So, DBN is the best algorithm for diabetic prediction from this review. Naïve Bayes provides the second level of good accuracy. HRPFNN delivers the next level of good accuracy.

Table 1 Various authors view diabetes prediction

Author's name	Methodology	Limitations	Paper number
Roxana et al.	The Ensemble Perceptron Algorithm (EPA) has been projected to enhance the PA classification of hidden data by combining it with the BA	Both algorithms are, EPAs require more execution time than Pas	4
Narendra et al.	SVM machine learning algorithm is used for diabetes prediction. The novelty of this task is to suggest the performance of the SVM kernel for diabetetic datasets for disease prediction	It is not compared to SVD evaluation indicators such as prediction, recall, and F measurements	5
Vijiya Kumar et al.	For finding the diabetes, random forest algorithm is used. The subset trees are formed by iterating the data many times	Application classification problem can be used to avoid over conference shortcomings and distinguish the essential options from the coaching dataset	6
Prabhu et al.	DBN algorithm is designed in three stages. The first is preprocessing and second is pre-training, and the last is DBN. After that, fine-tuning back propagation for supervised classification	Precision result of the DBN does not provide better results compared with other supervised learning algorithms	7

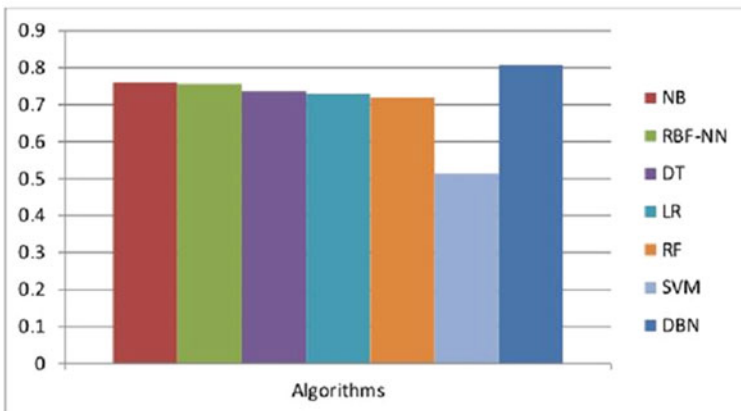


Fig. 4 Comparison of different algorithms for diabetes prediction

7 Conclusion

Non-insulin-independent diabetes presents inevitable and prominent problems to the public during periods of inadequate effectiveness, timeliness, and non-invasive price for the poor quality of life. Today, the risk factor for diabetes is increasing day by day. This review presented a complete study of current methods of predicting diabetes. The Deep Belief Predictive Model has been successfully validated with the Performance Benchmark using the diabetes dataset. The resulting comparison illustrates that the deep belief new model is more effective in accuracy, memory usage, and the WA measure.

References

1. <https://www.who.int/news-room/fact-sheets/detail/diabetes>
2. Qurat-UI-Ain, Anwar F, Ejaz MY, Mosavi A (2020) A comparative analysis on the diagnosis of diabetes mellitus using different approaches—a survey. *Inform Med Unlocked* 21:100482
3. Indoria P, Rathore YK (2018) A survey: detection and prediction of diabetes using machine learning techniques. *Int J Eng Res Technol (IJERT)* 7(3)
4. Sumitha J, Devi T, Ravi D (2017) Comparative study on gene expression for detecting diseases using optimized algorithm. *Int J Hum Genet* 17
5. Sumitha J, Devi T (2016) Breast cancer diagnosis through analysis of BRCA gene using machine learning algorithms. *Pakistan J Biotechnol* 13(4)
6. Mirshahvalad R, Zanjani NA (2017) Diabetes prediction using ensemble perceptron algorithm. In: 9th international conference on computational intelligence and communication networks (CICN), Cyprus, Turkey
7. Mohan N, Jain V (2020) Performance analysis of support vector machine in diabetes prediction. In: 4th international conference on electronics, communication, and aerospace technology (ICECA), IEEE
8. Vijiya Kumar K, Lavanya B, Nirmala I, Sofia S (2019) Random forest algorithm for the prediction of diabetes. In: IEEE international conference on system, computation, automation, and networking (ICSCAN), IEEE
9. Prabhu P, Selvabharathi S (2019) Deep belief neural network model for prediction of diabetes mellitus. In: 3rd international conference on imaging, signal processing and communication (ICISPC), IEEE
10. Choubey DK, Paul S (2016) GA_MLP NN: a hybrid intelligent system for diabetes disease diagnosis. *Int J Intel Syst Technol Appl*
11. Syed R, Gupta RK, Pathik N (2018) An advance tree adaptive data classification for the diabetes disease prediction. In: International conference on recent innovations in electrical, electronics & communication engineering (ICRIEECE)
12. Jayalakshmi T, Santhakumaran A (2010) A novel classification method for diagnosis of diabetes mellitus using artificial neural networks. In: International conference on IEEE data storage and data engineering (DSDE)
13. Komi M, Li J, Zhai Y, Zhang X (2017) Application of data mining methods in diabetes prediction. In: 2nd international conference on image, vision, and computing (ICIVC), IEEE
14. Erdem MB, Erdem Z, Rahnamayan S (2019) Diabetes mellitus prediction using multi-objective genetic programming and majority voting. In: 14th international conference on computer science & education (ICCSE)
15. Sumitha J, Nathiya S (2021) A comparative study on breast cancer prediction using optimized algorithms. In: 2nd international conference on smart electronics and communication, ICOSCE

16. Sumitha J, Mallika R (2011) Cancer classification in microarray data using gene expression with SVM OAA and SVM OAO. *Int J Adv Res Comput Sci* 2
17. Chandy A (2019) Smart resource usage prediction using cloud computing for massive data processing systems. *J Inform Technol Digit World* 1:108–118

Sound Mind: Detecting Depression and Anxiety in Humans



S. Revathi, K. Keerthanaa, A. R. Ranjitha, and Priyanka Balli

Abstract Depressive disorder commonly known as depression is an illness or disease that makes the individual to feel constant sadness, emptiness and a lack of interest in day-to-day activities. Around 3.8% of the global population is affected by depressive disorder, wherein 5% are adults and 5.7% are older than 60 years. Globally, around 280 million people suffer from depression. This research work attempts to detect the level of depression and anxiety in humans with the help of speech recognition technology by utilizing the audio recording of the participants, which is usually composed of both depressed and healthy audio notes. The proposed study analyses the implementation of Naive Bayes (NB) algorithm to detect the level of depression. It uses certain repeated words as parameters in the given audio clip to determine the depression level of a person. In near future, this research study will be further extended to incorporate real-time recording in order to detect the level of depression in a person.

Keywords Depression · Speech recognition · Humans · Acoustic features · Anxiety · Emotions · Machine learning · Naive Bayes algorithm

1 Introduction

Depression is by far the most serious global challenge. Though it is said that the earlier detection of depression increases likelihood of people returning to traditional life and performance, developing and utilizing effective and possible strategies for depression detection remain difficult to achieve [1]. Speech-based technologies have shown enormous potential and possess the additional benefits of being non-invasive and widely accessible. Combining these discoveries with the potential of speech analysis has revealed numerous new openings for implementing an effective efficacious depression and anxiety detection technique based on real-life exploitation time period information. On the other hand, smartphone sound-based apps must cross

S. Revathi · K. Keerthanaa (✉) · A. R. Ranjitha · P. Balli
NHCE, Bengaluru, India
e-mail: keerthanaa1502@gmail.com

few hurdles, such as loud areas in the surrounding and wide range of characteristics. The fluctuation of voice quality in unattended environment poses big challenges for standard acoustic options such as spectral, glottal, and conveyance outlines [2].

Hence, when compared to the research centre grade clean discourse data, acoustic preferences may perform skill-degraded execution due to their susceptibility towards natural calamities and mobile inconstancy. Recently, the discourse milestones, which are referred as the occasions associated with discourse explanation, have demonstrated the viability for analyst available in a practical context [2, 3].

The significant component of the emotion detection framework is the selection of appropriate preferences to categorize different emotions. The completely covered boundaries, such as the energy of pitch, signal, timer and MFCC, are more popular than words. As these boundaries shift indefinitely over time, the sound gets separated into appropriate-sized covers. It is also expected that the boundaries mentioned above will not change significantly over the course of a casing [4]. Depending on the severity of disease in several humans with various issues, they can be divided into three categories, namely mild, moderate and severe. It is the most vastly used categories for the categorization of depression and anxiety. Mild depression is actually having a low interest in the surroundings around the affected individual, but it doesn't affect the normal life and social exchanges. Moderate depression includes emotional fluctuations, but the state of mind is typically depressed, where the thought process and actions have a slower response, which negatively impacts the work life. The extreme form of a major depressive disorder is suicide [1]. Normally, mild depression can be cured by being mindful, considerate, patient and surrounding oneself with positive people and engaging in activities that improve the emotion and by practicing mindfulness. Moderate and severe depression conditions should be taken more seriously and seek professional treatment in time. As per reports of cases, this disorder gets split into one episode depression type and recurrent episodes of depression. One episode depression type happens rarely for the diseased individual, and it doesn't resurface upon finishing the prescribed treatment, but this is not clinically typical. Recurring depressive disorder refers to the recurrence of symptoms even after giving treatment, and the issue hasn't been eradicated [5].

2 Related Works

Little kids and teens with emotional disorders and anxiety, adult depression and senile depression are the various age groups of individuals suffering from this disorder [1]. The age span of it is large. This however can't be said that young people do not suffer from it. Therefore, the categorization of this view is made according to the different age groups of this subject [6]. Depression patients have tell-tale speech characteristics in speech signals, and the further study on recognizing this illness is based on speech signals which have a solid theoretical basis; hence, the depression recognition based on speech signals is considered as the best method. In contrast to the complex data collection process of physiological signals and behavioural signs, currently, there are

Table 1 Table to identify the depression severity

Measurement score	Analysis of measurement results	Depression treatment recommendations
0–4	No depression	None
5–9	Mild depression	Wait for observation
10–14	Moderate depression	Counseling or mild medical treatment
15–19	Moderate to severe depression	Active medication and psychotherapy
20–47	Major depression	Recommended for professional treatment immediately

more number of open source and authoritarian depression speech datasets. Rosenthal and Ambady’s study on a human’s behaviour has revealed that a brief observation of human’s behaviour can predict their actual behaviour with a higher probability than random predictions. There lies several more studies available on speech slice in the domain of speech recognition. The effect of relative slices of time is considered as best algorithm for performing speech-based emotion recognition. It was discovered experimentally that when reading out loud by using the starting of the speech has produced a better depression recognition effect than using the entire speech. Research of Moore on glottic features shows that when using glottic features for depression recognition, only, 10–20 s of speech portions is utilized to achieve a recognition effect comparable to 2–3 min of speech portions [7]. Table 1 shows the scale of scores for the various levels of depression.

2.1 Machine Learning Model for Recognition

Currently, the most vastly used depression recognition models based on machine learning are GBDT, KNN, SVM, NBM, RF, etc. The flowchart for depression recognition by using machine learning technique is given in Fig. 1.

2.2 Deep Learning Model for Recognition

As our information becomes very diverse, the data processing becomes more prominent and vital. DNNs are powerful in processing the data and are considered as the powerful tool for processing complicated data. Pitts puts forth the human brain operation to process and store complicated information. It processes data in a parallel manner and obtains more useful info from it. CNN is one of the vastly used depression recognition models. The flowchart for depression recognition using deep learning is given in Fig. 2 [8].

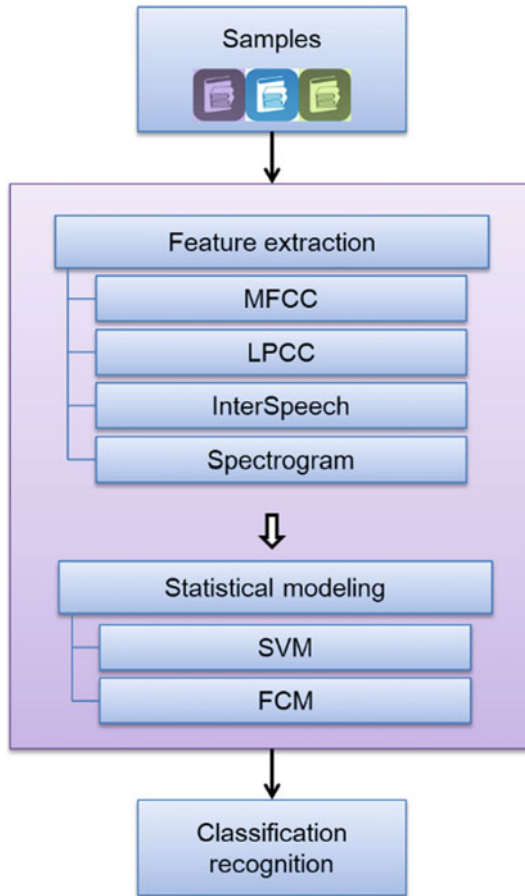


Fig. 1 Machine learning model

2.3 Recognition of Traditional Depression

Typical tools of diagnosis are interview assessments, Hamilton depression rating scale or HAM-D and depression scale. These assessments and the severity of behavioural symptoms will provide a depression score to the patients. The diagnosis is then performed by this method which is very complex as it relies mainly on the human’s ability, their patience, willingness, honesty and seriousness whilst communicating. However, since their thinking and motives are impacted and affected by depression, collecting the diagnostic information consumes more time. In addition, human training, practices, and validations are required to achieve results. At the primary treatment stage, this method of diagnosing depression is extremely difficult. Depression diagnosis can also be very broad. Furthermore, the results of such classification are much generalized and can only inform patients whether they have

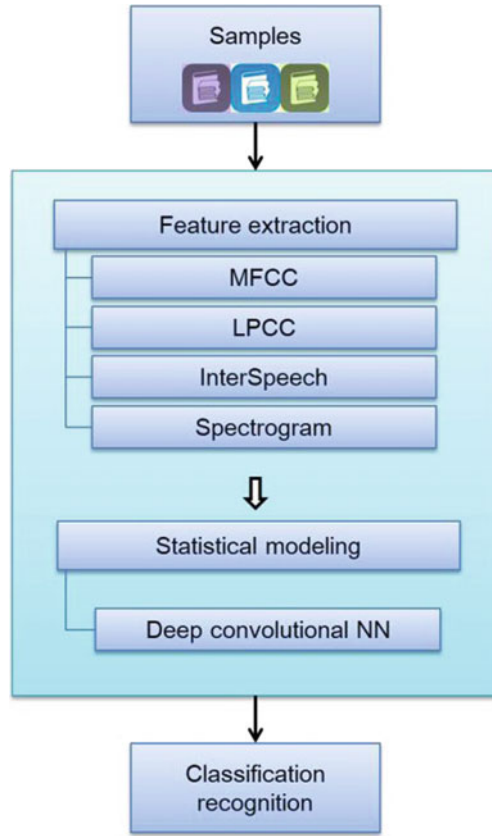


Fig. 2 Deep learning model

high or low levels of depression, which can easily lead to misdiagnosis and waste of treatment time. Simultaneously, due to the obscured nature of depression’s physical symptoms, it cannot always be detected accurately. These are referred as the notable drawbacks [9].

2.4 Deep Neural Network

The features that are primarily used to define and display the content are investigated in pattern recognition. These features are then extracted from the data by using this algorithm have an impactful effect on the model’s performance. Typical feature methods of extraction are to primarily select features depending on various data designs and tasks, such as Gabor features, SIFT, local binary patterns and so on. However, designing a decent feature isn’t easy, specifically when dealing with a large

dataset; it requires a significant amount of time and effort. DNNs refer to a subclass of deep learning in terms of DL multi-layer nonlinear function relationships [10]. A DNN is a neural network that maps a learned function to a highly complicated function by using a nonlinear combination. DNN is referred to as a neural network with multiple obscured layers. Deep neural network (DNN) utilizes minimal quantity of neurons to achieve better capabilities at generalization [11].

2.5 Depression Severity Scales

Many scales are utilized for describing how severe the depression in a human is. This study has made use of a couple of the most vastly used scales, namely Beck Depression Index (BDI) and Hamilton Depression Rating Scale (HAM-D). Both of these scales are often used by clinical practitioners, but they have clear differences. The BDI is a self-assessed questionnaire that centres around on the human's negative self-esteem symptoms, where the HAM-D score can only be determined by a psychiatrist. As a result, the HAM-D score is regarded as the more objective in both the methods. Determining the HAM-D score is hard, but it is not because it needs to be done by a medical professional, but also because the questionnaire can be completed in a short period of around five minutes, whilst the HAM-D score can be detected in around a half hour. It is crucial to observe that the both these scales are used to measure the severity of depression in various ranges [12].

3 Literature Survey

Ding et al. [1] discuss about detecting the different depression levels of college students, and a novel method has been proposed to detect the student's depression by using the Sina Weibo data (DISVM) to make the model more stable and improve the accuracy. NLP has also been used, and the drawback is that the accuracy can be improved. Kiss and Jenei [2] have investigated the different ways in which the depression severity can be accurately detected by processing the speech signals. The developed model can comprehend the depression qualities and the impact on speech. This acoustic model has also predicted the depression severity with a lower error rate than the BDI questionnaire (the classic way to identify depression). The drawback is that the size of dataset is limited and can be further expanded. An et al. [3] have explored the mental health detection method by making use of CNNs for processing the speech signals. It highlights certain advantages which CNNs offer for the detection of mental health. Huang et al. [4] propose a method to ensure the accurate diagnosis of emotion disorder. The methods used here possess bottleneck features and HMM-based verification, which provides a decent accuracy. Kiss and Vicsi [5] have proposed a cross-lingual investigation to check the possibilities of depression prediction. Here, the acoustic elements are selected as the info vector,

and numerous mono and cross-lingual examinations were performed. Huang et al. [6] proposed a novel framework to analyse the discourse as a series of acoustic occasions followed by an investigation on the applications used for detecting depression. Song et al. [7] have discussed about a proposal from one viewpoint, which recognizes the human conduct automatically. Two different methods have been used to achieve a good improvement when compared to previous estimations. Long et al. [8] have introduced a novel classifier framework to perceive depression, which was then evolved and combined different speech types and emotions. Yang et al. [9] have proposed new written, sound and video highlights which hybridizes the deep and shallow models to gauge depression and arrange from sound, video and written descriptors. Alghifari et al. [10] investigated the optimum speech segment length to provide a fast and accurate detection of depression. The proposed ANN system uses a single speech feature to detect depression. Govindaswamy and Palanichamy [11] detected the depression of different users, which has been shared on social media. Han and Wang [12] have developed a new emotion recognition technique based on DL and kernel nonlinear PSVM. Shih et al. [13] introduced weights to train examples and successfully modify the traversal of trajectory in parameter space whilst executing the learning process. Suganya and Charles [14] proposed that the model's performance should be analysed on USC-IEMOCAP and EmoDB datasets. This CNN model consists of 9 layers of weight. Harar et al. [15] utilized the voice activity detection (VAD) algorithm to get rid of silent segments and detect the depression state of the user. Pinto et al. [16] presented an emotion detection model based on CNNs, by considering the data from an existing dataset and classifying emotions into different groups. Gong and Poellabauer [17] discussed about displaying a system for recognizing the second-level states of human emotions by using voice recordings. This makes use of SVM. Harati et al. [18] have developed a predictive model based on emotion-related features by utilizing the connection amongst present moment, long-term emotions and state of depression to detect the current level of depression. Sefara and Mokgonyane [19] have discussed about using the machine and deep learning algorithms to analyse frequency, time and spectral features of emotions in order to detect any existing depression. Swamy et al. [20] have discussed about developing a new platform to detect the presence of depression without requiring any physical presence of a professional.

4 Methodology

Naïve Bayes computation is a directed learning estimation by considering Bayes speculation. This is used for handling the data collection challenges. This classifier is considered as one of the essential and best classification computations which assist in building advanced AI models, which can make quick conjectures. It is a probabilistic classifier, which means it makes predictions based on the likelihood. As a result, this type of computation is used to identify the problem of depression in individuals. The Naive Bayes computation is made up of two words: Naive and

Bayes, which can be written as *Naive*: it anticipates that the event of a specific component will be liberated from the event of other components. As a result, each component independently continues to add in order to recognize that it is a specific thing without relying on others. It considers each feature as distinct from the others. *Bayes*: depends on the rule of Bayes speculation. The proposed study utilizes a training and testing model. First, the discourse sound clips will be considered in the preparation model, which will then use specific words as boundaries to distinguish whether a specific recording contains some degree of sadness. If the boundary word has rehashed a certain number of times, it belongs to one class of wretchedness; if it does not rehash, it belongs to another class of misery, and so on. The testing model is then initiated by entering any of the recordings, and the downtime forecast gets started. This calculation is probably the best option for predicting despondency with reasonable accuracy. As a result, the proposed study employs this specific Naive Bayes calculation.

5 Architecture

The depression detection and analysis takes place in a series of steps, out of which the first and foremost one is to collect the data samples for performing observation and training. Once the data collection is performed, it is followed by the data cleaning and normalization process, which typically results in removing the unnecessary noise and audio bits. After completing these steps, the feature extraction process has been performed. Here, the required words were selected as parameters for classifying the audio sample into different categories, and if present, the frequency of the word is calculated and noted. This is followed by the category classification of whether a person is depressed, not depressed or borderline. These observations are then fed as training data to the model, which then performs the analysis and produces the outcome of the detection. Figure 3 shows the architecture diagram which is shown below.

5.1 Data Flow Diagram

The DFD, called the bubble chart, is a fairly simple graphical format used to represent the system according to the data input to the system, the numerous processes performed on this dataset and the result data generated by this model system. This DFD is an important modelling tool for simulating the system. The process of this system, the data this process uses, as well as an external entity that interacts with this system and the information which flows in this system are all considered as components.

DFD displays how this info is moved along in the system and how it gets modified by a series of transformation. It is a graphical technique used to depict the information

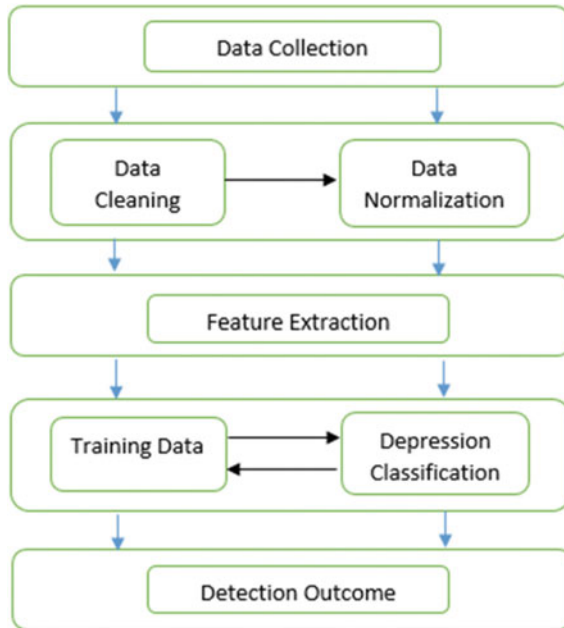


Fig. 3 Overall architecture of detecting depression in humans

flow, and the transformations are then applied as the data, which moves from input to the output. A DFD can be used to represent any level of abstraction in a given system. This can also be portioned into different levels, which represent increasing information flow and functional detail. Figure 4 illustrates the same.

5.2 UML Diagram

UML is an abbreviation for unified modelling language (UML). In the field of object-oriented software engineering, UML is a standardized general-purpose modelling language. The object management group has managed and created the standard.

The main goal of UML is to become a standard language for modelling object-oriented computer software. UML is currently made up of a couple of major components which are a notation and a meta-model. In future, a process can be added to UML. The UML is a standard language used for specifying, visualizing, constructing, and documenting software system artefacts and also the modelling of business and other non-software systems. It represents a collection of best engineering practices, which has been to be successful in the modelling the complex and huge systems. The UML plays a very important role in developing the software that is object-oriented

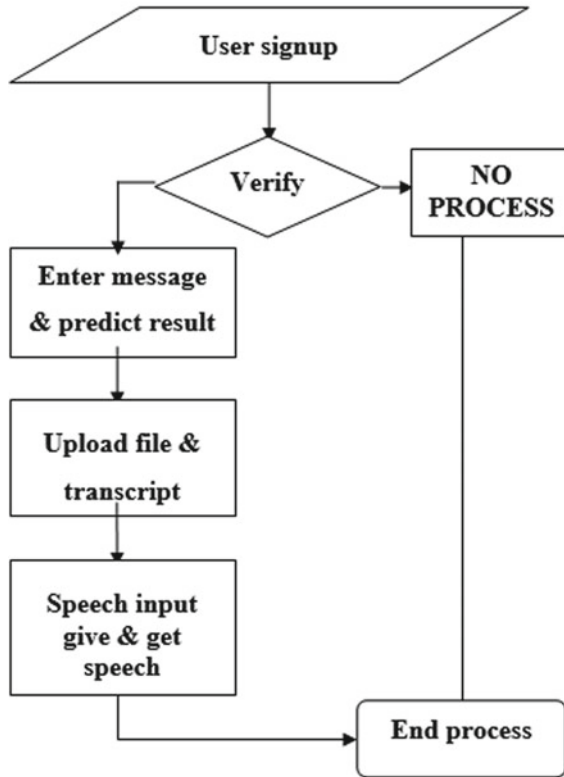


Fig. 4 Data flow diagram architecture

and the development process of this software. It also uses notations of graph to express the software projects design.

The main goals in the design of the UML are as follows:

1. Give users a ready-to-use, expressive visual modelling language which they can utilize to exchange and build meaningful models.
2. Offer extendibility apart from specialization mechanisms so that they can extend it for the core concepts.
3. To become independent of certain programming languages and processes of development.
4. To understand the modelling language, offer a formal basis.
5. OO tools in the market need to be encouraged.
6. High-level concepts like framework, patterns, collaborations and components have to be supported.
7. Best practices must be integrated.

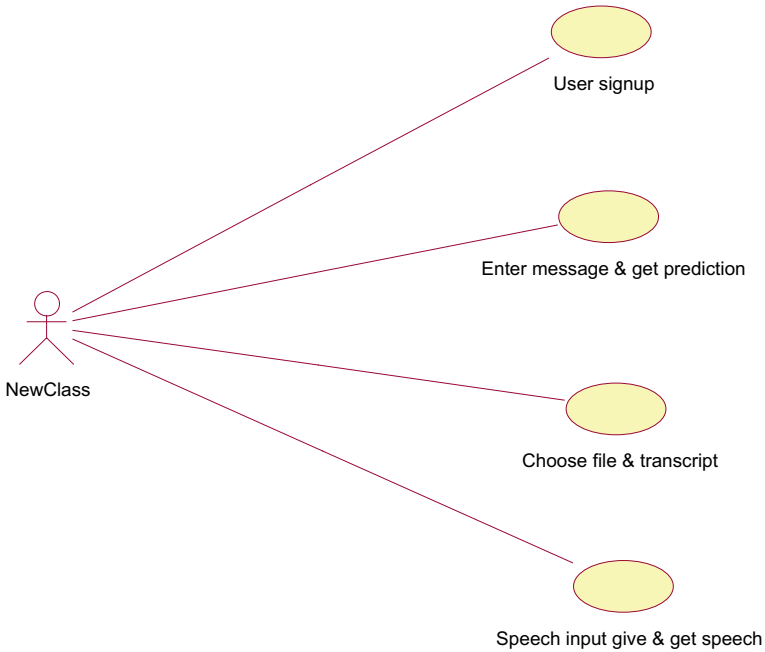


Fig. 5 Use case diagram architecture

5.3 Use Case Diagram

This diagram in the UML is a behavioural diagram type that is defined by and created from a use case analysis. The main motive is to give an overview in the form of a graph about the functionality given by the system which is actors and their goals, as well as any dependencies amongst the use cases. The main cause of this diagram is to display the functions performed of the system and by which actor it has been done. The actors’ roles in the system can also be depicted. Figure 5 depicts the same.

5.4 Class Diagram

This Fig. 6 is used for refining the use case diagram to define an elaborate design of the system. This diagram takes the actors from the use case and sets them into classes that are inter-related. The association between the classes can be an “is-a” or “has-a” relationship. Every class in the diagram is capable of providing certain functionalities. These functionalities provided by the class are termed “methods” of the class. Apart from this, each class might have certain “attributes” that uniquely identify the class.

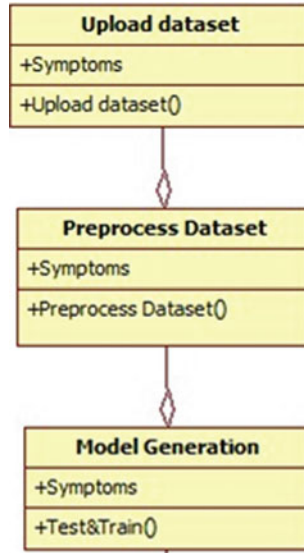


Fig. 6 Class diagram architecture

5.5 Activity Diagram

This diagram Fig. 7 captures the flow of process happening in the system. It is quite similar to the state diagram as it also has actions, activities, transitions and the initial-final states as well as the guard conditions.

5.6 Collaboration Diagram

This Fig. 8 groups the interactions between various objects. These interactions are numbered, which is used to trace the interaction sequence. The diagram helps to determine all the potential interactions that each object could have with other objects.

5.7 Sequence Diagram

This Fig. 9 showcases the interaction between different objects present in the system. The crucial part of this is the fact that it is time oriented, which means that there is a step-by-step representation of the sequence interaction amongst the objects. Message passing mechanism is used to interact with the objects.

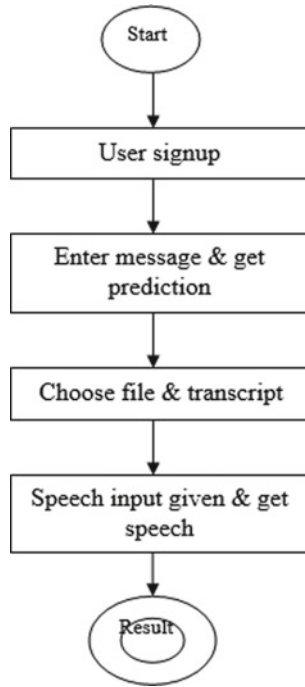


Fig. 7 Activity diagram architecture

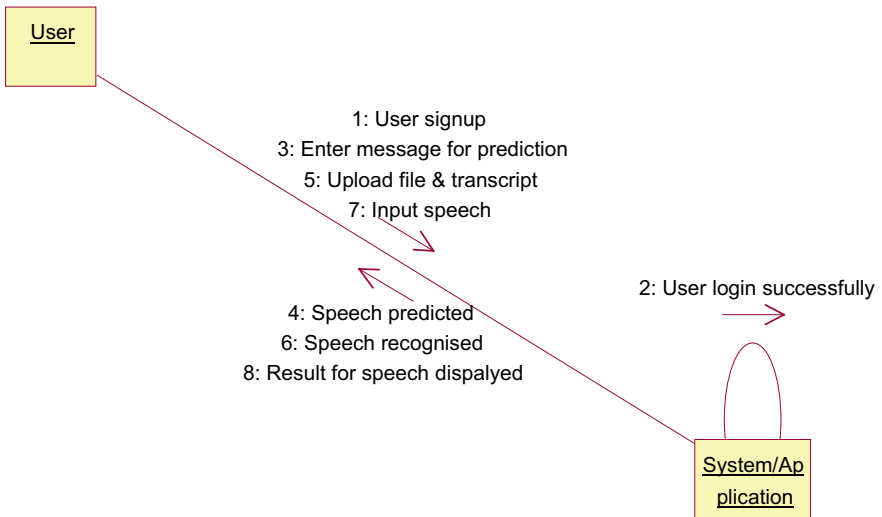


Fig. 8 Collaboration diagram architecture

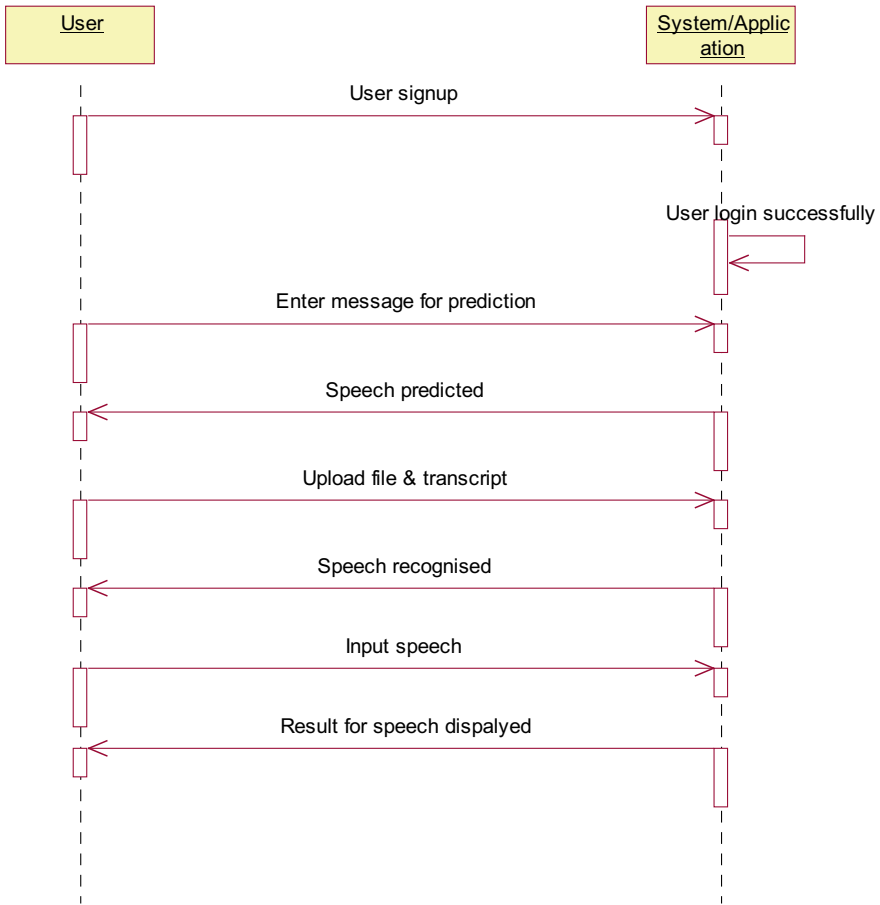


Fig. 9 Sequence diagram architecture

5.8 Component Diagram

This Fig. 10 showcases the high-level parts that build up the system. This diagram displays what are the components that form parts of the system and their connections. A component diagram shows the components compiled after the system has undergone the construction phase.

6 Studies of Algorithms

Acknowledgment precisions of four correlation calculations on the preparation set are given below. We have seen the various architectures and methods used for performing

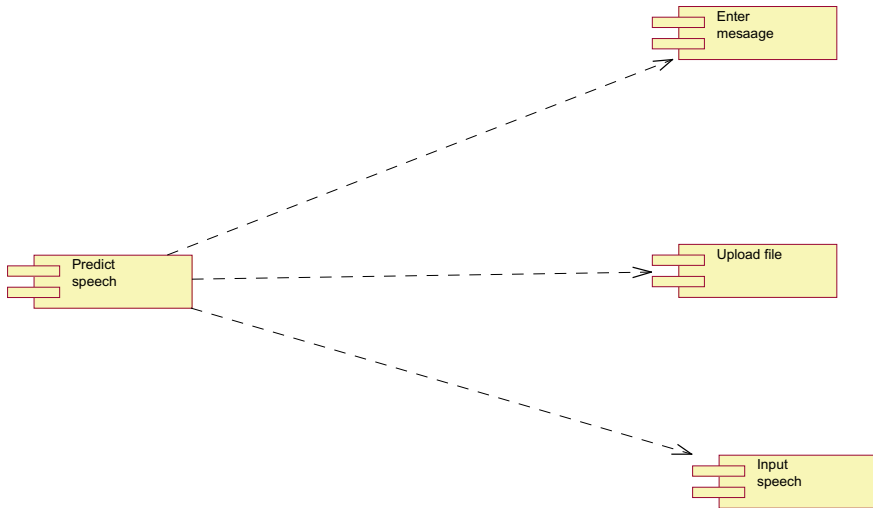


Fig. 10 Component diagram architecture

Table 2 Studies of different algorithms

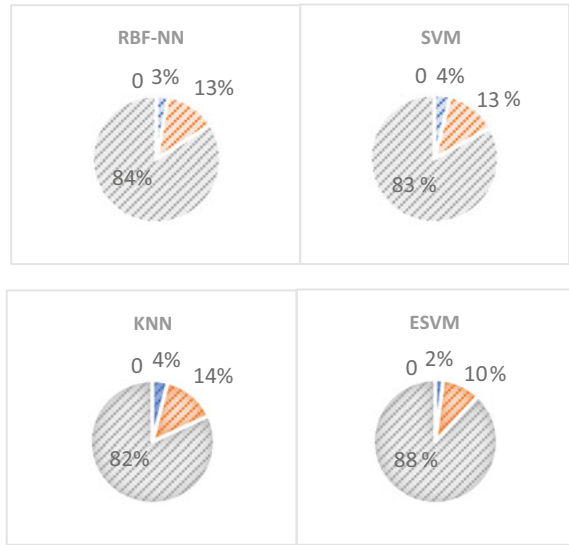
Algorithm	Normal users identified as depression users	Depressed users identified as normal users	Users classified correctly	Precision
RBF-NN	73	19	471	0.8366
SVM	76	21	466	0.8277
KNN	81	23	459	0.8152
DISVM	56	11	496	0.8810

the same analysis. Now, we will do a comparison study of the different algorithms used for detecting depression [13]. Table 2 and Fig. 11 show the studies of numerous algorithms, their precision percentages.

6.1 Analysis and Comparison

The request eventual outcomes of a couple of assessment computations on the arrangement set are shown in Table 2 and Figs. 11 and 12. In this planning set, there are 563 clients, 78 of whom have a pity propensity and 485 of whom are ordinary. Table 3 and Fig. 12 demonstrate the accuracy of recognition on the test set [14, 15]. The test group consisted of 130 persons, 30 of whom were gloomy and 100 of whom were normal. Each classifier’s introduction on the planning set is usually superior to the test set [16, 17]. This DISVM computation put forth by this paper can

Fig. 11 Precision percentage



give low weight to models of poor quality, thereby increasing the model confirmation rate. The table’s results show that the proposed model has a good affirmation sway in both the readiness and test sets. Distinguishing proof exactness of the four correlation calculations are available on the test set [18]. Table 3 shows the comparison amongst numerous algorithms, and Fig. 12 compares their performances. Table 4 provides the analysis of the research.

Naive Bayes: This is an authentic collection based on the Bayes theorem, in which each component in the class is treated as independent of various attributes. A Naive Bayes classifier expects one item in a class to have a certain effect released from a different area of the body. Unknowingly, Bayes engages in a probabilistic strategy for predicting the test class data and a rule for predicting different classes. Regardless, Naive Bayes has a flaw in admitting the traits. The first condition depicts the state-of-the-object and is considered as one of the elements that are unlikely to have an impact in the authenticated version.

The Naive Bayes model is considered. Let us take note of $C.x_n$ is the component inputs and is a class imprint.

$$P(x|c) = P(x|c) = P(x|c) = P(x|c) = P(x|c) P(c)/P(x)$$

- $P(c|x)$ = the class’s back probability where C is the goal and x is the marker characteristics.
- $P(c)$ = the previous probability of the class
- $P(x|c)$ = Percentage of markers in each class.

The probability class is a subset of the probability class.

Fig. 12 Performance comparison

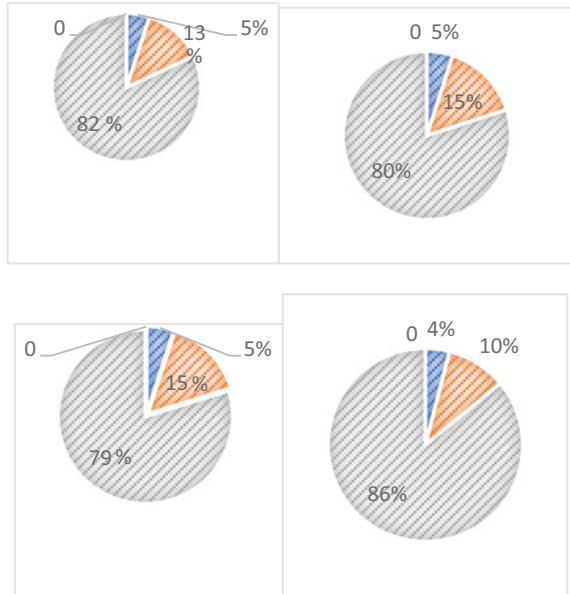


Table 3 Comparison of different algorithms

Algorithm	Normal users identified as depression users	Depressed users identified as normal users	Users classified correctly	Precision
RBF-NN	17	6	107	0.8231
SVM	19	7	104	0.8000
KNN	20	7	103	0.7923
DISVM	13	5	112	0.8615

- $P(x)$ = prior probability of the pointer

The probability is processed using this classifier following a few key developments:

Stage 1: Determine the prior from the given class names probability.

Stage 2: Calculate the probability of each class. Every trademark in the dataset has a chance.

Stage 3: Add consideration to the probability calculation in the Bayesian formula.

Stage 4: The class with the highest probability based on the data has a place with the higher likelihood class on given data.

From the analysis and comparison of the above papers, we can see that a number of methods and algorithms are used to detect the presence of depression in humans.

Table 4 Research analysis

Research paper	Calculation approach	Discoveries	Restriction
	<ul style="list-style-type: none"> Choice tree Straight SVM Guileless bayes Calculated regression 	Innocent Bayes gives high accuracy	<ul style="list-style-type: none"> Use of the old dataset Zero in on client admission
	<ul style="list-style-type: none"> SVM Gullible bayes 	Credulous Bayes gives high exactness	<ul style="list-style-type: none"> The dataset is in Arabic
	<ul style="list-style-type: none"> Choice tree K-nearest neighbour SVM 	The choice tree gives the best results	<ul style="list-style-type: none"> The information centre on the remarks
	<ul style="list-style-type: none"> NB-SVM hybrid model 	The accuracy of the model is 85%	<ul style="list-style-type: none"> Sets aside effort to compare the long-short scraps
	<ul style="list-style-type: none"> SVM Choice tree Innocent bayes 	SVM gives high precision	<ul style="list-style-type: none"> Use of the distinctive bit in SVM Can't keep away from overfit information

We can come the conclusion that overall the performance of Naïve Bayes algorithm has a better accuracy compared to the other procedures. Hence, in this paper, we will further investigate the utilization of Naïve Bayes and work on improving the accuracy to successfully detect depression in people through speech recordings.

The reason for choosing Naïve Bayes algorithm is due to its speed and the ease of its operation compared to other algorithms in this domain whilst simultaneously giving an accurate prediction [19, 20]. This algorithm performs well in real-world situations such as our depression detection in humans and also works best with categorical input which is the input type used in this project. The dataset required is also considerably less hence simplifying the whole procedure whilst still maintaining the quality and high-level classification [21]. Figure 13 shows the frequencies of the expressions, and Fig. 14 gives the formula used by Naïve Bayes algorithm. Table 5 shows the survey of various papers to review their algorithms and methodologies.

The equation utilized in this order calculation is as per the following:

7 Implementation

This paper reviews the Naive Bayes algorithm after comparing the reference papers as we can see from Table 5, which is referred as a classification algorithm to detect the level of depression in the person. It uses certain repeated words as parameters in the given audio clip to determine if someone has depression. The parameter words represent negative emotions such as anger, hurt and sadness, and these emotions can be conveyed through a wide variety of words out of which we have selected the

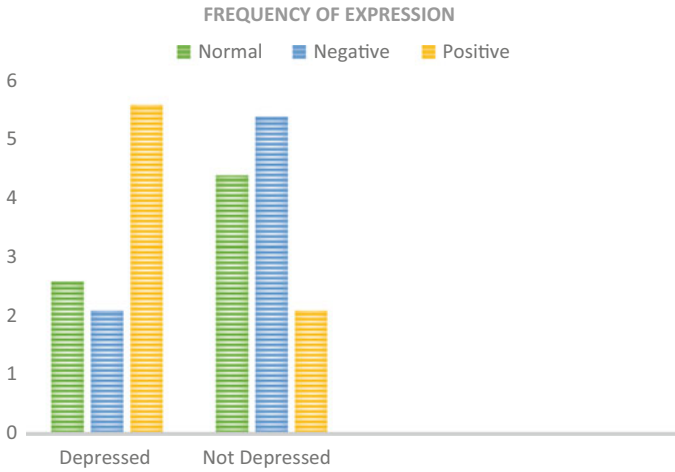


Fig. 13 Frequency of expression

Fig. 14 Naïve Bayes formula

The Formula For Bayes' Theorem Is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)}$$

where:

$P(A)$ = The probability of A occurring

$P(B)$ = The probability of B occurring

$P(A|B)$ = The probability of A given B

$P(B|A)$ = The probability of B given A

$P(A \cap B)$ = The probability of both A and B occurring

most commonly used words and phrases. The implementation has two phases, which will be observed in detail. The proposed dataset consists of various audio recordings collected from the common people, which will be given as input data to the training phase.

7.1 Dataset

This project uses a wide variety of dataset in the form of text messages in the system; pre-recorded audio files can also be uploaded into the system along with a provision for live recording and transcribing simultaneously. The Naive Bayes algorithm analyses the input data and detects the category of speech into one of the three categories “Depressed,” “Not Depressed” or “Borderline.” This dataset broadens the scope of detection by supporting both speech-based and transcribe-based input.

Table 5 Survey of algorithms and methodology

References	Year of publication	Author	Review	Algorithm/methodology	Advantages	Limitations
1	2019 IEEE	K. Y. Huang, C. Wu, M. Su, C. Chou	Proposes a method to make sure mood disorder is diagnosed correctly	Deep bottleneck features and HMM verification	Helps patient get correct diagnosis	The highest accuracy achieved is still not great, so better methods need to be found and used
2	2019 IEEE	Z. Huang, J. Epps, D. Joachim, V. Sethu	Proposes a framework to analyse speech like series of acoustic events and to investigate that application to detect gloom	Natural language processing algorithm	Non-invasive and has broad accessibility	Accuracy can be improved
3	2020 IEEE	Y. Ding, X. Chen, Q. Fu, S. Zhong	Proposes a strategy to successfully identify whether college students have depression or not	Uses Sina Weibo data for the detection of depression. Uses DISVM	Model is stable and accurate	Detecting the depression takes time here
4	2017 IEEE	G. Kiss, A. Z. Jenei	Investigates how depression severity can be accurately detected based on processing speech signals	Pre-processing, segmentation, acoustic features, model training, testing	Model was able to learn the depression characteristic	Can use some expansion as size is limited

(continued)

Table 5 (continued)

References	Year of publication	Author	Review	Algorithm/methodology	Advantages	Limitations
5	2018 IEEE	L. Yang, D. Jiang, H. Sahli	Takes 35 number of participants and classifies them as depressed or not depressed, out of which 6 of them were not rightly classified	Uses a deep CNN and a DNN	This model successfully increases accuracy of the detection estimation of depression	The performance is still not satisfactory and has scope for improvement
6	2020 IEEE	M. Niu, J. Tao, B. Liu, J. Huang, Z. Lian	Speech is divided to amplify spectrum into segments of fixed lengths in STA network. It employs eigen evolution pooling method and also represents using multimodal information	Novel spatial-transient consideration or STA organization and multimodal consideration include combination or MAFF technique	The quality of the representation is improved	The performance has room for improvement
7	2019 IEEE	M. F. Alghifari, T. S. Gunawan, M. A. W. Nordin, M. Kartwi L. Borhan	It looks into the length of optimum speech segment	Uses an ANN or artificial neural network	It is accurate and fast. Accuracy is 97.5–98.3%	Needs improvement in recognizing new speakers as input

(continued)

Table 5 (continued)

References	Year of publication	Author	Review	Algorithm/methodology	Advantages	Limitations
8	2017 IEEE	M. Ghai, S. Lal, S. Duggal, S. Malik	Feeling acknowledgment is done based upon different algorithms are presented	Mel frequency cepstral coefficients or MFCC	A high accuracy of 81.05% was achieved by using random decision Forest classification algorithm	Single emotion prediction is tough since it is difficult to determine their boundaries. Presence of noise is another issue
9	2020 IEEE	H. An, X. Lu, D. Shi, J. Yuan, R. Li, T. Dish	Speech that is segmented is showed as a spectrogram in STFT	CNN or convolution neural networks and STFT or short-time Fourier transformation	This method can directly be used by interchanging speech of emotions	
10	2020 IEEE	H. Long, Z. Guo, X. Wu, B. Hu, Z. Liu, H. Cai	Proposes a new multiplier classifier method for detection of depression	Short-time energy, LPC, MFCC, ZCR and LSP	Good accuracy of 78.02%	Accuracy can be increased further in future attempts
11	2021 SCOPUS	Jaya R., Thirukkumaran R., Saravanan S., Sivabalan N	Applying ML methodologies to predict the maintainability	Programming framework or part can be adjusted to address issues, further develop execution or different qualities or adjust to a changed climate	It assists professionals with working on the nature of programming frameworks and in this way streamline support costs	The performance is still not satisfactory and has scope for improvement

7.2 Training Phase

This is the phase where the proposed model is trained by feeding it with the input data so that the model can process it and understand what it needs to analyse. It uses all the audio files to train itself by referring to the parameter words and then detects whether the person in the recording suffers from depression or may be on the verge of falling into depression. The parameters referred here are certain words and phrases that display emotions associated with depression such as extremist words like “always, never, everything, nothing,” negative emotion words like “lost, stuck, helpless, stupid, worthless, meaningless” and so. Specific phrases like “what’s the point?”, “I give up,” “I feel better” are also considered to decide whether a person could suffer from possible depression.

7.3 Testing Phase

This is a phase where one particular audio file is entered, uploaded or recorded live for the system to analyse and detect the outcome of recording. This model has three categories and places each audio file into the respective category after analysing the input data. The categories are namely “Depressed,” “Not Depressed” and “Borderline.” The testing is done by comparing the parameters against their frequency. If a certain parameter word is repeated n number of times, then that audio file is declared as belonging to the depressed category. If those words are repeated for less than n number of times, then it belongs to the borderline category, and if the repetitions are negligible or non-existent, then the audio is labelled as not depressed. By using the Naive Bayes algorithm and its training and testing phase, we can detect the level of depression in humans with a decent accuracy whilst maintaining a more simplified approach requiring less time and effort compared to other algorithms of machine learning domain.

7.4 Results

Finally, in the final results as we can see from Figs. [15](#), [16](#), [17](#), [18](#), [19](#), [20](#), [21](#), [22](#) and [23](#), the given input is processed, analysed and then assigned to the most appropriate category. People can use this to learn about their mental health from the privacy and comfort of their own homes. As a result, it is considered as a non-intrusive method and a user-friendly platform to assist people in identifying any problems that they may have by taking appropriate action. The results of this study have been implemented by using machine learning algorithm, and it paves way for future innovations in the speech recognition technology by striving to improve the accuracy to its maximum

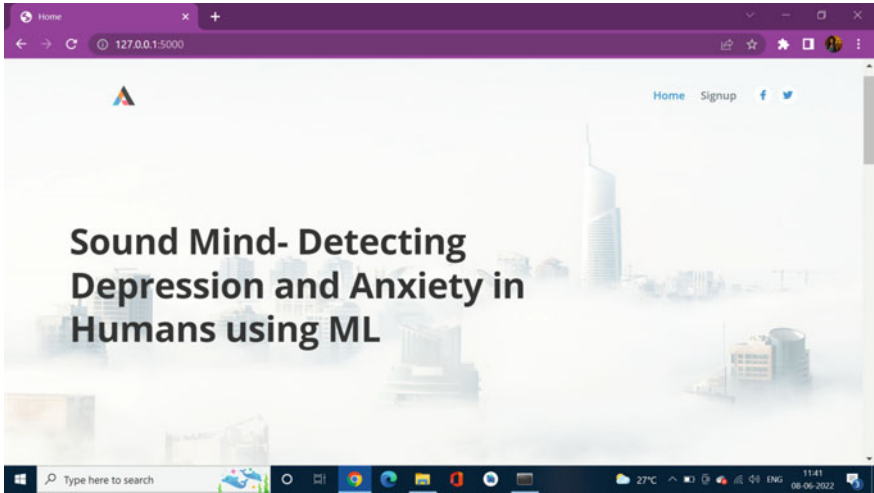


Fig. 15 Homepage part 1

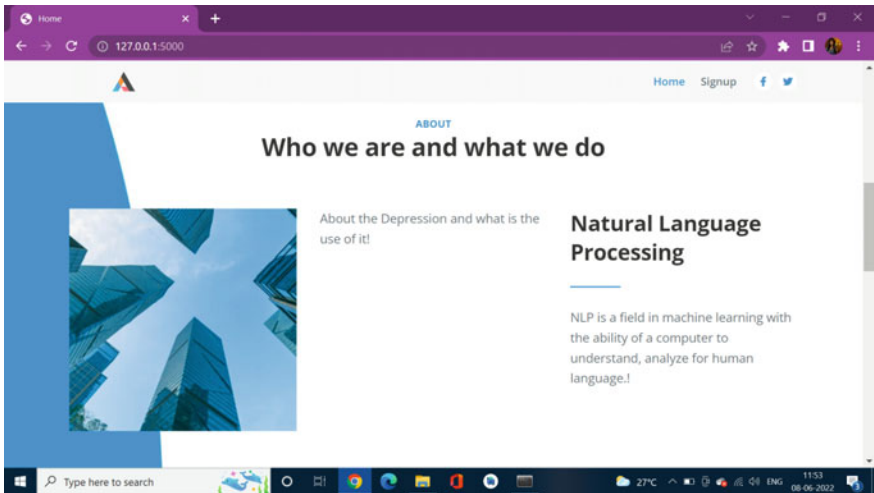


Fig. 16 Homepage part 2

capacity, expand the dataset to improve efficiency of learning to detect depression and to make the whole process more fast and user-friendly.

8 Conclusion

The proposed study has successfully presented a method for detecting depression in humans and categorizing the severity of the depression into three main labels:

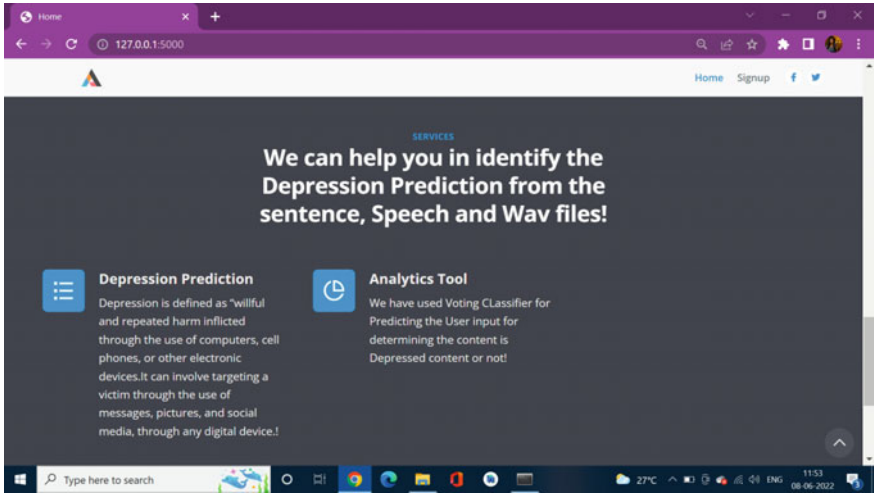


Fig. 17 Homepage part 3

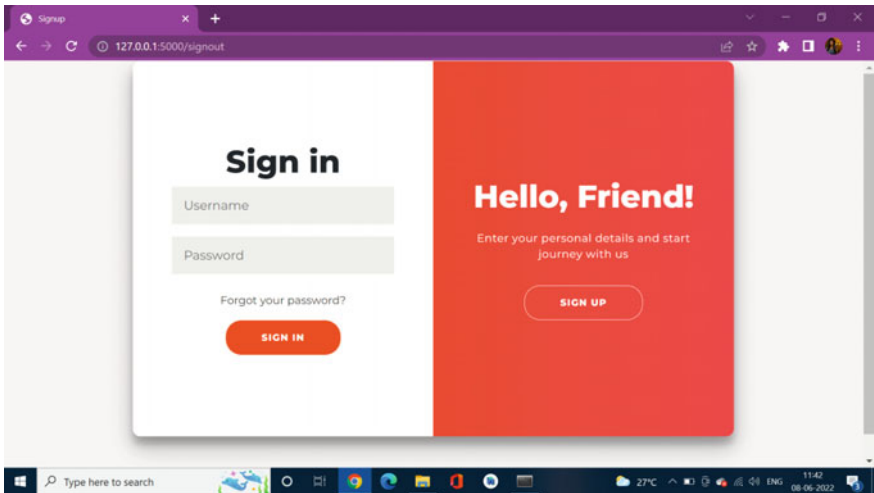


Fig. 18 Login page

“Depressed,” “Not Depressed” and “Borderline.” This research work has demonstrated that it is able to predict the severity of depression based on speech processing and recognition with precision by comparing different algorithms, and that the Naive Bayes algorithm is considered as the best option to detect the disorder of depression because it is simpler to implement than other algorithms whilst being nearly equal in terms of efficiency. The future work will be to look for more appropriate and sustainable techniques to improve on the current precision of this model, as well

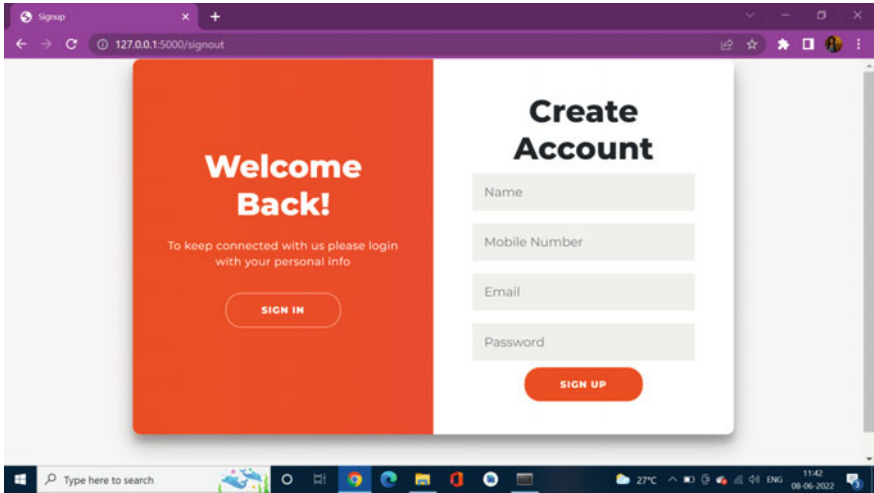


Fig. 19 Signup page

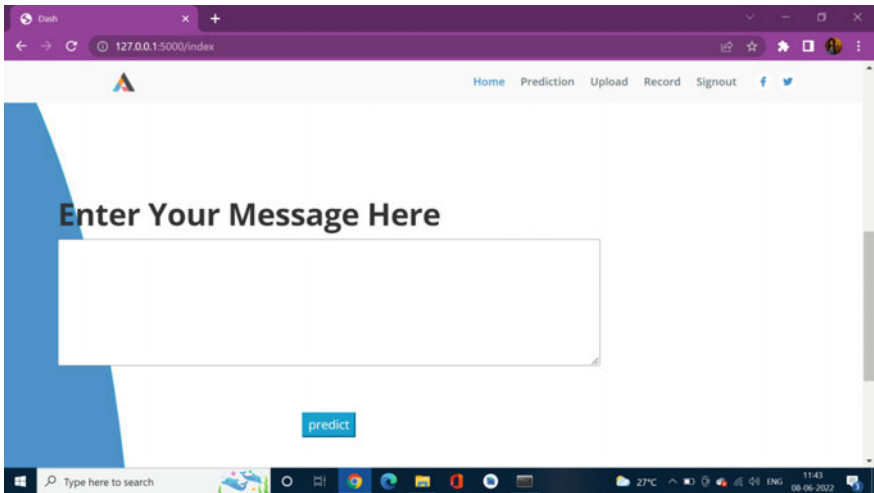


Fig. 20 Text box input

as to incorporate the concept of real-time recording and simultaneous detection of depression, as opposed to the current method of feeding speech files to the training model and then testing it to give the result. Finally, this research study intends to collect additional data on this disease in order to broaden the dataset and improve accuracy to make the interface more interactive with the user. This project's future scope could include incorporating real-time recordings and instant depression analysis done concurrently and provided to the user. The proposed system can also be

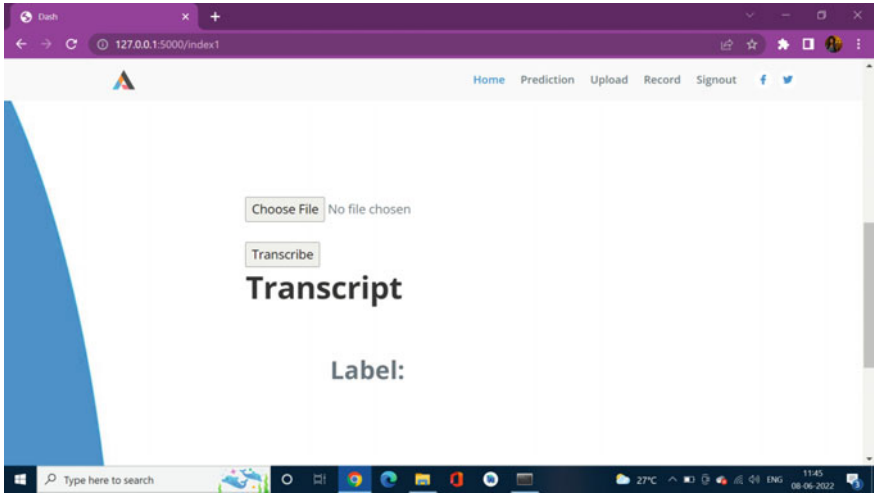


Fig. 21 Audio file input

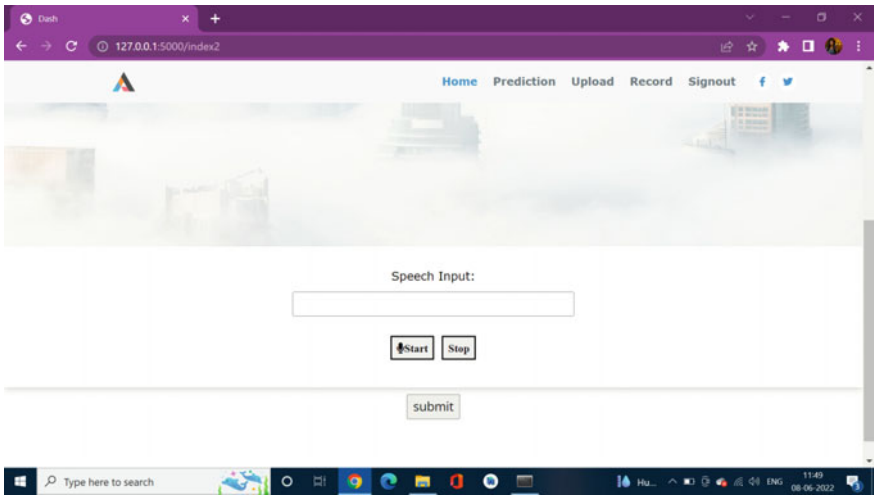


Fig. 22 Live recording input

improved to detect the depression directly by using the modulation of voice instead of using words as parameters. The scope for improvement and expansion is very vast for this project, and it can be implemented in various different ways.

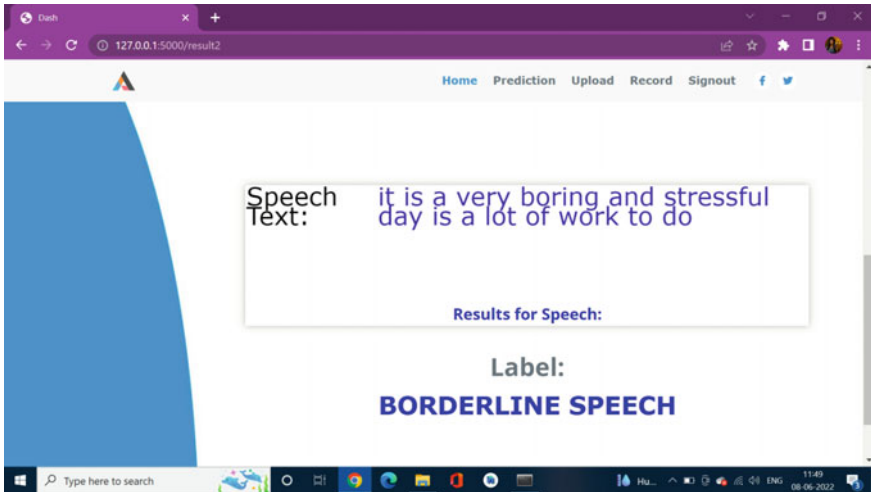


Fig. 23 Output

References

1. Ding Y, Chen X, Fu Q, Zhong S (2020) A depression recognition method for college students utilizing deep integrated support vector algorithm. *IEEE Access* 8:75616–75629. <https://doi.org/10.1109/ACCESS.2020.2987523>
2. Kiss G, Jenei AZ (2020) Investigation of the accuracy of depression prediction based on speech processing, 2020 43rd international conference on telecommunications and signal processing (TSP), pp 129–132. <https://doi.org/10.1109/TSP49548.2020.9163495>
3. Jenei AZ, Kiss G (2020) Possibilities of recognizing depression with convolutional networks applied in correlation structure. In: 2020 43rd international conference on telecommunications and signal processing (TSP), pp 101–104. <https://doi.org/10.1109/TSP49548.2020.9163547>
4. An H, Lu X, Shi D, Yuan J, Li R, Pan T (2019) Mental health detection from speech signal: a convolution neural networks approach. In: 2019 international joint conference on information, media and engineering (IJCIME), pp 436–439. <https://doi.org/10.1109/IJCIME49369.2019.00094>
5. Huang K, Wu C, Su M, Chou C (2017) Mood disorder identification using deep bottleneck features of elicited speech. In: 2017 Asia-Pacific signal and information processing association annual summit and conference (APSIPAASC), pp 1648–1652. <https://doi.org/10.1109/APSIPA.2017.8282296>
6. Huang Z, Epps J, Joachim D, Sethu V (2020) Natural language processing methods for acoustic and landmark event-based features in speech-based depression detection. *IEEE J Select Top Signal Proces* 14(2):435–448. <https://doi.org/10.1109/JSTSP.2019.2949419>
7. Song S, Shen L, Valstar M (2018) Human behaviour-based automatic depression analysis using hand-crafted statistics and deep learned spectral features. In: 2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018), pp 158–165. <https://doi.org/10.1109/FG.2018.00032>
8. Kiss G, Vicsi K (2017) Mono-and multi-lingual depression prediction based on speech processing. *Int J Speech Technol* 4:919–935
9. Han Z, Wang J (2019) Speech emotion recognition based on deep learning and kernel nonlinear PSVM. *Chin Control Decis Conf (CCDC) 2019*:1426–1430. <https://doi.org/10.1109/CCDC.2019.8832414>

10. Shih P, Chen C, Wang H (2017) Speech emotion recognition with skew-robust neural networks. In: 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 2751–2755. <https://doi.org/10.1109/ICASSP.2017.7952657>
11. Suganya S, Charles EYA (2019) Speech emotion recognition using deep learning on audio recordings. In: 19th international conference on advances in ICT for emerging regions (ICTer), pp 1–6. <https://doi.org/10.1109/ICTer48817.2019.9023737>
12. Jaya R, Thirukkumaran R, Saravanan S, Sivabalan N. Applying machine learning techniques to predict the maintainability. *J Phys Conf Ser Art Scopu*
13. Anida M, Premalatha K (2017) An application of Fuzzy Normalization in miRNA data for novel feature selection in cancer classification. *Biomed Res* 28(9):4187–4195 (Impact Factor:0.226, SJR:0.16)
14. Yang L, Jiang D, Xia X, Pei E et al. Multimodal measurement of depression using deep learning
15. Govindasamy KA, Palanichamy N (2021) Depression detection using machine learning techniques on twitter data. In: 2021 5th international conference on intelligent computing and control systems (ICICCS). <https://doi.org/10.1109/iciccs51141.2021.943>
16. Alghifari MF, Gunawan TS, Nordin MAW, Kartiwi M, Borhan L (2019) On the optimum speech segment length for depression detection. In: 2019 IEEE international conference on smart instrumentation, measurement and application (ICSIMA), pp 1–5. <https://doi.org/10.1109/ICSIMA47653.2019.9057319>
17. Dibeklioglu H, Hammal Z, Cohn JF (2017) Dynamic multimodal measurement of depression severity using deep autoencoding. *IEEE J Biomed Health Inform*
18. Abdel-Hamid O, Mohamed A-R, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional neural networks for speech recognition. *IEEE/ACM Trans Audio Speech Lang Proces* 22(10):1533–1545
19. Lahaie O, Lefebvre R, Gournay P (2017) Influence of audio bandwidth on speech emotion recognition by human subjects. *IEEE Glob Conf Signal Inform Proces (GlobalSIP)* 2017:61–65. <https://doi.org/10.1109/GlobalSIP.2017.8308604>
20. Lee KH, Kyun Choi H, Jang BT, Kim DH (2019) A study on speech emotion recognition using a deep neural network. In: 2019 international conference on information and communication technology convergence (ICTC), pp 1162–1165. <https://doi.org/10.1109/ICTC46691.2019.8939830>
21. Cummins N, Epps J, Ambikairajah E (2013) Spectro-temporal analysis of speech affected by depression and psychomotor retardation. In: 2013 IEEE international conference on acoustics, speech and signal processing, pp 7542–7546

A Survey of Different Identification and Classification Methods for Medicinal Plants



Shashank M. Kadiwal, Venkatesh Hegde, N. V. Shrivathsa, S. Gowrishankar, A. H. Srinivasa, and A. Veena

Abstract Pharmaceutical firms are increasingly resorting to therapeutic plants since they are less costly and have fewer negative effects than existing medications. Based on these findings, some people are concerned about automated medicinal plant recognition. There are several techniques to creating a reliable classification system that can categorize medicinal plants in real time. The study includes the usefulness and reliability of many image processing algorithms, machine learning, and deep learning algorithms for plant classification based on the leaf images used in the recent years with their benefits and drawbacks. The effectiveness of these algorithms in recognizing leaf images based on plant characteristics like shape, grain, texture, and combination of many aspects is evaluated. Our paper looks at both publicly available and self-captured leaf datasets for automated plant identification, and it concludes with a summary of current research and areas for improvement.

Keywords Medicinal plants · Deep learning · Machine learning · Convolutional neural network (CNN) · Support vector machine (SVM) · Computer vision · Multi-classification · Leaf patterns

S. M. Kadiwal (✉) · V. Hegde · N. V. Shrivathsa · S. Gowrishankar · A. H. Srinivasa · A. Veena
Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology,
Bengaluru, Karnataka 560056, India
e-mail: shashankkadiwal2000@gmail.com

V. Hegde
e-mail: venkateshhegde18@gmail.com

N. V. Shrivathsa
e-mail: vatsanv917@gmail.com

S. Gowrishankar
e-mail: gowrishankarnath@acm.org

A. H. Srinivasa
e-mail: srinivasaah.cs@drait.edu.in

A. Veena
e-mail: veenaal@acm.org

1 Introduction

Plants are necessary for all life on Earth because they supply food, oxygen, and medicines. To aid in the identification of new medicinal plant species in order to help the pharmaceutical business, the environment, and agricultural production and sustainability, a complete understanding of medical plants is essential. According to the Botanical Survey of India, India possesses 8000 medicinal plant species. India is the main supplier of medicinal herbs, with a global herbal trade of \$120 billion. Medicinal plants are considered to be fairly safe to use with few or no side effects. The fact that these treatments are in harmony with nature is a great advantage. Medicinal plants provide a rich source of chemicals that can be used in the production of pharmacopoeia, non-pharmacy, and synthetic drugs. Changes in leaf characteristics, for example, are used by botanists and agronomists in their plant study as a comparison tool. This is because leaf traits in deciduous, annual plants, and evergreen perennials may be observed and analyzed throughout the year. Despite several attempts (including powerful computer vision algorithms), medicinal plant identification remains a challenging and unsolved problem in computer vision. India is a biologically diverse country. Several articles have suggested potential solutions to the problem. Medicinal plant identification has been advocated for medication development, drug research, and the preservation of rare medicinal plant species. The first component of this survey investigation looks at the current methods used by agronomists to identify medicinal plants. In Sects. 2 and 3, we classified some of the papers related to medicinal plant recognition by the type of data utilized to identify the medicinal plants, namely structured and unstructured data. For identifying the medicinal plants using structured data approaches such as SVM, extreme learning machine, artificial neural networks, K-nearest neighbors, and random forests [1–6] have been experimented for recognizing the medicinal leaves. Convolutional neural networks with various different architecture such as MobileNet, AlexNet, VGG-16 [7–9], and many other approaches have been tried for identifying the medicinal plants using unstructured data such as images. In this paper, we look at all the prior strategies and see how effective they were in overcoming this obstacle. The survey article's last section summarizes what we gained from reading all of that research, as well as how the methodologies have evolved over time.

2 Different Approaches Based on Data Type Format

2.1 Structured Data

Kan et al. [1] propose an automatic identification method using images of leaves of medicinal plant. The leaf is shot in color at first (Fig. 1).

The images are sharpened and denoised after the petiole is removed. After grayscale conversion, images are binarized. Edges, texture properties, and shape

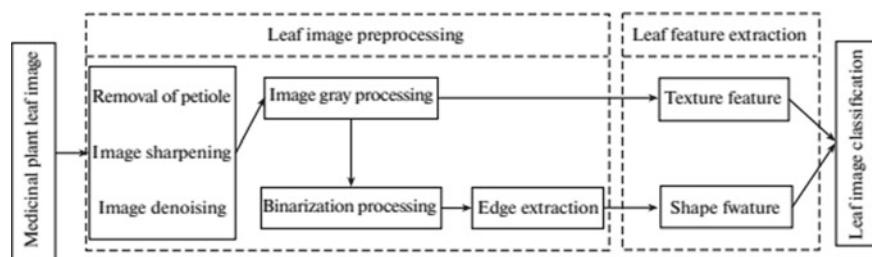


Fig. 1 Experiment flowchart (source classification of medicinal plant leaf image based on multi-feature extraction, p. 2)

features are all gathered. Threshold segmentation is used to separate the leaves from the background and get the contours of the leaves. Pixels with values greater than the threshold Q ($Q = 0.7$) are white, and the rest are black. From medicinal plant leaf images, the approach derives three form features (SF) as well as five texture characteristics (TF). Some of the form features are eccentricity, rectangleness, and circularity. Haralick et al. introduced GLCM as a well-known second-order statistical method for characterizing texture qualities of images. The main concept behind this approach is to create a matrix based on the probability that two pixels will meet a particular displacement relationship and the gray value of the image. The distance and angle between the two pixels define it. GLCM can be used to characterize the texture features of images of medicinal plant leaves. The five texture qualities are contrast, entropy, correlation, energy, and inverse difference moment. Invariant moments help to highlight an image's geographical differences. Moment invariant is a valuable indicator for defining picture aspects since it has traits like translation, rotation, and scale invariance. In this experiment, seven invariant moments are used to characterize the leaf shape characteristics of medicinal plants. The researchers utilized 240 leaf photographs from 12 different medicinal plants. After that, split each medicinal plant's 20 leaf photographs into two groups: 10 training and the rest 10 testing samples. Finally, SVM model was trained using 120 training examples before being utilized to categorize the 120 samples. Repeat the second and third steps several times to get the average recognition rate and classification result. BP network classifiers, PNN, and KNN classifiers were also used in the comparative experiments. The average recognition rate of all the four classifiers exceeds 70% when using only 10 SF functions, and the recognition rate of PNN is the highest at 78.3%. For all classifiers, the average identification rate surpasses 50% when just 5 texture characteristics (TF) are utilized, with the SVM classifier with the highest detection rate of 58.3%. When both texture and shape data (TF + SF) are used, the average detection rate of the four classifiers has improved significantly, and the SVM classifier has achieved a detection rate of 93.3%. This is a higher value than the other three classifiers. The findings show medicinal plants may be automatically categorized by

combining multi-feature analysis of images of the leaves with SVM. This publication gives a valuable conceptual framework for the research and the development of classification models of medicinal plants.

The Flavia leaf dataset (which comprises 1907 photos of leaf belongs to 32 distinct plants) is utilized in Muammer Turkoglu and Davut Hanbay's research article [2], and the authors offer a method for extracting attributes using segmentation rather than collecting data from the full leaf. The overall image's features are then established by merging the information obtained from each piece. Examples include vein-based features, GLCM, color, and Fourier descriptors. These methods help to determine the characteristics of each leaf. The ELM approach was used to classify and test these feature parameters. Leaf pictures were preprocessed by removing the backdrop and converting RGB photographs to gray scale. To segment leaf pictures, a trial-and-error method of identifying suitable thresholding settings was applied. Erosion and filling processes were also used to decrease potential pixel blockages. The image data are normalized once the characteristics have been retrieved. ELM hidden layer input weights and biases start randomly and, unlike traditional networks, are stable throughout the process. The determination of output weights is done by least squares method. Seven leaf photographs were created from a 1907 leaf sample after being randomly selected prior to each intersection using a ten-fold cross-validation model. With this strategy, an average success rate of 98.36% was reached. A comparison between ANN and LS-SVM was done. The suggested approach minimized the influence of deformations on the identification and classification performance of certain plant leaves. Unlike earlier studies, the presented strategy solves problems and can be improved by adding more feature extraction methods.

In the experiment done by Yigit et al. [3], ANN, NB algorithm, random forest algorithm, KNN, and SVM were used for autonomous identification of leaves. They applied 637 leaves from 32 distinct plants were applied in this investigation. Image processing methods are used to extract the visual features of every leaf. Texture, pattern, dimension, and color are the four groupings of 22 VF. To explore the impact of these clusters on the classification recital, 15 alternative combinations of four groups were created. After that, the models are trained using data from 510 plant leaves, and their accurateness is calculated using data from 127 leaves. SVM model is the most accurate identifier for all-category combinations, according to the testing, with an accuracy of 92.91%. With an accurateness of 87.40%, a combination of D#6, C#6, and P#5 groups comes in second. Because employing the fewest number of VF is the most significant component of the classification process, the CFS method is utilized for selection of the 16 most effective VF for identification. The SVM model also delivers the greatest results for these 16 VF, with an accuracy of 94.49%. The efficiency of the proposed method is then put to the test in order to identify sick and defective leaves. As a result, 637 healthy leaves have been combined with 33 damaged or infected leaves. The SVM model recognized 536 randomly selected leaves for training and 134 leaves for testing with an accuracy of 92.53%, accounting for 80% of all leaves. The P# is the 5-feature group is demonstrated to be the most effective feature group quantitatively in this investigation. Furthermore, the edge Fourier transform feature has been shown to be the effective attribute in P#5 groups.

The findings imply that, if AIT replicas were built and skilled appropriately, they may reliably identify plants even when samples are ill or damaged.

In their paper, Azadnia and Kheiralipour [4] offer an ANN for categorizing medicinal plants. Forty sites were chosen for the investigation, with four plant specimens obtained from each. Six collections of medicinal plants were collected and numbered and characterized from A1 to A6. The photos were preprocessed before the features were extracted. Preprocessing includes picture segmentation, which divides the sample photos into various segments and treats just the desirable portions of the images. Certain texture, shape, and color information were recovered following the classification of significant attributes. Before extracting the features, the established image processing approach created different color spaces for the photographs, such as I1I2I3, HIS, CrCb, and NRNGNB. To discover crucial color-based features, the skewness, kurtosis, mean, variance, and standard deviation metrics were extracted from the investigated medicinal plant photographs. Homogeneity, entropy, energy, and correlation are all texture-related metrics that have been quantified. After acquiring the GLCM of the pictures, the texture properties were assessed. Following feature extraction, the best features were picked for classification. The medicinal plants under investigation were identified using effective qualities gleaned from sample images. ANN is utilized for the categorization of the many types of medicinal plant samples. To analyze the ANN models, the obtained characteristics were classified into three parts: training, testing, and validation sets, each including 60, 20, and 20% of the dataset. The ANN models were assessed using some statistical metrics, such as the correlation coefficient of test data (r), CCR, and MSE. The results showed that the algorithm can properly categorize different types of medicinal plants. In agricultural commodities, texture, shape, and color extraction algorithms have been widely and successfully used for several applications, such as classification, identification, and so on. The design and implementation of robustness algorithms for color, texture, and shape extraction was a major contribution of the study. As a result, the integration of the suggested machine vision system has a lot of potential for categorizing and identifying various agricultural goods based on valuable texture, form, and color features.

In their paper, Pacifico et al. [5] report a mechanized characterization of restorative plant species in view of variety and surface elements. The first step was to collect images of a variety of medicinal plant species, which were subsequently processed. The first step in preprocessing was to eliminate the backdrop. Once the backdrop is removed, the RGB shots are converted to grayscale images. After that, the grayscale photographs are converted to binary images. After the image has been preprocessed, many color properties are extracted/calculated, including mean, standard deviation, skewness, and kurtosis. The gray-level co-occurrence matrix is used to compute textural properties for each leaf image segment (GLCM). The GLCM counts how many times certain gray-level combinations appear in a photograph. Weighted K-nearest neighbors (WKNNs), random forest (RFC), and a multilayer perceptron trained with backpropagation technique were used to classify medicinal plant species (MLP-BP). The trials employed a ten-fold cross-validation method. The scientists discovered that, in contrast to the other literature-based classifiers,

RFC and MLP-BP can get the highest scores in all four-classification metrics, with MLP-BP doing marginally well than RFC. Furthermore, all KNN and WKNN slants were surpassed by the DT classifier. The authors also observed that when the k value increases, both KNN and WKNN perform worse. Both RFC and MLP-BP get the best results in all four stated classification metrics when compared to the other carefully chosen classifiers from the prose, with MLP-BP somewhat outperforming RFC. Furthermore, all KNN and WKNN slants were surpassed by the DT classifier. The authors observed that, when the k value increased, the performance of both KNN and WKNN dropped. Subsequent to leading a general assessment that considered both grouping measurements and normal execution time, the creators inferred that the arbitrary timberland classifier would be the most ideal decision among the chose classifiers for making a programmed order framework for the proposed informational collection.

Sharma et al. [6] study's major goal was to use image processing techniques to identify the leaves. The writers acquired the photographs first, then utilized the equation to convert them to grayscale images.

$$C_{\text{gray}} = 0.2989 * C_{\text{red}} + 0.8570 * C_{\text{green}} + 0.1140 * C_{\text{blue}} \quad (1)$$

where C is the channel, red, green, and blue address the different variety channels of the picture.

The distinctive properties of leaves help to identify them. The histogram of oriented gradients (HOGs) approach was used for this. The essential thought behind hoard is that the light power of privately circulated slopes is the most effective way to depict an article's shape and structure. The picture might be separated into cells that show parcels that are equally compared and isolated. These areas can be separated into wedges, which are then normalized to preserve consistency in the face of illumination or photometric changes. It is calculated to normalize extracted properties across blocks. The 900 criteria calculated include length, width, aspect ratio, perimeter, form factor, circularity, and compactness. K-nearest neighbors and artificial neural network backward propagation algorithms are the two supervised learning methodologies categorized in the final stage. K-nearest neighbors have a 92.3% accuracy rate. When the two methodologies were compared, artificial neural networks were found to be the better option, with a 97% accuracy.

2.2 Unstructured Data

In the article proposed by Varghese et al. [10], the system was created to detect the plants using transfer learning. The six medicinal plant groupings considered in the study were *Coleus Aromaticus*, *Annona Muricata*, *Hemigraphis Colorata*, *Aloe Barbadosensis* Miller (*Aloe vera*), and *Azadirachta Indica* (neem). Each class had about 500 photographs. Once the features have been obtained, the model is trained using the MobileNet architecture. The component extraction method is done utilizing the

convolutional brain organization, or CNN. Extra convolutions, for example, pooling layers, completely associated layers, and normalizing layers, which are alluded to as covered up layers since their feedback and result sources are hidden by the actuation work and the last convolution, are often trailed by a ReLU layer. Back engendering is as often as possible used in the last convolution in this method, bringing about a rising amount of exact weight. The input image is processed over several kernels in the convolution layer, apiece with its unique function. Color identification is handled by a few filters, while edge separation is handled by others. Likewise, each filter has a certain role. Lines, corners, and other essential parts will be discovered in early stages, and a full edge, for example, will be discovered in the final layers. The item's shape and other features will be revealed in the final layers. More features will be obtained as the number of convolutional layers increases, and these data will be arranged into feature maps. After each convolutional layer, the pooling stage is performed. For instance, a $2 * 2$ -pixel channel will slide multiple pixels in a steady progression, with the greatest impetus coming from those four pixels being saved in one pixel of the accompanying layer. This allows for the reduction of dimensions and a more complete examination of key traits. A totally associated layer has each neuron connected to each and every other neuron, and the result is gotten from the last layer. The preparation model is then changed over completely to TensorFlow light viable organization using a TensorFlow light converter. The plant name was the output of the label with the highest probability. The model was trained in two stages, each with 12 epochs each. The research work completed by these individuals was 95% accurate, which is comparable to human performance on the dataset.

Pechebovicz et al. [7] suggest a smartphone app to help universities, students who have never encountered the species, and healthcare professionals recognize Brazilian medicinal plants. They looked at a collection of 10,162 downloadable photographs that encompassed different stages of plant life, visual backdrops, and ecosystems. The Google Photos dataset was created using a download service that allows users to download photographs based on keywords or key phrases. They describe how the database was created using the Brazilian Service of Wellbeing's rundown of therapeutic and normal, unsafe plants. Data augmentation methods such as rotation, cropping, mirroring, random zooms, random distributions, and skewing were developed using Python's augmentor module. After the augmentation, the absolute number of photographs was augmented to 151,128 through 5230 in the hazardous group and a characteristic of 2000 in the other classes. In this work, the authors employed the Mobilenetv2 architecture, using beginning weights learnt from the ImageNet dataset with values between 1 and 224. The model has an aggregate of 2,629,032 params, with 371,048 teachable and 2,257,984 non-teachable. Simply by preparing the top layers of the Mobilenetv2 with the new classifier, the model was tweaked to increment execution. After the layers 101 through 155 were relaxed and the model was reinstructed, the model was retrained. This stage has 2,233,640 teachable and 395,392 non-teachable params. Several epoch sizes (5, 10, 20, 30) were used in both the training and fine-tuning phases, and a comparison was made. After 20 epochs of training, the final output is fine-tuned. The absolute approval exactness

was 0.4254, though the main 5 precision was 0.685. In the implanted model in a cell phone, the loads from the last preparation and adjusting step are utilized.

Using deep features to characterize the original plant leaf image, Prasad and Singh [8] constructed a data interchanged from thing identification to plant hereditary research. The authors suggested utilizing VGG-16 to construct features, which were then condensed using PCA to appropriately display medicinal plant leaves for categorization (Fig. 2). During preprocessing, the leaf is photographed against a dark background using the recommended unique approach for obtaining leaf samples quickly and clearly. To create classification feature maps, the image was transformed to 1 color space and voted for through fully allied convolutional layers. The book also goes into how to capture good photographs. The photo capture requires a basic transparent triangular glass structure set aside over a black backdrop with LED illuminations to prevent the dispersion consequence. The camera head is situated opposite to the board with L distance to best catch the leaf wedged between the straightforward sheets. The distance between the camera and the leaf is L, and the point between the camera and the item plan is. To make it reliable, this paper utilizes $= 90^\circ$ and $L = 1$ m.

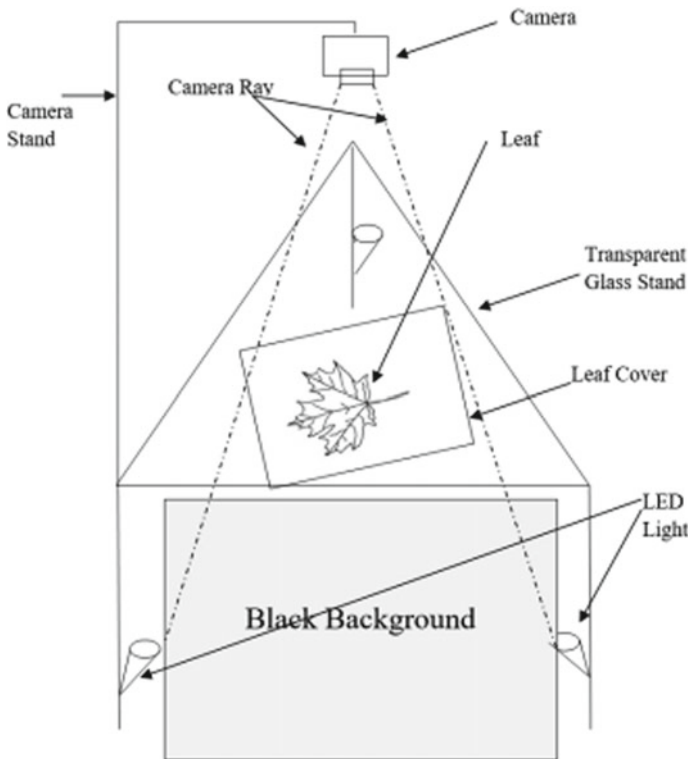


Fig. 2 Proposed leaf capture setup block diagram (source medicinal plant leaf information extraction using deep features, p. 2)

The image is in the RGB color space. This is a device-dependent color space that needs to be converted to a device-independent CIE-1 color space using equations.

$$I \xrightarrow{xyz} I_{xyz} \xrightarrow{l\alpha\beta} I_{l\alpha\beta} \tag{2}$$

where I is the caught picture,

l is the gentility coefficient going from 0 (compares to unadulterated dark) to 100.

For any deliberate shade of power l , the directions (α, β) on a rectangular-coordinate framework opposite to the l hub at l find the variety credits (tint and immersion).

Since there is no immediate change among RGB and $l\alpha\beta$ and we initially change RGB to XYZ variety space utilizing the condition:

$$\begin{aligned} x &= R \times 0.4124 + G \times 0.3576 + B \times 0.1805 \\ y &= R \times 0.2126 + G \times 0.7152 + B \times 0.0722 \\ z &= R \times 0.0183 + G \times 0.1192 + B \times 0.9505 \end{aligned} \tag{3}$$

where R, G, B are various channels of the picture.

$$\begin{aligned} l &= 116 * Y' - 16 \\ \alpha &= 500 * X' - Y' \\ \beta &= 200 * Y' - Z' \end{aligned} \tag{4}$$

The gentility coefficient, l , goes from 0 (complete dark) to 100 (complete white) (analyzes to pure white). The variety qualities for each assessed shade of solidarity l are found utilizing the directions of focuses on a rectangular direction network that are opposite to the l hub at area l (shade and immersion). At the grid origin $(= 0, 0)$, the color is achromatic, meaning that it does not show any color. The grayscale image has a range of colors, from black (the darkest color) to white (the lightest color), with positive suggesting reddish-purple hue, negative representing bluish-green hue, positive indicating yellow, and negative indicating blue. We must first convert RGB to XYZ color space because there is no direct conversion from RGB to l . After the picture has been transformed to a device-independent lab color space, the VGG-16 feature map was developed. This element map is re-projected to PCA subspace to work on the adequacy of species recognizable proof. To illustrate the sturdiness, the researchers used two sorts of plant leaf collections. The tests employed two distinct kinds of leaf datasets. The first is the straightforwardly open ICL leaf benchmark dataset. The learning rate is 0.0015. The results are derived from their articles and contrasted with our techniques. The authors' accuracy percentage was 94.1%. The researchers devised two methods for recognizing plant leaves: l -VGG-16 and l -VGG-16+PCA, both of which were 94% accurate. This paper presents a better profound organization design for mechanized plant leaf species acknowledgment. A VGG-16 design in PCA space performs better using l variety space as an information. This study presents a

novel 2+ layered catching instrument that takes into consideration the exact recording of shape and surface data with negligible leaf folds. Leaf data are separated using a multi-scaled profound organization, bringing about a rich component vector. At long last, the vector is decreased using PCA to lessen the grouping cost.

Naresh and Nagendraswamy [11] exhibited a representative way for perceiving plant leaves. Modified local binary patterns (MLBPs) have been proposed for retrieving textural information from plant leaves. Plant leaves of comparable species can have an assortment of surfaces concurring on their age, obtaining, and environment. Thus, the bunching idea is utilized to pick countless class delegates, and intra-group changes are recorded using the stretch esteemed type emblematic qualities. The mean and standard deviation of a 3×3 framework of pixels inside the picture are utilized as neighborhood limit boundaries. After that the whole picture is handled to make a changed LBP histogram, which is then used as a surface descriptor to describe the picture. The chi-square distance is used to compare unknown leaf samples of a species with the reference leaf provided in the knowledge base. The plant leaves’ textural qualities are retrieved first, following by emblematic portrayal. From that point onward, the matching strategy is utilized to sort the submitted test. To make categorization easier, a basic nearest neighbor classifier is utilized.

$$LBP_{P,R} = \sum_{i=1}^P S * 2^{i-1} \tag{5}$$

where $S = \begin{cases} 1 & \text{if } (Gi - Gc) \geq 0 \\ 0 & \text{otherwise} \end{cases}$

P is the similarly divided pixels with sweep R from the middle pixel and where Gc and Gi are the dim upsides of the middle pixel and neighborhood pixel separately.

The essential LBP depends on a hard limit laid out by the dim worth distinction between the middle and adjoining pixels. Rather than using a hard threshold, the authors developed a strategy that takes into consideration the closeness of adjoining pixels to the mean (μ) and standard deviation (σ) of the whole area. The dark upsides of pixels in the quick area influence the upsides of the image, which is versatile in nature and addresses the underlying connection between the dim upsides of pixels in the prompt area. Let Gi be the dim worth of the encompassing pixels around the middle pixel Gc in a geography of a circle with range R and P number neighbors. The mean and the standard deviation of the dim upsides of the relative multitude of pixels are figured as follows:

$$\mu = \frac{\sum_{i=1}^P Gi + Gc}{(P + 1)} \tag{6}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^P (Gi - \mu)^2 + (Gc - \mu)^2}{(P + 1)}} \tag{7}$$

where μ is the mean and σ is the standard deviation of the entire area.

The creators use progressive bunching to pick various class agents for each class by gathering comparable leaf tests in view of textural designs. The similarity score is calculated by comparing the test leaf sample to all of the knowledge base reference leaf samples. The chi-square distance was utilized to work out the disparity between reference leaf tests communicated as span esteemed type emblematic component vectors and a test leaf test with fresh element vectors. The researchers utilized a standard 1-closest neighbor arrangement technique to classify the given test leaf test as one of the known plant species. A range of datasets, including the UoM medical plant dataset, Flavia dataset, Foliage dataset, and Swedish leaf dataset, were used to evaluate the given approach. Overall exactitude for the suggested method was 98%. When compared to traditional crisp representation tactics, the nominated emblematic delegacy approach has been demonstrated to be effective in terms of decreasing the quantity of reference leaf tests expected to prepare the framework while holding grouping exactness. Other well-known classification approaches have been surpassed by the proposed classification system. Furthermore, the proposed method provides equal computational and storage efficiency when compared to the current systems.

Tan et al. [9] proposed D-Leaf, another CNN-based approach. To pre-process the leaf pictures and concentrate the characteristics, three unique convolutional brain organization (CNN) calculations were utilized: pre-prepared AlexNet, calibrated AlexNet, and D-Leaf. The support vector machine (SVM), artificial neural networks (ANNs), k-nearest neighbor (k-NN), Naive Bayes (NB), and CNN were the five AI techniques utilized to distinguish these properties. The leaf tests for this examination come from three unique areas at the College of Malaya in Kuala Lumpur, Malaysia. A total of 43 tropical tree species were collected, with each piece containing 30 samples. The leaf pictures were reproduced into square aspects, which our CNN models required as inputs. All of the pictures were first converted to gray scale from RGB. Sobel was used to isolate the area of interest from the photos. After division, the pictures were present handled and skeletonized to acquire a reasonable vascular engineering. CNN was used to retrieve the leaf's properties. The authors produced a model termed D-Leaf, which was compared to pre-trained AlexNet, fine-tuned AlexNet, and D-Leaf. The recommended model was built to extract characteristics from photographs rather than fine-tuning the AlexNet model (Fig. 3). This model contains three levels that are all connected. The 1290 neurons of the primary completely associated layer (FC4) and the second completely associated layer (FC2) accepted C3's result (FC5). The third completely associated layer (FC6), which had 43 neurons, mirrored the quantity of classes in the dataset. The D-Leaf model was prepared utilizing the stochastic inclination drop with energy approach with a clump size of 100. A preparation subset called mini-batch changes the weights and evaluates the loss function gradient, number of finishing focuses, number of branches, number of areoles, and different factors. The leaf region was utilized to process the thickness of veins, fanning destinations, end focuses, and areoles. The D-FC5 Leaf's layer qualities were then arranged utilizing five (5) distinct classifiers: SVM, ANN, k-NN, NB, and CNN. The D-Leaf with SVM had a precision of 82.75%, ANN

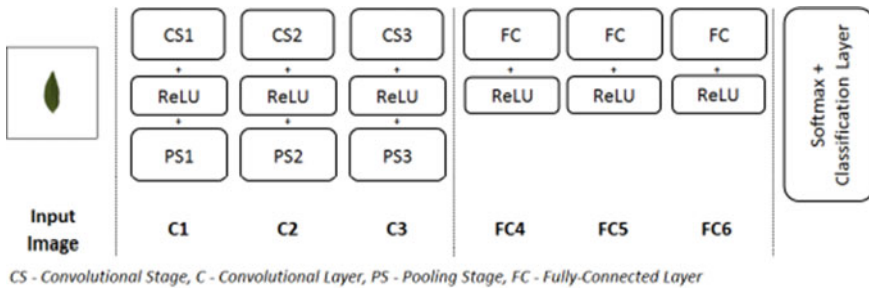


Fig. 3 Architecture of AlexNet (*source* deep learning for plant species classification using leaf vein morphometric, p. 6)

94.88%, KNN 82.44%, NB 81.86%, and CNN 79.03%. Cross-validation (CV) was used to forestall overfitting troubles on the grounds that the suggested dataset just included 30 examples for every species. The cross-validation process produces numerous miniature train-test parts utilizing the underlying preparation information. The information is divided into k subsets, known as folds, in regular k -overlap cross-validation. The technique is then prepared over and over on $k-1$ folds with the excess crease filling in as a test set. With cross-validation, the model is calibrated with various hyperparameters utilizing just your unique preparation information. This decreases overfitting by keeping your test set as a totally obscure dataset for picking your last model. The 5-fold and 10-fold CV were investigated utilizing the D-Leaf strategy. The accuracy of fivefold cross-validation was 93.15%, and tenfold cross-validation was 93.31%, respectively. Aside from the dataset, D-Leaf was evaluated on the Malaya Kew, Flavia, and Swedish datasets, with accuracy of 90.38, 94.63, and 98.09%, respectively. The scientists found that consolidating CNN with ANN which included feed forward brain network with a solitary secret layer which comprised of 80 neurons that utilized the default scaled conjugate gradient work and the accomplishment of a base inclination as the halting models produced the best outcomes.

In the Western Ghats of India, Sabu et al. [12] nominated a PC vision technique for comprehending ayurvedic restorative plant species. Picture catch, picture pre-handling, highlight extraction, and order are the four stages of the proposed framework. A 14.1 megapixel camera was used to take leaf photos. The picture maximum area boundary is then examined for preprocessing, which may correlate to the leaf area. This region is clipped off in step two of the preprocessing and captured as leaf pictures, which are then transformed to binary and morphological procedures are conducted. After that, all of the cropped photos were resized to the same height of 1200 pixels. The authors evaluated two techniques for feature extraction: accelerate vigorous element and histogram of arranged slopes. An assortment of interest focuses is first found from the picture utilizing the SURF highlight descriptor, then the strongest 20 are chosen, resulting in a SURF feature vector of $20 * 64$ dimensions. The dimensionality of HoG features varies with picture resolution since every point in an image has an oriented gradient. The authors limited the HoG feature vector to a

maximum of 100,000 values. The element vectors are then joined to make a solitary 101,280-pixel standardized highlight vector. SURF qualities are extricated using the indispensable picture idea, which is resolved using the accompanying condition.

$$S(x,y) = \sum_{i=0}^x \sum_{j=0}^y I(x, y) \quad (8)$$

where S addresses the SURF highlights x , y addresses the directions of the picture and I addresses the picture.

In computer vision and picture handling, the histogram of arranged inclinations is a component descriptor for object location. This approach counts how often an inclination direction shows up in a specific segment of a photo. To increment precision, hoard is figured across a thick matrix of consistently separated cells using a covering neighborhood contrast standardization procedure. The dataset was then parted fifty, with one half used to prepare and the other half used to test the KNN classifier. The melded highlight vectors and relating class names were processed using the KNN classifier, and the results appear to be sufficient for constructing real-world apps. The presentation of the proposed framework was assessed using measurable techniques, including the computation of the metrics precision, recall, F-measure, and accuracy. When tested with a KNN classifier, the proposed technique achieves near-perfect accuracy by combining SURF and HOG features.

Dileep and Pournami [13] worked effectively of distinguishing therapeutic plants in light of leaf elements like shape, size, variety, and surface. The proposed dataset incorporates leaf tests from 40 therapeutic plants. Ayurleaf, the proposed own dataset, has 2400 pictures of in excess of 30 leaves from 40 different plant species. A comparative naming show is utilized to check each picture: plant species name followed by an interesting grouping number. Because specimen guarantees that the dataset is varied in terms of plant species, the algorithm may produce more accurate classification results. Using the GIMP image editor, just the leaf region is selected and cropped, and each image is saved in the jpg (Joint Photographic Experts Group) format, with the naming method used to label the images. They used RGB pictures, and if the photos weren't in the $227 \times 227 \times 3$ format, they padded the image to make it $N \times N$ dimensions, then scaled the padded image to the $227 \times 227 \times 3$ dimension. The Ayurleaf CNN framework is created and constructed using the AlexNet. The underlying layer, the information layer, decides the size of the info pictures. In the second layer of the convolution layer, 90 (7×7) channels with a step size of 2 were utilized. A ReLU layer limits the result, which is then trailed by a maximum pooling layer with a 2×2 channel size. This layer lessens the result size to around 50% of its unique size. After the maximum pooling layer, a subsequent convolution layer chips away at 512 pieces with a size of 5×5 and a step of 2. From that point forward, there is a ReLU layer at that point, a maximum pooling layer with a 3×3 channel size and step 1. The following two layers are consecutive convolution layers with the accompanying arrangements. The first and second use 3×3 portions with a step of 2, with 480 and 512 bits, separately. A ReLU layer and

a maximum pooling layer with channel size 2×2 and step 2 are added after these layers. The result from the last max-pooling layer is gotten by the main completely associated layer, which involves 6144 neurons. The second altogether coupled layer comprised of 6144 neurons. The third completely associated layer has 40 neurons, which compares to the quantity of restorative plant groupings to be characterized. Lastly, the production of the third fully connected layer, which generates classification probabilities for each species, is sent to the SoftMax classification layer. The recommended CNN model performs four tasks: picture obtaining, picture preprocessing, include extraction, and arrangement. Picture capture using a flatbed scanner requires sampling the photos and guaranteeing image quality. For feature extraction, variable CNN models with varied numbers of layers, filters, and training parameters were created. The Ayurleaf dataset was used to compare the results of AlexNet, fine-tuned AlexNet, and the D-Leaf CNN model. On a training graph, they presented their final findings. The accuracy and loss functions have evolved through time, as seen in the graph. According to the study, the recommended technique AyurLeaf based on AlexNet has an average accuracy of 96% on the validation set.

In this review, Shah et al. [14] utilize a double-way profound convolutional brain organization (CNN) to learn shape-subordinate highlights and surface qualities. These paths eventually join to form a multilayer perceptron, which integrates balancing info on silhouette and quality to categorize leaf images more effectively. The leaf picture and the surface fix are made by the creators and contribution to the CNN model's two ways. The texture-patch input is made by increasing the original leaf photo by $2\times$, sharpening, and cutting the core section of the sprig to produce a 144,192-pixel reinforcement that substantially captures elements relevant to leaf consistency and venation. The CNN's element extraction part is comprised of convolutional and max-pooling layers. Each convolutional layer is trailed by a group standardization (BN) layer. With the exception of the last layer, the corrected straight unit is the enactment utilized in all stages (ReLU). On the way to create the leaf's joint shape-texture representation, the components gathered from both pathways are integrated. This composite representation is sent into a multilayer perceptron for classification. The images were improved to increase accuracy while avoiding overfitting. On datasets including Flavia, Leafsnap, and ImageClef, the model performed well, with accuracy of 99.28, 95.61, and 96.42%, respectively. They have also introduced the reduced shape context, a clever shape include that is computationally proficient and outperforms standard hand-crafted features, making it superior to most traditional approaches. As leaf classification techniques, they looked at dual-path CNN, uni-path CNN, texture-patch CNN, marginalized shape context with an SVM classifier, curvature histogram, and multiscale distance matrix. The proposed technique outperforms all others with an average accuracy of 97%.

Samuel Manoharan [15] recommends a two-stage framework for recognizing the restorative plants or spices. The initial step of the stage one includes edge-based natural plant recognition and order-based restorative plant discovery in the subsequent advance (Fig. 4). The creators have utilized different edge location strategies like Prewitt, Canny, Laplace, and Sobel edge detection activity for distinguishing the

edge in the primary stage which thus helps in the quicker ID. Then, an information-based regulator has been picked the edge discovery administrator in view of the estimation, which can work on the exactness of the identification cycle. In the subsequent stage, input crude pictures are handled at first to further develop the picture lucidity. When the picture is preprocessed, the variety, shape, length, and width of the photos are utilized to separate the highlights. The chose highlights are removed utilizing chi-square strategy to work on the order. The CNN classifiers are then utilized for arrangement of information extricated pictures by leaf. The shape data are gone through the visual cortex by the form of the picture, and afterward different elements like shape, color, and measurement of passes on shipped off tertiary visual cortex. Later, CNN is utilized with numerous convolution channels to handle elements to acquire suitable component vectors for wonderful characterization. The factorization takes different boundaries to work on the viability of the model. The inception structure in the pooling layers is utilized to recognize the home grown plant leaf. The consequences of the Phase 1 and Phase 2 are contrasted and XOR door activity. The total distinguishing proof is reported once the indistinguishable outcome is gotten in both phase of ID, i.e., which follows the XOR door activity. In this review, 250 leaf tests with the front and back of the picture were utilized with 80% of the information being preparing information and 20% of the information being utilized as the test information. The different gatherings of the highlights of natural plant leaf structures are contrasted and given input plant leaves. The recommended approach considered by the creators doesn't represent the way that all leaves have a similar perspective proportion, centroid, and roundness. The writers additionally look at the different edge location procedures and recommend that Sobel operator is more exact when the length of the article (leaf) is thought of while canny edge operator performed well when broadness of the item (leaf) was thought of. The above concentrate on brought about precision of 92% outpassing the single classifiers, for example, CNN, canny edge detector, and image segmentation technique subsequently performing better compared to other proposed approaches.

Comparison of different models for classification of medicinal plants is shown in Table 1.

3 Conclusion

Identifying medicinal plants has a wide scope of moves that actually should be tended to. In this paper, we have distinguished a thorough rundown of improvements that can be made in recognizing the medicinal plants using deep learning methodologies. We have also discussed some popular methods implemented in overcoming these challenges in part or as a whole. The impact of identifying medicinal plants using technologies like deep learning will be big for agronomists, farmers, doctors, and the pharmaceutical industry. The applications once developed for identifying the medicinal plants are going to give the agronomists, pharma industry, and ayurvedic practitioners a fabulous chance to revolutionize the way the medicine industry works

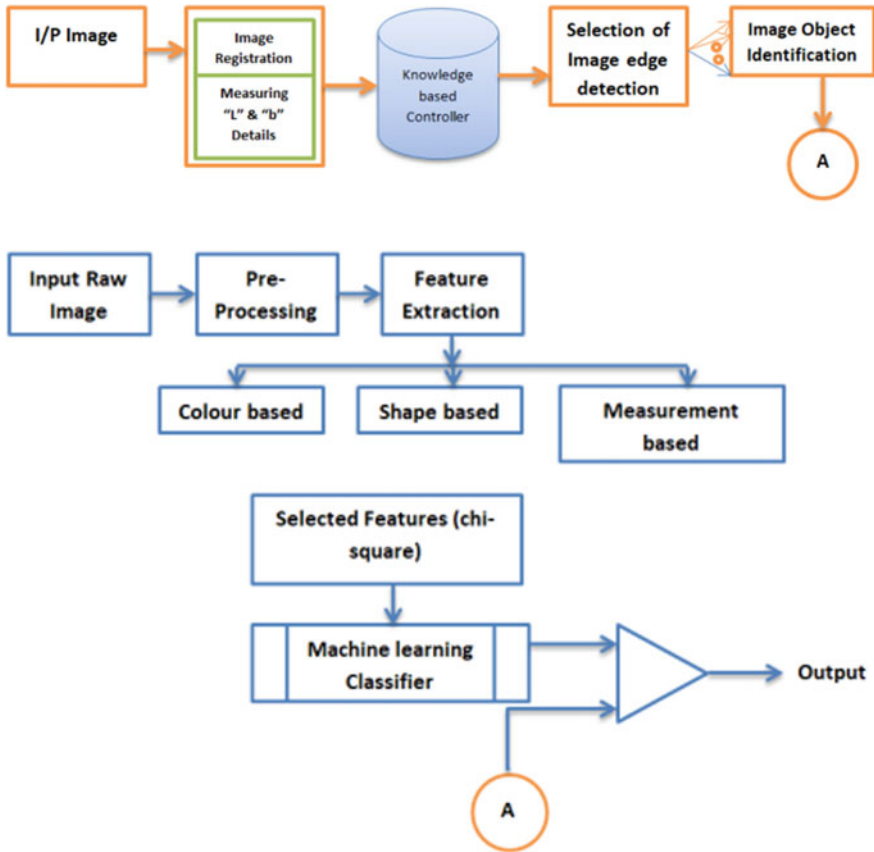


Fig. 4 Phases 1 and 2 architecture (source flawless detection of herbal plant leaf by machine learning classifier through two stage authentication procedure, p. 131)

currently. Every one of the papers examined above have tracked down their own interesting approach to moving toward the issue, utilizing various kinds of information or learning models. As an outline, these techniques demonstrate that there are new potential ways of fostering a mechanized framework to distinguish therapeutic plants. These elements add to our objective recorded as a hard copy this review study, which is to reveal new responses and a hole on the lookout for per users to load up with their inventive and unique answers for this overall issue. According to our viewpoint of examining the exploration hole, we can obviously see that each of the previously mentioned approaches have drawbacks; however, the longing to find the better strategies is consistently the main impetus for us, and this study paper is doubtlessly a supportive device.

Table 1 Comparison of different models for classification of medicinal plants

Reference No.	Model	Advantage	Disadvantage	Performance metrics
[1]	Support vector machine, BP neural net, PNN, KNN	Uses both shape and texture features	It cannot achieve a reasonably good recognition rate in light of just leaf shape or surface elements only	Accuracy—93.3%
[10]	Convolutional neural network	Predicted with a low-level computational requirements	–	Accuracy—95%
[7]	Convolutional Neural Network (Mobilenet v2)	Can easily run-on edge devices	Requires more dataset variety	Accuracy—94%
[8]	Convolutional neural network (VGG-16)	Proposes a method to capture leaf images in an efficient manner for data collection	–	Accuracy—94.1%
[11]	Local binary patterns	Better computational efficiency	There are more intra-class variations than between-class variations	Accuracy—95% (average, precision, recall)
[9]	SVM, ANN, KNN, NB, CNN	Comparative study of different algorithms	Not useful for compound leaves	Accuracy—94.8%
[12]	KNN	Lighter in computation and can be used in edge devices	Training time is high	Precision, recall, F-measure and accuracy—99.6
[13]	Convolutional neural network	Higher accuracy, computationally efficient	–	Accuracy—96.76%
[2]	Extreme learning machine	Implementation is easier with faster training	Performance is not up to the mark	Accuracy—99.10
[14]	Convolutional neural network	A better approach to gain higher accuracy with less computational requirements	–	Accuracy—97.10 (average)
[4]	Artificial neural networks	Novel, simple, robust, low cost	–	Accuracy—100%

(continued)

Table 1 (continued)

Reference No.	Model	Advantage	Disadvantage	Performance metrics
[15]	Edge detection algorithms + CNN	The edge-based detection provides better accuracy than other existing methods	The computation time is very large, and also, the pixel-based process consumes more storage space in the array of the algorithm used for pixel	Accuracy—92%

References

1. Kan HX, Jin L, Zhou FL (2017) Classification of medicinal plant leaf image based on multi-feature extraction. *Proc J Pattern Recogn Image Anal* 27(3):581–587
2. Turkoglu M, Hanbay D (2019) Recognition of plant leaves: an approach with hybrid features produced by dividing leaf images into two and four parts. *Proc J Appl Math Comput* 352:1–14
3. Yigit E, Sabanci K, Toktas A, Kayabasi A (2019) A study on visual features of leaves in plant identification using artificial intelligence techniques. *Proc J Comput Electron Agric* 156:369–377
4. Azadnia R, Kheiralipour K (2021) Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier. *Proc J Appl Res Med Aromat Plant* 25
5. Pacifico LDS, Britto LFS, Oliveira EG, Ludermir TB (2019) Automatic classification of medicinal plant species based on color and texture features. In: *Proceedings of 8th Brazilian conference on intelligent systems (BRACIS)*, Salvador, Brazil, pp 741–746
6. Sharma P, Aggarwal A, Gupta A, Garg A (2019) Leaf identification using HOG, KNN, and Neural Networks. In: *Proceedings of international conference on innovative computing and communications*, VSB-Technical University of Ostrava, Czech Republic, Europe, pp 83–91, 21–22 Mar 2019
7. Pechebovicz D et al (2020) Plants recognition using embedded Convolutional Neural Networks on Mobile devices. In: *Proceedings of IEEE international conference on industrial technology (ICIT)*, Buenos Aires, Argentina, pp 674–679, 26–28 Feb 2020
8. Prasad S, Singh PP (2017) Medicinal plant leaf information extraction using deep features. In: *Proceedings of IEEE Region 10 conference (TENCON)*, The Wembley St. Giles Premier Hotel, Penang, Malaysia, pp 2722–2726, 5–8 Nov 2017
9. Tan JW, Chang SW, Abdul-Kareem S, Yap HJ, Yong K-T (2020) Deep learning for plant species classification using leaf vein morphometric. *Proc IEEE/ACM Trans Comput Biol Bioinform J* 17(1):82–90
10. Varghese BK, Augustine A, Babu JM, Sunny D, Cherian S (2020) INFOPLANT: plant recognition using convolutional neural networks. In: *Proceedings of the 4th international conference on computing methodologies and communication (ICCMC)*, Erode, India, 11–13 Mar 2020, pp 800–807
11. Naresh YG, Nagendraswamy HS (2016) Classification of medicinal plants: an approach using modified LBP with symbolic representation. *Proc J Neurocomput* 173(Part 3):1789–1797
12. Sabu A, Sreekumar K, Nair RR (2017) Recognition of ayurvedic medicinal plants from leaves: a computer vision approach. In: *Proceedings of fourth international conference on image information processing (ICIIP)*, Jaypee University of Information Technology, Wagnaghat, District Solan, Himachal Pradesh, India, pp 1–5, 21–23 Dec 2017

13. Dileep MR, Pournami PN (2019) AyurLeaf: a deep learning approach for classification of medicinal plants. In: Proceedings of IEEE Region 10 conference (TENCON), Kochi, India, pp 321–325, 17–20 Oct 2019
14. Shah MP, Singha S, Awate SP (2017) Leaf classification using marginalized shape context and shape+texture dual-path deep convolutional neural network. In: Proceedings of IEEE international conference on image processing (ICIP), China National Convention Center, Beijing, China, pp 860–864, 17–20 Sept 2017
15. Samuel MJ (2021) Flawless detection of herbal plant leaf by machine learning classifier through two stage authentication procedure. Proc J Artif Intel Capsule Netw 3:125–139

SARS-CoV-2 Mortality Risk Prediction Using Machine Learning Algorithms to Aid Medical Decision-Making



Manish Sewariya and Sanjay Patidar

Abstract The coronavirus has affected the world in every possible aspect such as loss of economy, infrastructure, and moreover human life. In the era of growing technology, artificial intelligence and machine learning can help find a way in reducing mortality so, we have developed a model which predict the mortality risk in patients infected by COVID-19. We used the dataset of 146 countries which consists of laboratory samples of COVID-19 cases. This study presents a model which will assist hospitals in determining who must be given priority for treatment when the system is overburdened. As a result, the accuracy of the mortality rate prediction demonstrated is 91.26%. We evaluated machine learning algorithms namely decision tree, support vector machine, random forest, logistic regression, and K-nearest neighbor for prediction. In this study, the most relevant features and alarming symptoms were identified. To evaluate the results, different performance measures were used on the model.

Keywords SARS-CoV-2 · Mortality rate · Logistic regression · Random forest · Support vector machine · Decision tree · Machine learning

1 Introduction

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a novel version of coronavirus, began spreading in the Chinese province of Hubei in late 2019 and took many human lives [1]. The World Health Organization (WHO) classified the emerging coronavirus epidemic as a Public Health Emergency of International Concern (PHEIC) in January 2020 [2]. The infectious disease caused by the new coronavirus was given the formal name COVID-19 (Coronavirus Disease 2019) by WHO in February 2020, and a COVID-19 pandemic was announced in March 2020 [3].

M. Sewariya · S. Patidar (✉)

Department of Software Engineering, Delhi Technological University, New Delhi, India
e-mail: sanjaypatidar@gmail.com

As a result of the rise in positive cases, when a hospital's capacity is exceeded, the death rate rises [4]. To address the existing challenges in effective mortality risk predictions, timely clinical decision-making, and prevention of further fatality rates, machine learning is a tool that could be of great aid.

As a subfield of artificial intelligence (AI), machine learning (ML) allows for the extraction of high-quality prediction models from massive raw information [5]. It is a useful technique that's increasingly being used in medical research to enhance predictive modeling and uncover novel elements that influence a certain goal outcome [5, 6]. By providing evidence-based medication for risk evaluation, screening, prediction, and treatment planning, machine learning algorithms can minimize uncertainty and ambiguity. They promote trustworthy clinical decision-making and expect to enhance patient outcomes and quality of care [7, 8]. In this regime, scientists have been actively involved in publishing useful reports by utilizing ML algorithms for example Xiao et al. [9] reported that older people had a greater mortality risk, Sowmya Sundari and colleagues stated in 2021 that predictions made are able to construct a more accurate clinical decision diagnostic assistance system which avoids costly medical tests and checks, the patient saves money and time, and the patient may plan for appropriate treatment at the earliest stages of the condition as a preventative step [10]. Furthermore, in 2021, Sayed and team proposed that to help healthcare professionals anticipate the seriousness and fatality risk of a patient, efficient and accurate intelligent systems tools are required and developed a prediction model to forecast different degrees of severity risks for the COVID-19 patient based on X-ray scans [11, 12].

The goal of the present work is to create a COVID-19 mortality risk prediction model based on machine learning algorithms that used patients' regular clinical data. We are seeking answers to the following questions: (1) What are the most important determinants of COVID-19 in-hospital patient mortality? What is the most effective machine learning algorithm for creating a mortality prediction model? So, we used the dataset of 146 countries which consists of laboratory samples of around 2,670,000 confirmed COVID-19 cases for our study. We evaluated a few classified machine learning algorithms, and the respective accuracies were found on the above dataset namely decision tree (DT) (accuracy—89.10%), support vector machine (SVM) (accuracy—89.87%), random forest (RF) (accuracy—91.26% %), logistic regression (LR) (accuracy—89.24%), and K-nearest neighbor (K-NN) (accuracy—89.98% %) for predicting the mortality risk in patients infected with SARS-CoV-2 virus.

This research article discusses the development of a COVID-19 patient's mortality risk prediction model utilizing various machine learning approaches. Section 2 traverses through the research methods employed in the study. Sections 3 and 4 present the debate and findings on the viewpoints and importance of the selected features for prediction, while Sect. 5 brings the research to a close and provides a clear picture of future advances in the subject.

2 Research Methodology

This section describes the dataset and machine learning techniques used to make the predictions. Figure 1 depicts the linkages and information flow between the tasks that we employed to obtain the prediction outcomes. Data preprocessing is done, then features are extracted, and various techniques are applied.

2.1 Dataset

In this research, we used/employed the patients dataset of 146 countries from all over the globe [13] which consists of laboratory samples or we can say patients medical data samples of around 2,670,000 confirmed COVID-19 cases, where 307,382 labeled samples encompassing both female and male patients with 44.75 years of average age [13].

The presence of viral nucleic acid confirms the illness. Each patient had 32 data items in the original dataset, comprising demographic and physiological information which is represented in Table 1. While performing the data preprocessing steps, we deleted irrelevant and redundant data items. The unlabeled data samples have also been deleted. To manage missing values, data imputation methods include mean/mode/median replacement is used, and after that normalization is performed. We balanced our dataset in order to obtain an accurate and impartial model and then for training and testing the model we split the data into 70% training and 30% testing.

To train and evaluate our model, we constructed a balanced dataset with an adequate ratio of recovered and dead patients. The training dataset's data samples (patients) were chosen at random and are fully independent of the testing data. Table 1: initial features in the dataset.

2.2 Feature Extraction

The patient's health state is explained by numerous values on the outcome label. Patients who were done with treatment from the hospital or who were in a good state with no further symptoms were classified in the recovered category. At the moment



Fig. 1 System architecture

Table 1 Characteristics employed in the machine learning techniques are listed below

Symptoms	Pre-existing conditions	Demographics
<ul style="list-style-type: none"> • Anorexia • Fever • Shortness of breath • Chest pain • Gasp • Somnolence • Chills • Headache • Sore throat • Conjunctivitis • Kidney failure • Sputum • Dizziness • Myalgia • Cold • Expectoration • Myelofibrosis 	<ul style="list-style-type: none"> • Fatigue • Cough • Lesions on chest radiographs • Septic shock • Dyspnea • Obnubilation • Cardiac disease • Eye irritation • Respiratory distress • Rhinorrhea • Diarrhea • Hypertension • Emesis • Pneumonia • Hypoxia • Heart attack 	<ul style="list-style-type: none"> • Diabetes • COPD • Hypertension • Parkinson's disease • Chronic kidney disease • Asthma • Cerebral infarction • HIV positive • Cardiac disease • Dyslipidemia • Hypothyroidism • Cancer • Chronic bronchitis • Any chronic disease • Prostate hypertrophy • Coronary heart disease • Travel history • Hepatitis B
		<ul style="list-style-type: none"> • Age • Country • Province • Gender • Tuberculosis • City

of admittance to the hospital, healthcare personnel documented the symptoms. For better results in the prediction, we applied one-hot encoding to the dataset as this technique transforms categorical variables into a format that can be fed into machine learning algorithms to improve prediction accuracy.

After applying a one-hot encoding, a total of 1261 features were created and, on those features, we applied the feature extraction using PCA. As PCA follows dimensionality reduction where important and relevant features are extracted that captures most maximum variables. We retrieved features that include symptoms, physicians' medical reports, demographics, and physiological parameters from the original dataset after extracting.

2.2.1 PCA

PCA which is principal component analysis is a dimensionality-reduction approach that reduces the dimensionality of large datasets by converting a large set of variables into a smaller group that preserves the majority of the content in the bigger set. It is an uncontrolled learning approach for reducing size in machine learning. It is a mathematical procedure that converts the visibility of a connected element into a set of irrelevant feature lines via orthogonal conversion. Newly changed characteristics are key components. PCA seeks a low-dimensional environment in which to handle high-resolution data. Because the top characteristic reveals the appropriate differences across classes, and PCA minimizes the size by assessing the variability of each feature. It keeps key variables while rejecting less relevant ones since it is a way of extracting a feature. The key components used for PCA in our model are listed in 25 different values as: [2, 5, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, and 1150]. Every algorithm is being cross-validated on these set of features.

2.3 *Predictive Analytics Algorithm*

We built a prediction model using multiple machine learning methods after the feature extraction process. Support vector machine, random forest, decision tree, logistic regression, and K-nearest neighbor were among the techniques utilized in this study.

2.3.1 SVM

Support vector machine (SVM) is a machine learning approach that is typically used for regression and classification, although it is also used for pattern recognition and data analysis [14]. The SVM approach is used to determine the border with the maximum width. The model generated here is a classification model that is represented as a boundary where data of diverse attributes are allocated in one area. The

input data must be seen as a high-dimensional feature space in nonlinear classification, which may be done efficiently using kernel approaches [15]. In this study, kernel used is “rbf” and the training and testing accuracy was found to be 90.49 and 89.87%, respectively, at 350 features.

2.3.2 DT

Decision tree is a supervised machine learning approach for classification and regression problems. Decision tree being a supervised machine learning technique the dataset must be labeled. The categorization is done using a set of criteria using the indecision tree approach. In a decision tree, a node represents a characteristic, a branch represents a procedure, and a leaf node represents the conclusion. It can be expressed as a tree-like structure, which provides more accuracy and stability. In the first stage, a tree will be built with input features as its nodes. It will then select a feature for the next stage from the input features. It will then select a feature from the input features to predict the output, resulting in the greatest growth in knowledge apply the procedures outlined above to the subtrees are created by utilizing characteristics that are available and were never used before [16]. In this study, the training and testing accuracy was found to be 93.49% and 89.10%, respectively, at 100 features.

2.3.3 LR

In statistics, the generalized linear regression model which is logistic regression used for the prediction with the target class having 2-level events likely alive or dead, true or false, and lose or win [17]. To put it another way, chances are the proportion of the likelihood of an event occurring divided by the likelihood of it not occurring. If p stands for probability, then the equation $\frac{p}{1-p}$ stands for odds. The natural logarithm of the chances is used as a predictive factor in the logistic regression model, and the Eq. (1) expressed mathematically is [9].

$$l = \log_b \frac{p}{1-p} \quad (1)$$

In this study, the training and testing accuracy was found to be 89.70 and 89.24%, respectively, at 450 features.

2.3.4 RF

A machine learning ensemble model called a random forest. Which helps in solving both regression and classification problems. It is an ensemble model, which means it combines many machine learning techniques to outperform others. By randomly

picking a subset of the training dataset, random forest builds different decision trees. Using decision trees, it will predict the class of test class objects [18]. In this study, the training and testing accuracy was found to be 91.50 and 91.26%, respectively, at 800 features.

2.3.5 K-NN

For both classification and regression tasks, the simplest supervised learning approach, K-nearest neighbor, is used. It is a supervised learning method, which means the data must contain both input and output parameters for the model to be trained. For a given value of k , the K-NN method will locate the k closest data points. The class of the data point will then be determined based on the class of the biggest collection of data points with the same class. It employs the Euclidean distance to get the K-closest neighbor. The Euclidean distance Eq. (2):

$$d(x, y) = \sum_{j=1}^k \sqrt{(x_j - y_j)^2} \quad (2)$$

The methodology is the same for regression problems, except instead of neighbor classes, target values are used. Choosing an appropriate k is one of the most difficult tasks in KNN. The choice border will be more irregular if k is less, while a greater value of k will result in a smoother decision boundary [19]. In this study, value of $k = 5$ and the training and testing accuracy were found to be 92.16 and 89.98%, respectively, at 150 features.

2.4 Model Evaluation

After performing normalization on all models, they were trained and tested by cross-validating the number of features in a range. The ROC curve and the below confusion matrix are shown in Figs. 3 and 4.

The models were cross-validated on 2, 5, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, and 1150 features which can be seen in Table 2. After this, the optimal value of the number of features was found. The accuracy versus the number of features graph obtained for all models is shown in Fig. 2.

Table 2 Accuracies computed by models on each iteration

No. of features	SVM		LR		KNN		DT		RF	
	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.
2	87.94	87.91	85.02	84.97	87.37	87.30	93.50	88.02	88.86	88.91
5	88.39	88.52	87.75	87.88	90.28	89.99	93.50	87.96	89.49	89.77
50	89.52	89.30	88.90	88.77	90.78	89.99	93.50	88.94	91.24	91.46
100	89.85	89.88	89.01	89.02	91.09	88.08	93.50	89.10	91.15	91.46
150	89.96	89.88	89.28	89.16	92.17	89.99	93.51	88.08	91.15	91.24
200	90.09	89.82	89.42	89.13	88.17	87.77	93.51	89.07	91.21	91.21
250	90.19	89.82	89.54	89.13	90.81	89.57	93.51	88.66	91.06	91.10
300	90.44	89.88	89.59	89.10	92.19	88.02	93.51	88.13	91.29	91.29
350	90.49	89.88	89.68	89.13	90.65	90.02	93.51	88.88	91.30	91.43
400	90.47	89.88	89.67	89.18	92.20	89.93	93.51	88.41	91.28	91.15
450	90.47	89.88	89.71	89.24	90.44	89.74	93.51	88.74	91.23	91.26
500	90.46	89.77	89.71	89.16	91.78	87.69	93.51	89.05	91.31	91.49
550	90.50	89.77	89.73	89.18	90.52	89.68	93.51	88.94	91.22	91.40
600	90.54	89.82	89.77	89.21	90.52	88.21	93.51	88.57	91.28	91.49
650	90.58	89.79	89.80	89.21	90.25	89.93	93.51	88.71	91.23	91.35
700	90.65	89.79	89.77	89.16	88.82	88.05	93.51	88.77	91.30	91.15
750	90.69	89.74	89.79	89.16	90.30	89.35	93.51	88.30	91.25	91.29
800	90.72	89.71	89.81	89.13	92.22	89.55	93.51	88.41	91.50	91.26
850	90.78	89.77	89.81	89.10	90.42	88.02	93.51	87.91	91.25	91.46
900	90.79	89.77	89.81	89.13	90.44	89.85	93.51	88.69	91.25	91.35

(continued)

Table 2 (continued)

No. of features	SVM		LR		KNN		DT		RF	
	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.	Train Acc.	Test Acc.
950	90.87	89.79	89.83	89.10	91.36	87.77	93.51	88.74	91.28	91.40
1000	90.92	89.77	89.83	89.13	90.47	89.85	93.51	88.63	91.11	91.21
1050	90.96	89.77	89.83	89.13	87.65	89.24	93.51	88.77	91.17	91.24
1100	90.96	89.77	89.84	89.13	91.24	89.32	93.51	88.82	90.98	91.04
1150	90.96	89.77	89.84	89.13	91.24	89.32	93.51	88.33	91.10	91.07

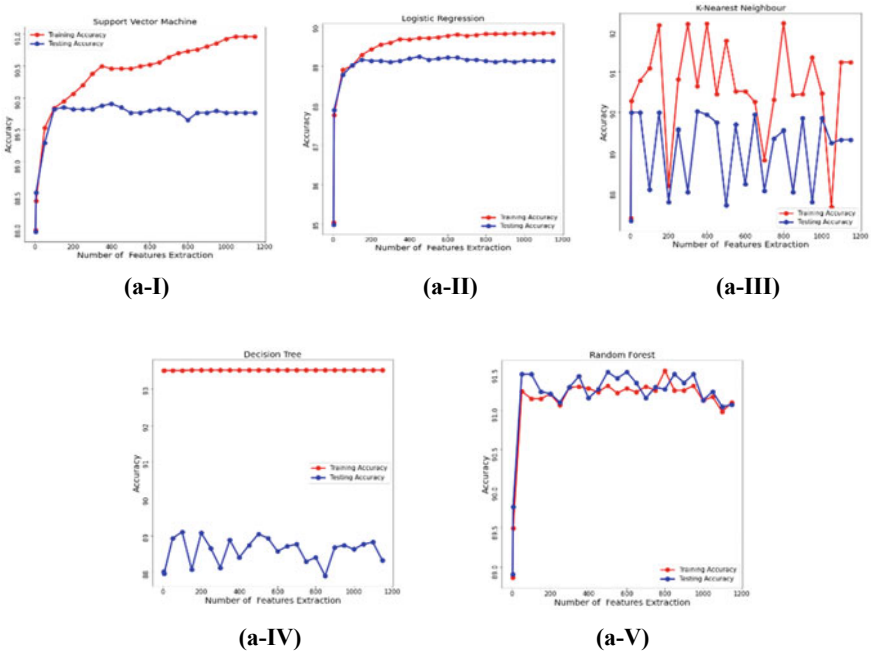


Fig. 2 a Accuracy versus number of feature graphs of all the ML techniques. **a-I** Graph of SVM algorithm performed on different number of features. Initially, we observed that model underfits, probably due to low amount of features that have low variance, but that is solved as we increase the number of features; **a-II** Graph of LR algorithm performed on different number of features. Initially, we observed that model underfits, probably due to low amount of features that have low variance, but that is solved as we increase the number of features. **a-III** Graph of KNN algorithm performed on different number of features. Initially, we observed that model underfits, probably due to low amount of features that have low variance, but that is solved as we increase the number of features; **a-IV** Graph displays cross validation of DT model on different number of features. Also, it can be observed that model is overfitting a bit. **a-V** Graph displays cross validation of RF model on different number of features. Also, it can be observed that model resolves the overfitting of the DT model

3 Results

In this research, the coronavirus infection rate was compared to the five models' predictions. Decision tree (DT), support vector machine (SVM), random forest (RF), logistic regression (LR), and K-nearest neighbor were used to assess the created model, which included accuracy, precision, recall, F1-score, ROC, AUC, and confusion matrix. Table 3 summarizes all of the findings, demonstrating the accuracy of several machine learning algorithms in predicting death in COVID-19 patients. Random forest was found to be the most effective and has the highest accuracy rating

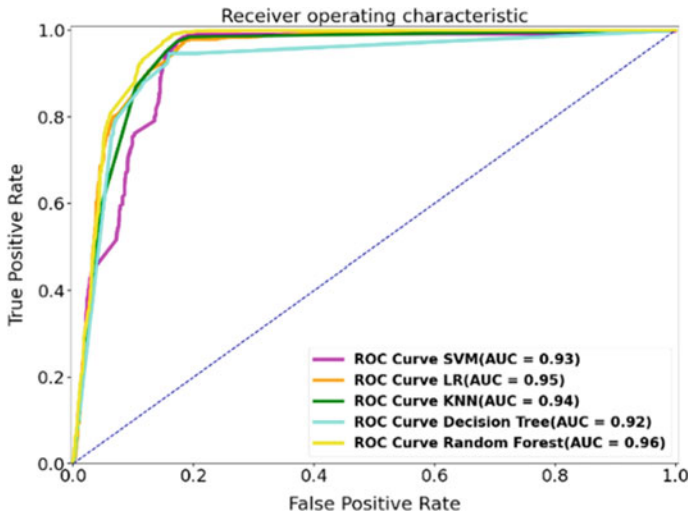


Fig. 3 ROC curve for all the algorithms

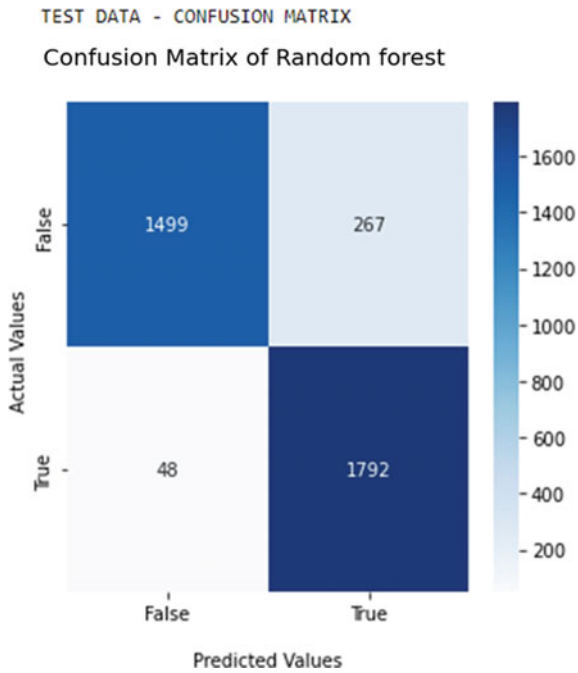


Fig. 4 Confusion matrix of RF

Table 3 Results of all the techniques used using different performance metrics

S. no.	Algorithm/measures	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	ROC (%)
1	Support vector machine (SVM)	89.87	90.44	89.87	89.82	92
2	Logistic regression (LR)	89.24	89.24	89.58	89.20	95
3	K-nearest neighbor (KNN)	89.98	90.34	89.98	89.95	94
4	Decision tree (DT)	89.10	89.33	89.10	89.07	93
5	Random forest (RF)	91.26	91.86	91.26	91.22	96

of 91.26%. Each machine learning method's ROC curves and AUC are displayed and compared in this study (Fig. 3). Random forest method classifier's performance is described and shown using a confusion matrix (Fig. 4), which also provides insight into what the model misclassifies.

4 Discussion

In this work, to estimate the death rate for COVID-19 patients, we enveloped cutting-edge machine learning methods using a huge dataset of positive cases gathered from all across the globe. The approach for finding the relevant and important features we developed a model using feature extraction which cross validate the different set of number of features on different algorithms to obtain good results. Several performance indicators were used to assess the created algorithms. The assessment findings show that the generated models are very accurate and effective. Other research has demonstrated that from clinical data and blood test results, it is possible to estimate the death rate in COVID-19 patients [20]. However, we focused on demographic data, physiological information, pre-existing conditions, and patient symptoms. Using the random forest model, we were able to achieve an exceptional accuracy of 91.26%. Furthermore, past research has mostly focused on data gathered in China [20, 21]. We combined hospital data from across the world to construct a more complete model that is relevant to the whole world's population, rather than training on data from a single location.

5 Conclusion and Future Work

The goal of this study was to develop a predictive model that would aid hospitals and medical institutions in improving survival rates by providing a precise and accurate tool to aid in medical decisions and better prioritization. Patients with COVID-19

during the pandemic, our approach can accurately assess the mortality risk in COVID-19 patients based on their physiological states, symptoms, pre-existing scenarios, and demographic data. This strategy or method can assist hospitals, health organizations, and caregivers in determining who should be handled foremost before other patients, prioritizing patients when the system is saturated, and reducing wait times for needed care. This study might be expanded to cover other diseases, empowering the healthcare system to react faster in the case of an epidemic or pandemic.

References

1. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Leung GM, Feng Z (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 382:1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
2. Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K (2020) Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. *J Travel Med* 27:taaa008. <https://doi.org/10.1093/jtm/taaa008>
3. World Health Organization (2020) Statement on the second meeting of the International Health Regulations Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), World Health Organization (WHO). Archived from the original on 31 Jan 2020
4. McConghy T, Pon B, Anderson E (2020) When does hospital capacity get overwhelmed in USA? Germany? A model of beds needed and available for coronavirus patients
5. Shanbehzadeh M, Ilam I, Nopour R, Kazemi-Arpanahi H (2021) Comparison of four data mining algorithms for predicting colorectal cancer risk. *J Adv Med Biomed Res* 29:100–108. <https://doi.org/10.30699/jambs.29.133.100>
6. Hernandez-Suarez DF, Ranka S, Kim Y, Latib A, Wiley J, Lopez-Candales A, Pinto DS, Gonzalez MC, Ramakrishna H, Sanina C, Nieves-Rodriguez BG, Rodriguez-Maldonado J, Feliu Maldonado R, Rodriguez-Ruiz IJ, da Luz Sant'Ana I, Wiley KA, Cox-Alomar P, Villablanca PA, Roche-Lima A (2021) Machine-learning-based in-hospital mortality prediction for transcatheter mitral valve repair in the United States. *Cardiovasc Revasc Med* 22:22–28. <https://doi.org/10.1016/j.carrev.2020.06.017>
7. Shipe ME, Deppen SA, Farjah F, Grogan EL (2019) Developing prediction models for clinical use using logistic regression: an overview. *J Thorac Dis* 11:S574–S584. <https://doi.org/10.21037/jtd.2019.01.25>
8. Gao Y, Cai G-Y, Fang W, Li H-Y, Wang S-Y, Chen L, Ye F, Gao Q-L (2020) Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun* 11:5033. <https://doi.org/10.1038/s41467-020-18684-2>
9. Xiao Z (2021) COVID 19 mortality rate prediction based on machine learning methods. In: 2021 IEEE international conference on computer science, electronic information engineering and intelligent control technology (CEI), pp 169–177. IEEE, Fuzhou, China. <https://doi.org/10.1109/CEI52496.2021.9574541>
10. Sowmya Sundari LK, Ahmed ST, Anitha K, Pushpa MK (2021) COVID-19 outbreak based coronary heart diseases (CHD) prediction using SVM and risk factor validation. In: 2021 innovations in power and advanced computing technologies (i-PACT), pp 1–5. IEEE, Kuala Lumpur, Malaysia. <https://doi.org/10.1109/i-PACT52855.2021.9696656>
11. Mostafiz R, Uddin MS, Alam N-A, Mahfuz Reza M, Rahman MM (2020) Covid-19 detection in chest X-ray through random forest classifier using a hybridization of deep CNN and DWT optimized features. *J King Saud Univ Comput Inform Sci* S1319157820306182. <https://doi.org/10.1016/j.jksuci.2020.12.010>

12. Sayed SA-F, Elkorany AM, Sayed Mohammad S (2021) Applying different machine learning techniques for prediction of COVID-19 severity. *IEEE Access* 9:135697–135707. <https://doi.org/10.1109/ACCESS.2021.3116067>
13. Xu B, Gutierrez B, Mekaru S, Sewalk K, Goodwin L, Loskill A, Cohn EL, Hswen Y, Hill SC, Cobo MM, Zarebski AE, Li S, Wu C-H, Hulland E, Morgan JD, Wang L, O'Brien K, Scarpino SV, Brownstein JS, Pybus OG, Pigott DM, Kraemer MUG (2020) Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci Data* 7:106. <https://doi.org/10.1038/s41597-020-0448-0>
14. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–297. <https://doi.org/10.1007/BF00994018>
15. Yan X, Chowdhury NA (2013) A comparison between SVM and LSSVM in mid-term electricity market clearing price forecasting. In: 2013 26th IEEE Canadian conference on electrical and computer engineering (CCECE), pp 1–4. IEEE, Regina, SK, Canada. <https://doi.org/10.1109/CCECE.2013.6567685>
16. Woldaregay AZ, Årsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, Hartvigsen G (2019) Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. *Artif Intell Med* 98:109–134. <https://doi.org/10.1016/j.artmed.2019.07.007>
17. Dreiseitl S, Ohno-Machado L (2002) Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform* 35:352–359. [https://doi.org/10.1016/S1532-0464\(03\)00034-0](https://doi.org/10.1016/S1532-0464(03)00034-0)
18. Liu S, Li H, Zhang Y, Zou B, Zhao J (2019) Random forest-based track initiation method. *J Eng* 2019:6175–6179. <https://doi.org/10.1049/joe.2019.0180>
19. Huang J, Wei Y, Yi J, Liu M (2018) An improved kNN based on class contribution and feature weighting. In: 2018 10th international conference on measuring technology and mechatronics automation (ICMTMA), pp 313–316. IEEE, Changsha. <https://doi.org/10.1109/ICMTMA.2018.00083>
20. Yan L, Zhang H-T, Goncalves J, Xiao Y, Wang M, Guo Y, Sun C, Tang X, Jin L, Zhang M, Huang X, Xiao Y, Cao H, Chen Y, Ren T, Wang F, Xiao Y, Huang S, Tan X, Huang N, Jiao B, Zhang Y, Luo A, Mombaerts L, Jin J, Cao Z, Li S, Xu H, Yuan Y (2020) A machine learning-based model for survival prediction in patients with severe COVID-19 infection. *Epidemiology*. <https://doi.org/10.1101/2020.02.27.20028027>
21. Du R-H, Liang L-R, Yang C-Q, Wang W, Cao T-Z, Li M, Guo G-Y, Du J, Zheng C-L, Zhu Q, Hu M, Li X-Y, Peng P, Shi H-Z (2020) Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *Eur Respir J* 55:2000524. <https://doi.org/10.1183/13993003.00524-2020>

A Comprehensive Study on Analysis and Prediction of Pollution Data Using Machine Learning



R. D. Aishwarya, C. Sahana, V. J. Deepa, J. Durgashree, S. Gowrishankar, and A. Veena

Abstract One of the most crucial global issues today is environmental pollution. Environmental pollution is caused by various human activities. It is destructive to all living organisms. Pollution is an essential environmental catastrophe. It causes diseases and is formidable to overcome. Due to the increase in pollution, it is now possible to monitor and predict the pollution data, mainly in urban areas. The main types of pollution are air, water, land pollution, noise, etc. Feature selection approaches have evolved. This tool helps to reduce the load of data that is transmitted to the model being used. It uses only the required data and eliminates the noise in the dataset. This paper summarizes the work of various researchers in the analysis of pollution data. It also reviews the different algorithms used by authors to study pollution data and compares the performance metrics of the algorithms. The algorithms are grouped based on the real-world applications of pollution analysis.

Keywords Time series · Data analysis · Pollution · Prediction · Machine learning

R. D. Aishwarya (✉) · C. Sahana · V. J. Deepa · J. Durgashree · S. Gowrishankar · A. Veena
Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology,
Bengaluru, Karnataka 560056, India
e-mail: 1da18cs010.cs@drait.edu.in

C. Sahana
e-mail: 1da18cs133.cs@drait.edu.in

V. J. Deepa
e-mail: 1da18cs043@drait.edu.in

J. Durgashree
e-mail: 1da18cs048@drait.edu.in

S. Gowrishankar
e-mail: gowrishankamath@acm.org

A. Veena
e-mail: veenaal@acm.org

1 Introduction

Pollution is caused because of the entry of contaminants into the environment. The contaminants can be natural or manmade but is majorly manmade. The pollution has increased due to the rapid increase in urbanization and industrialization [1]. The major types of pollution are air, water, land, noise pollution, etc.

Air pollution is caused due to the release of toxic chemicals into the atmosphere. Some primary pollutants produced by human activities consists of particulate matter (PM), nitrogen dioxide (NO_2), ozone, carbon monoxide, etc. [2, 3]. Particulate matter is a combination of liquids and solids, covering sulfates, nitrates, carbon and complicated organic chemicals, water, and mineral dirt suspended in the air. PM differs in size. But smaller pieces are the most damaging ones, which are known as PM10 and PM2.5. Nitrogen dioxide is a gas, and it is a vital element of urban air pollution. Top levels of NO_2 can inflame and irritate the edge of our airways, creating a flare-up of asthma or COPD and also there are symptoms such as difficulty in breathing and coughing. When the O_3 is at a high level, it can cause discomfort when you breathe, and it reduces your lung capacity and also triggers asthma symptoms. Air pollution can cause respiratory, cardiovascular diseases, difficulty in breathing, etc.

Water pollution is caused by the spilling of industrial waste and wastewater to surface water bodies. Research on water pollution has highlighted the chance of surface oil and other contaminants [4]. The three major measures of water pollution are biochemical oxygen demand (BOD), concentration of ammonium ions, and concentration of phosphate ions [5]. In broad cities, when the level of phosphate and ammonium is high, it can normally be detected in agricultural areas and could indicate sewage contamination. Some primary pollutants include Ph, alkalinity, and total phosphorus. Pollution can change a water's pH, which can harm plants and animals living in the water areas. In surface waters, if the alkalinity is higher, it will cushion acid rain and also other wastes of acid and can prevent pH changes that can damage the aquatic life. In excess amounts, problems related to quality of water can be led by phosphorus such as eutrophication and harmful algal growth.

A human ear can normally detect sounds from 0.02 to 20 dB. But the levels can increase to 45–55 dB. There are higher levels of sound produced. The frequency of this range is not good for the ear. This high range sound is called noise. Noise means the sound that is unpleasant to the ear. In the long run, noise may cause damage to the ears. Noise pollution can be caused in many ways like industry, automobiles, crackers, construction, etc. The lives of millions of people can be adversely affected by noise pollution. The noise pollution's potential health effects are countless, and it includes illnesses related to stress, high blood pressure, speech intervention sleep disruption, hearing loss, and loss of productivity [6].

In this paper, in Section one, we have discussed about pollution and its effects on human life. Because of the rapid increase in pollution in recent times, we need to determine pollution levels at an early stage to control pollution. For early detection, we can use various machine learning algorithms and deep learning techniques. In Section two, we have discussed the various machine learning algorithms that can be

used to analyze and predict the time series pollution data. Some of the algorithms used are convolutional neural network (CNN), long short-term memory (LSTM), autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), logistic regression, autoregression, and others, and finally, we conclude the paper.

2 Models Utilized for Classification

2.1 Models Utilized for Analysis and Prediction of Air Pollution

2.1.1 Gated Recurrent Units (GRU)

This paper [7] tests the forecasting approach using the database that consist of the PM2.5 concentration data of Beijing's US Embassy. It also takes data from Beijing Capital International Airport. The dataset considered here covers data on an hourly basis for a span of 4 years from 2010 to 2014. It contains seven features like concentrations of pollutants, especially PM2.5, humidity, temperature, pressure, direction of wind, and speed of wind and rain. It is very important to recognize the connection between the numerous factors that influence in increase of pollution and the concentration of PM2.5 to enlarge an efficient model that is used to make predictions. The author has tried to remove the irrelevant factors which are a burden for the model. In this paper, the CB-GRU model is used for forecasting future values of PM2.5 concentration. Firstly, concentration of PM2.5 data is passed as a model input. The future values of PM2.5 concentration is the output. It is used to carry out a multi-step prediction. This model has three parts as shown in Fig. 1. In the first step, a 1D convolutional neural network is used to execute local feature extraction and dimensionality reduction on the input variables. Low-dimensional feature sequences are created from original data by the process of convolution and pooling. In the second part, the feature sequences into the bidirectional GRU neural networks. The third part of the model has a fully coupled layer stacked to it. The output (last) layer of the model has only one neuron. This layer does not have any function for activation. It generates the PM2.5 concentration of the value predicted. Convents have certain properties such as convolution operation, data efficiency, feature extraction, etc. These properties of convents make it useful for computer vision and sequence processing. The bi-GRU is used for time series forecasting. For clarifying sequence data, recurrent neural network (RNN) is a special form of neural network used but it has some drawbacks like exploding gradient and vanishing gradient. Modification of RNN structure lea LSTM and GRU. LSTM can track long-term information through its three gates. The gates are input gate, forget gate, and output gate. GRU is basically an improvised version of LSTM. It has only two gates. They are the update gate and the reset gate. Having two gates gives GRU a simpler architecture. Therefore, it requires less computations. Hence, it can be trained faster.

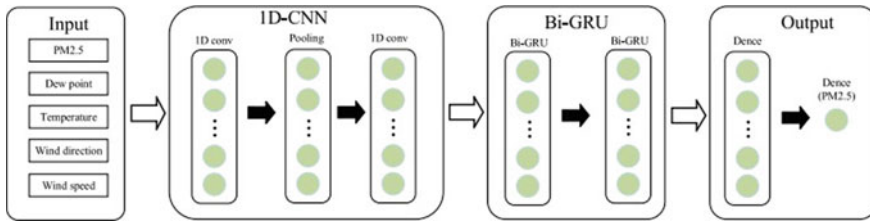


Fig. 1 Working of Bi-GRU on PM2.5 concentration levels [7]

2.1.2 ARIMA and SARIMA

ARIMA stands for autoregressive integrated moving average. ARIMA is a statistical analysis model. This model is primarily used for data that involves time series analysis.

In this model, autoregression is applied on the variables of interest. Integration refers to the differencing of the data to make it stationary. Moving average implies that this model uses errors of the previous predicted data, rather than the values.

SARIMA stands for seasonal ARIMA. It is an extension of ARIMA which supports time series data that has a seasonal component. ARIMA has only three parameters P , Q , and D for non-seasonal data. SARIMA has four parameters. The last parameter indicates the seasonality. They are

1. P : Autoregression order of the seasonal trend
2. Q : Difference order of the seasonal trend
3. D : Moving average order of the seasonal trend
4. m : The duration of a single seasonal period

There are two types of ARIMA models based on the difficulty of extracting seasonal trends from the dataset [1]. They are

1. Simple seasonal model: here, $P = D = 0$
2. Product seasonal model: Denoted as $ARIMA(P, Q, \text{and } D)$. This model is used to define the autocorrelation between many time-dependent random variables. It includes least squares, maximum likelihood, and moment estimation [1].

ARIMA avoids the problem of multivariate analysis, but it is computationally expensive compared to other models.

During the analysis, it was observed that PM2.5, PM10, O_3 , SO_2 , and CO are the main pollutants that cause air pollution. Especially, PM2.5 affects the human health and the climate of the earth adversely. The study also uses concepts from other domains such as environmental science, atmospheric chemistry, meteorology, and geochemistry to predict pollutant levels. The prediction and verification can be done on a daily/weekly/monthly/yearly basis.

Akaike's information criteria (AIC) and Bayesian information criteria (BIC) are parameters that are used for model selection. The author has measured these criteria for the ARIMA model. The AIC is approximately 856, and the BIC is 858 [1].

2.1.3 Logistic Regression and Autoregression

This work [8] describes a system that focuses on ml methods to predict and detect PM_{2.5} values using a data collection which includes conditions of the atmosphere in a particular area. The dataset taken to train the algorithm for air quality detection was collected from the University of California, Irvine repository. The suggested system performs two functions: (i) based on supplied atmospheric variables, detects PM_{2.5} levels and (ii) predicts PM_{2.5} levels for a specific date.

To evaluate if the user defined input is contaminated or unpolluted, logistic regression is used. When the dependent variable is dichotomous, logistic regression is the most appropriate to use. For instance, in this case, the dataset is divided into two categories: contaminated and unpolluted. This technique categorizes the training data into two, as either 1 (contaminated) or 0 (unpolluted) using the logit function and tests its correctness using the test data. The user input also yields a 0/1 outcome. In order to determine the readings of the future, an autoregressive model uses data from the past as input. When a relationship exists between the values in a time series and the values that come after and before them, autoregression is used to forecast. Logistic regression works best with standard deviation accuracy of 0.000612 and mean accuracy to be 0.998859. The mean squared error (MSE) for autoregression was 27.00.

2.1.4 Convolutional Neural Network

In this article [9], deep learning models for forecasting the quality of air focusing primarily on particulate matter (PM_{2.5}) were proposed. The modules that lay the foundation to the model comprise two deep neural networks. They are one-dimensional convolutional neural networks (1D-CNNs) and bidirectional long short-term memory networks (Bi-LSTM). It was reported that this was the first study to incorporate various one-dimensional CNNs and bidirectional LSTM. Bidirectional LSTM through two independent hidden layers processes the time series data in two indications, and these data are chained and taken forward to the output layer. The dataset used to perform forecasting was collected from Beijing air quality from UCI every hour. An alongside framework for deep air quality forecasting (DAQFF) was proposed which by using hybrid deep learning model conveys the dynamic, spatial-temporal, and nonlinear features of multivariate time series data of air quality.

The model can identify the trend pattern of local using one-dimensional CNNs and extract the characteristics and lengthy spatial-temporal reliance of multivariate time series data related air quality using bidirectional LSTM from both previous and future context by applying time series data in both further as well as reverse directives, for instance, PM_{2.5}, temperature, and wind speed. The extracted countless features of one-dimensional CNNs are chained and fed into given bidirectional LSTMs. From the experiment conducted using the datasets of real-world, the reported conclusions showed that the model is efficient in handling PM_{2.5} forecasting of air pollution.

2.1.5 KNN and Naïve Bayes Theorem

The data from the sensors is transferred to the webpage and ThinkSpeak (IoT). To check accuracy, the author has compared this data against the trained data. In this paper, they merge IoT and machine learning algorithms for monitoring the pollution level. The machine learning algorithms, KNN, and Naïve Bayes used here, fall under supervised learning, specifically under classification. KNN considers K-nearest neighbors (Datapoints) to predict the class or continuous value for the new Datapoint. Naïve Bayes algorithm is a group of algorithms which follow the same principle. The principle is, when the features are classified into pairs, they are independent of each other. Also the other assumption is that all the features make equal contribution to the outcome. The author has used various components such as temperature and humidity sensor, carbon monoxide gas sensor, analog to digital converter, and global system for mobile communication (GSM). All these components are interfaced with raspberry pi microcontroller using series cable. These various sensors are used to sense the presence of harmful gas, humidity, temperature, sound level, and continuously this data is transmitted to raspberry pie. Then, the microcontroller processes this data and sends it to server over IoT. KNN and Naïve Bayes are used for data mining in the webpage. Lastly, models are compared to see which of them have made more accurate predictions [10].

2.2 Models Utilized for Analysis and Prediction of Water Pollution

2.2.1 LS-SVM

To prognosticate water quality, the LS-SVM is used to build a nonlinear time series forecasting model [11]. Figure 2 represents the implementation framework to predict the water quality. To decrease modeling errors, first perform processing on raw data followed by zero-mean normalization, then get a matrix from 1D time series by conversion using Takens' theory to acquire the link between data relations, and thus gaining access to as much information as viable. Least squares support vector machine technique implements RB function instead of RBF kernel function, and LSSVMLab Matlab toolbox framework was used as the basic package, parameter $\text{Gam} = 10$, $\text{sig}2 = 0.5$, and $\text{Type} = \text{'function estimation'}$ [11]. To train the LS.

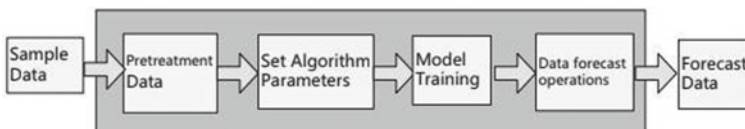


Fig. 2 Implementation framework to predict the water quality [11]

SVM technique, the train LS-SVM function is used. To predict total phosphorus, this algorithm is tested on monitoring data of river water. The result obtained were contrasted with the RBF and BP networks. The calculated relative error of the proposed model LS-SVM was 0.064% when compared to the RBF and BP network, which was 0.529 and 0.552%, respectively. The MRE and RMSE of the LS-SVM method are substantially fewer than those of the BP and RBF network technique, demonstrating that the LS-SVM has high accuracy in terms of prediction. The LS-SVM technique outperforms the RBF and BP (multi-layer) neural networks in the small sample condition with noise.

2.2.2 STL Model

Wan and Li [12] analyzes the parameters of water quality trend applying seasonal trend decomposition and LOESS (STL) of the canals in south Florida. They also predict the total value of nitrogen and phosphorous and turbidity of four canals using exponential polishing and additive Holt-winters method from 2016 to 2020 using the dataset downloaded from South Florida Water Management District (SFWMD) from January 1979 to April 2014.

Using decomposition model, the time series has been forecasted, the future values for each component are calculated, and it is added back together to obtain a prediction. Here, the challenge would simply be finding the best model for each component and it does not handle trading day or calendar variation automatically, and it provides facilities only for additive decompositions.

The STL method uses locally fitted regression models to break down into three components: (1) trend, (2) season, and (3) residual. Locally estimated scatterplot smoothing (LOESS) is used by STL to extract smooths estimates of three components. LOESS is weighed as effortless and a nonparametric statistical method. The trend component is considered as moving propensity with least frequency. The components of seasonal is viewed as alteration with stable seasonal disturbance with high frequency. The residuals component which has random disturbance is considered as uneven variation. STL method which implements LOESS of three times and it is a recursive process and changing average process in each recursion. It was found that a suitable method for time series with consistent seasonal variations was Holt-winters method and additive Holt-winters method. This model has been used in forecasting the variation of total nitrogen, total phosphorous, and turbidity in the 5 years of 2016–2020. As a result, the total nitrogen concentration was found to be decreasing trend and total phosphorous concentration was developing trend.

2.3 Models Utilized for Analysis and Prediction of Noise Pollution

2.3.1 LSTM

The LSTM (two-layer) network was developed for environmental noise prediction in the presence of enormous data volumes, and four sets of IoT system were deployed in a Shaoxing administrative district, China to collect environmental noise data in this study [13]. Identifying essential hyperparameters is crucial, so a streamlined dataset is used to improve the rate of parameter adjustment. To optimize the learning rate, Adam algorithm was used. Stationary test of time series was conducted using the augmented Dickey–Fuller technique. The LSTM system used has one input, two hidden, and one output layers. A three-dimensional tensor is fed an input which includes batch size-number of input sample count, window size-time series length, and number referring to dimension of output. On the test datasets with 10-min intervals, the MAE readings are 0.31–0.69 dB and RMSE readings are 0.48–0.93 dB [13]. Three conventional prediction models were utilized in this study to examine the LSTM model's prediction accuracy: SVM, RW, and SAE are their names. The RMSE and MAE values of LSTM were found to be lowest when compared to RW, SVM, and SAE, hence proving that noise forecasting results are obvious.

2.3.2 TRANEX

This study [12] develops its own models and methods for evaluating traffic noise exposure (TRANEX) to estimate exposures of noise for the resident London population from 2003 to 2010. They have used a method called calculation of road traffic noise (CoRTN) to calculate traffic noise estimates. This method was enlarged in open-source GIS and also using geographic information system (GIS) preprocessing of special evidence was attempted. TRANEX evaluates by estimates of noise which were correlated with the measurements of noise built in Leicester and Norwich which are the cities in British. TRANEX evaluates by estimates of noise by making use of noise measurements data which were as a part of earlier accomplishment gathered and attempted by the authors in the fifth framework of EU funded Program called HEAVEN in 2002 and HEARTS in 2005. To evaluate models, a sequence of working statistics was used such as the coefficient of determination, Spearman's rank correlation coefficient, root mean square error (RMSE), and the difference of computed and modeled levels of noise. The datasets for the implementation of noise models in London were collected from London's king's college and downloaded from digimap.

This model has predicted a minimum night time noise level that exceed one hour of 38 dB for an individual small road. It is been found that the least possible value modeled of night time that exceeds one hour in London was 42.4 dB. TRANEX in Leicester and Norwich has been evaluated from the information obtained from local traffic models and land use datasets. In TRANEX, traffic information is used

to assigned traffic origin points. From each of these points to receptors, a continuous path is produced. Exposures of the native London population was modeled by applying the postcode's locations which has the population of 8.61 million and the conclusions was presented for the year 2008. The end results obtained suggest that to daytime road traffic noise levels 12% people are revealed greater than or equal to 65 dB and to night time road traffic noise levels 19% people are revealed greater than or equal to 55 dB.

2.3.3 CNOSSOS-EU

Noise pollution has various effects on health and affects the individual differently. Noise pollution can be controlled by estimating an individual's exposure to noise from various sources within a considered area [14]. To estimate this exposure, the author has considered various factors such as traffic flow, land cover, and road network layouts based on their residential address. For models to produce efficient results data availability, precision and affordability matters. By considering a broader geographical scale for estimation, the input is standardized but may not be precise. Hence, the estimation is mostly done on a smaller scale restricted to a city or a municipality area. In this paper, they consider common noise assessment methods (CNOSSOS-EU) noise modeling framework for Europe. This model makes it easier to directly compare results from different countries. This modeling framework is standardized as it can take varying inputs from different data sources over a broad geographical area. CNOSSOS-EU is an efficient tool used for noise mapping all over Europe. The algorithms of this model were executed in PostGIS. GIS stands for geographical information system. PostGIS is a freely available framework. It is an extension of PostgreSQL database. The working of model is as follows: for each receptor, GIS finds all roads within a 500 m radius and divides them into source points at 20 m intervals. This division is done in series for each segment of road. Then, rays are projected from receptor to source point in a direct path. For each such ray path, noise level is calculated at each source using the data available about flow of traffic and the empirical relationships defined by CNOSSOS-EU. After this, noise levels are calculated at the receptor as per CNOSSOS-EU sound propagation algorithms. At last, all the noise level estimates from each path of the projection of the receptor are summed together logarithmically to provide a single patterned noise exposure value at each receptor point. For this model, two assumptions were made. Firstly, they had assumed an initial noise level, which corresponded to a few hundred vehicles at a 10 m distance from each receptor point per day. Second, the calculations included the corrections to regulate the noise level at the source according to geometrical divergence, barriers, and relative heights of barriers and receptor points. For model evaluation, this paper compares predicted noise levels with the measured noise levels (Table 1).

Table 1 Summary of various algorithms used in time series forecast

No.	Model	Application	Advantage	Disadvantage	Performance metrics
1	GRU [7]	Air	Controls the flow of information. It is computationally more efficient	Slow convergence Low learning efficiency	1. Root mean square error 2. Symmetric mean absolute percentage error
2	SARIMA [1]	Air	This model can be used for large amounts time series analysis of data It also avoids the problem of multivariate models	It is computationally expensive	1. Mean absolute percentage error
3	Logistic regression and autoregressive model [8]	Air	Autoregressive model predicts the value of PM2.5 a week prior to the present date	Logistic regression can only be used to perform analysis if the dependent Variable is dichotomous The independent variable should not have multicollinearity	1. Mean squared error
4	CNN [9]	Air	CNN recognizes the extensive features without any human supervision and has the highest accuracy	CNN needs a huge dataset to process and train the neural network, and it is significantly slower due to certain operations	1. Mean absolute error 2. Root mean square error
5	LS-SVM model [11]	Water	The prediction accuracy of LS-SVM method is high, and it works satisfactorily for real-time small quantities of data samples		1. Mean relative error 2. Root mean square error

(continued)

Table 1 (continued)

No.	Model	Application	Advantage	Disadvantage	Performance metrics
6	STL [15]	Water	STL will handle any type of seasonality, it is not like it handles only monthly data and quarterly data. The change of rate can be measured by the user and the allowance of seasonal component to switch over time	In particular, STL does not hold trading day or calendar deviations automatically, it supplies facilities only for additive decompositions	1. Mean absolute error 2. Mean square error
7	LSTM [13]	Noise	The result of this model has high precision prediction, avoids long-term dependency, and outperforms three existing classic models (SAE, RW, and SVM)	As the model was not sufficiently trained and along with few datasets, this led to poor performance in the thirty minutes interval dataset	1. Mean absolute error 2. Root mean square error
8	TRANEX [12]	Noise	TRANEX is built in such a way that with adequate detailed resources on traffic flows, speeds, and composition, it is ready to be transferable to many other different cities and areas in the UK	TRANEX is not dynamic with respect to representing the fluctuating/shifting automobile speeds along road links	Road traffic noise calculation (CoRTN)

3 Conclusion

In recent years, the global environmental pollution has increased rapidly and is a global concern that needs to be addressed urgently. If pollution keeps increasing like this, it would be difficult to sustain health of various living organisms. Hence, early detection of pollution levels is required, which will be helpful in the control and reduction of pollution levels and it also helps in waste minimization.

Each model as discussed above presents machine learning algorithms that deal with various time series datasets involving different types of pollution like air, water,

and noise. The algorithms are also used to make predictions of future values. Every model gives a different forecast of the pollution levels. Some of them deal with small amount of data and some others deal with large amount of data. A model's efficiency can be measured using performance metrics associated with it. The comparison of performance metrics helps to find out the optimal one.

References

1. Bhatti UA et al (2021) Time series analysis and forecasting of air pollution particulate matter (PM_{2.5}): an SARIMA and factor analysis approach. *IEEE Access* 9:41019–41031
2. Choudhary MP, Garg V (2013) Causes, consequences and control of air pollution. In: All India seminar on methodologies for air pollution control, held at MNIT
3. Caraka RE, Chen RC, Toharudin T, Pardamean B, Yasin H, Wu SH (2019) Prediction of status particulate matter 2.5 using state markov chain stochastic process and HYBRID VAR-NN-PSO. *IEEE Access* 7:161654–161665
4. Huang L, Feng H, Le Y (2019) Finding water quality trend patterns using time series clustering: a case study. In: 2019 IEEE fourth international conference on data science in cyberspace (DSC), pp330–337
5. Oke SA (2004) On the environmental pollution problem: a review. *J Environ Eng Lands Manage* 12(3):108–113
6. Gulliver J, Morley D, Vienneau D, Fabbri F, Bell M, Goodman P, Beevers S, Dajnak D, Kelly FJ, Fecht D (2015) Development of an open-source road traffic noise model for exposure assessment. *Environ Model Softw* 74:183–193. ISSN 1364-8152
7. Tao Q, Liu F, Li Y, Sidorov D (2019) Air pollution forecasting using a deep learning model based on 1D ConvNets and bidirectional GRU. *IEEE Access* 7:76690–76698
8. Aditya CR, Deshmukh C, Nayana K, Gandhi P, Astu V (2018) Detection and prediction of air pollution using machine learning models. *Int J Eng Trend Technol* 59:204–207. <https://doi.org/10.14445/22315381/IJETT-V59P238>
9. Du S, Li T, Yang Y, Horng S-J (2021) Deep air quality forecasting using hybrid deep learning framework. *IEEE Trans Knowl Data Eng* 33:2412–2424
10. Ramana Reddy M (2020) IoT based air and sound pollution monitoring system using machine learning algorithms. *J ISMAC* 2(1):1–25. <https://irojournals.com/iroismac/>
11. Tan G, Yan J, Gao C, Yang S (2012) Prediction of water quality time series data based on least squares support vector machine. *Procedia Eng* 31:1194–1199. ISSN 1877-7058
12. Gulliver J, Morley D, Vienneau D, Fabbri F, Bell M, Goodman P, Beevers S, Dajnak D, Kelly FJ, Fecht D (2015) Development of an open-source road traffic noise model for exposure assessment. Received 1 Aug 2014, Revised 18 Dec 2014, Accepted 21 Dec 2014, Available online 8 Jan 2015, Version of Record 22 Oct 2015
13. Zhang X, Zhao M, Dong R (2020) Time-series prediction of environmental noise for urban IoT based on long short-term memory recurrent neural network. *Appl Sci* 10(3):1144
14. Morley D, de Hoogh K, Fecht D, Fabbri F, Bell M, Goodman P, Elliott P, Hodgson S, Hansell AL, Gulliver J (2015) International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. *Environmental pollution (Barking, Essex: 1987)* 206:332–341. <https://doi.org/10.1016/j.envpol.2015.07.031>
15. Wan L, Li YC (2018) Time series trend analysis and prediction of water quality in a managed canal system, Florida (USA). Published under licence by IOP Publishing Ltd. IOP conference series: earth and environmental science, The 4th international conference on water resource and environment (WRE 2018), 17–21 July 2018, Kaohsiung City, Taiwan, vol 191, p 012013

Utilization of Waste Banks System in the COVID-19 Pandemic



Riyan Pradanang, Bima Aditya, Adhitya Rayhan, Michael Adrian Winarlie,
Ford Lumban Gaol, and Tokuro Matsuo

Abstract People's perspectives and behaviors altered in a variety of ways during the pandemic period, most notably in areas related to health, environment, and most notably waste management. Many people still prefer to work from home and this in turn leads to a large accumulation of garbage without having any proper means of disposal. As a result, the garbage area will serve as a breeding ground for illness. Garbage is usually composed of a residual waste generated from a product or object that has outlived its useful life, but can still be recycled into a valuable item. On this basis, the solution is to use a waste bank; it is a collection point for sorted garbage. It is claimed that using a waste bank can help in preventing the garbage accumulation. The main purpose of this study is to determine the contribution of the waste bank to end garbage accumulation. The information gathered through a review of the literature and direct interviews with multiple informants. According to the findings of the literature review, waste banks can provide a variety of benefits to users, as well as the community and surrounding environment. The proposed research study's findings show that the changes in community activity patterns contribute to an increase in garbage, and that the community plays an important role in waste

R. Pradanang · B. Aditya · A. Rayhan · M. A. Winarlie
School of Information System, Bina Nusantara University, Alam Sutera, Tangerang, Indonesia
e-mail: Riyan.Pradanang@binus.ac.id

B. Aditya
e-mail: Bima.Aditya@binus.ac.id

A. Rayhan
e-mail: Adhitya.Rayhan@binus.ac.id

M. A. Winarlie
e-mail: Michael.Winarlie@binus.ac.id

F. L. Gaol (✉)
Computer Science, Binus Graduate Program, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aait.ac.jp

control. From the performed research study, it is evident that many individuals are interested about garbage banks and their benefits based on this presumption.

Keywords Waste bank · Recycling · Reuse

1 Introduction

In recent times, people continue to coexist with COVID-19, which referred to as new normal. According to the Indonesian government, new normal entails the establishment of a new system in order to comply with COVID-19 [1, 2]. According to official data, the total number of COVID-19 cases in Indonesia is 4,272,421 out of which 8775 positive cases are still active. Furthermore, the government announced recently that on an average 598 patients rehabilitated every day. As a result, the total number of COVID-19 cases found is 4,119,472 [3]. As evidenced by this data, we are continuing to collaborate with current COVID-19 epidemic, known as the new normal (LBM Eijkman). To combat the current epidemic situation, people should incorporate three new habits in their lifestyle, these habits are generally termed as 3 M (wearing masks, maintaining social distance, and meticulously washing hands), and they should maintain environmental hygiene during the new normal period. Ten trials with two viruses (SARS-Co-2 and SARS-CoV-1) conducted in five different conditions (aerosol, stainless steel, plastic, copper, and cardboard) [4–7].

SARS-CoV-2 remained in an aerosol state throughout the experiment (3 h). These are similar to the circumstances encountered with SARS-CoV-1 [8–10]. According to the findings of this study, the coronavirus has the chance of transmission through solid media such as waste and air. Garbage management in the community is now difficult due to a public concern on controlling the waste, and this could become a source of coronavirus transmission [7, 11–13]. People are no longer collecting and depositing dry garbage, rather preferring to dispose it in temporary trashcans due to coronavirus [14]. The results of data obtained from BPS regarding the classification of waste sources in DKI Jakarta is as shown in Fig. 1.

Figure 1 shows that the majority of waste in Jakarta comes from households, which has recorded 37.33% proportion as the highest compared to other sources during past year.

Many waste banks actively used in DKI Jakarta in this new normal era are scattered in several areas as follows:

According to Fig. 2, five colors used: blue for North Jakarta, green for Central Jakarta, purple for West Jakarta, yellow for South Jakarta, and orange for South Jakarta (East Jakarta).

The primary goal of establishing a waste bank is to assist in garbage processing; the secondary objective is to promote public awareness on the importance of maintaining a healthy, orderly, and clean environment. Furthermore, waste banks are established to convert garbage into something more useful to society, such as handicrafts and monetary-valued fertilizers. Garbage cans benefit both humans and environment

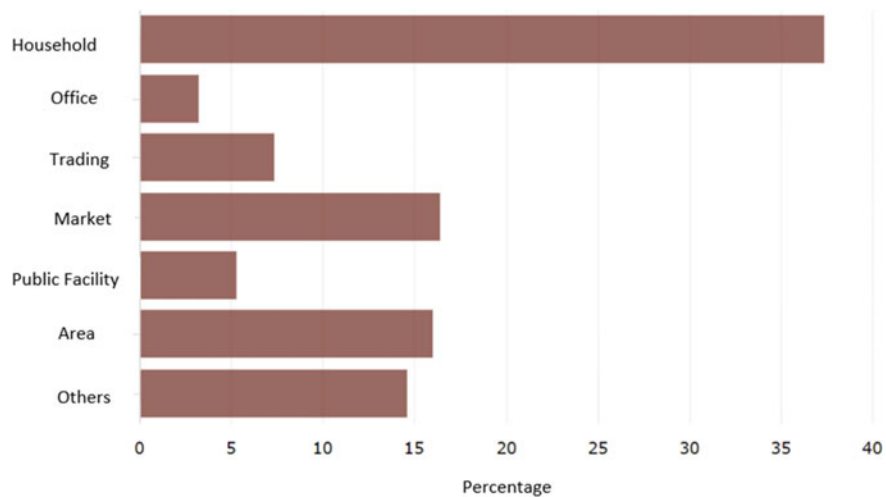


Fig. 1 Source Central Bureau of Statistics (BPS) [15]

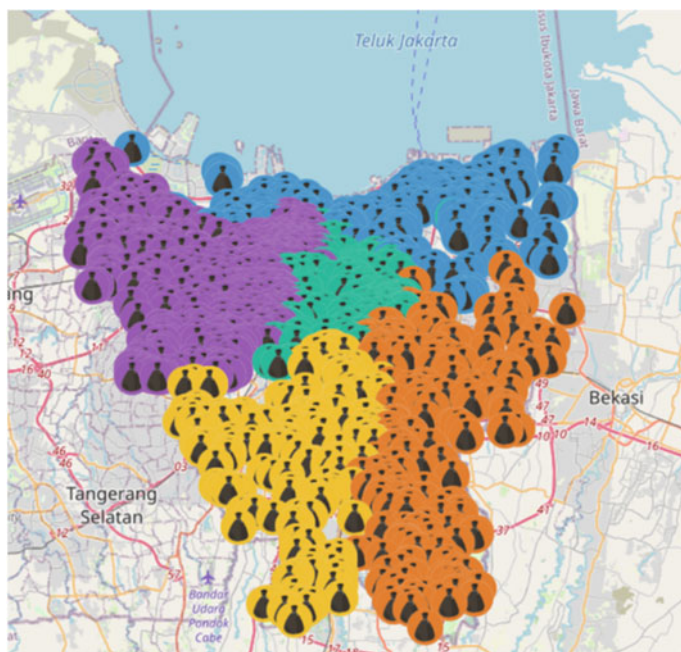


Fig. 2 Source Jakarta Provincial Environmental Service [16]

in different ways, including cleaner environments, increased hygiene awareness, and waste conversion into economically valuable items. A waste bank benefits the community by increasing people's income because when they trade their junk, they receive a reward in the form of money in their bank account. When individuals have a sizable sum of money in their savings, they have the option of withdrawing it at any time. Other than money, some of the incentives provided to savers are in the form of staples such as sugar, soap, oil, and rice. Garbage banks are also beneficial for financially challenged people; some schools have included waste payment as a manner of paying school fees.

The main purposes of this research is to establish a waste bank to raise the level of the Indonesian economy, especially for the lower middle class by means of which people can sell the waste they have collected for money. In addition, the surrounding environment becomes cleaner and well maintained.

Research question:

1. What is the impact of changes in community behavior that affect waste in this new normal era?
2. What type of garbage accumulates the most?
3. Are people familiar with the waste bank?
4. Do people know what the benefits of a waste bank are?
5. How can a waste bank affect people's lifestyle in managing waste?

Waste bank is a banking system concept in which the saving is not money but waste that has already been collected and sorted. Customers, also known as savers, are given a bankbook in which they can borrow money, which is then returned with waste. The customer's waste is then weighed and valued; the waste will then be sold to manufacturers or recycling agencies or it can be sent to local recycling and reuse agents for processing [17].

2 Research Technique and Method

The proposed study has conducted a qualitative research to investigate the significance of waste banks in Indonesia. By conducting direct interviews with informants involved in the distribution of waste bank in Indonesia, it is highly anticipated that a direct investigation of the field will highly assist in obtaining the required information in an accurate and useful manner.

The research methodology implemented here is by directly interviewing the informants with questions marked on paper or phones, and then all activities between the interviewer and informant are recorded by using a video recorder on the phone, particularly during the interview. We will select participants between the ages of 17 and 50 who live in Jakarta. After gathering all of the necessary information, the final step in our research is to analyze all the available data.

The interview questions asked to the participants:

1. Do you know what a waste bank is?
2. Do you know what the benefit of a waste bank is?
3. What is the type of waste that your household throw the most?
4. Do you believe that a pandemic situation that changes community behavior has an impact on waste?
5. Do you think that the waste bank can help people manage their waste?
6. Do you think the community needs a waste bank, especially during the pandemic?

3 Research Result

According to Fig. 3, we can learn about the interview results: 78% or 36 people from total population knows about the waste bank, while 22% or 10 people do not.

Based from the interview result as in Fig. 4, from 46 participant 54% or 25 people know about the benefit of a waste bank and 46% or 21 people does not know about the benefit of a waste bank.

3. What is the type of waste that your household throw the most?

Based from the interview result, from 46 participant, the waste that they throw the most are 25 people answered plastic, 2 people answered organic, 16 people answered paper, 2 people answered glass, and 1 people answered metal. The detail of types of household throw the most was presented on Fig. 5.

Do you know what is a waste bank?

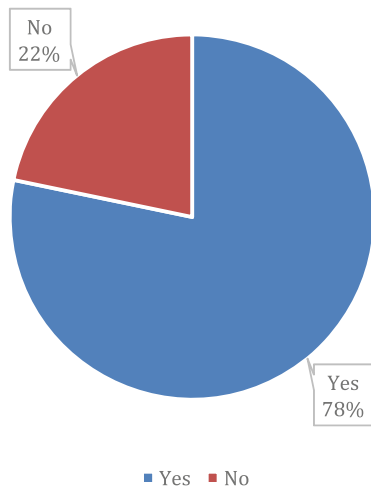


Fig. 3 About waste bank

Do you know what is the benefit of a waste bank?

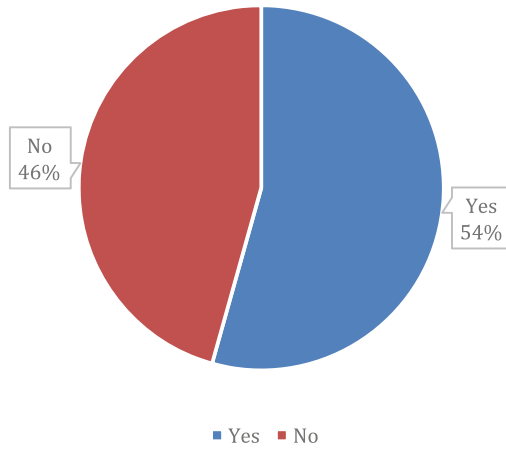


Fig. 4 Proportions about the benefit of a waste bank?

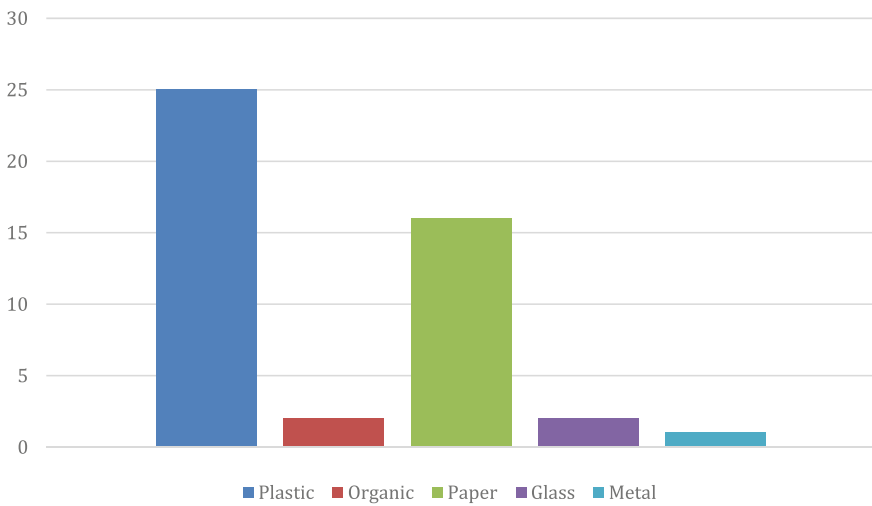


Fig. 5 Types of household throw the most

4. Do you believe that a pandemic situation that changes community behavior has an impact on waste?

Based on the interview result, from a total of 46 participants 87% or 40 people think that the waste is affected by pandemic situation that change community behavior and 13% or 6 people think that the waste is not affected by pandemic situation that change community behavior. The detail of distribution of result of the impact of changes community behavior has an impact on waste was shown on Fig. 6.

Do you think that the waste is affected by pandemic situation that change community behaviour?

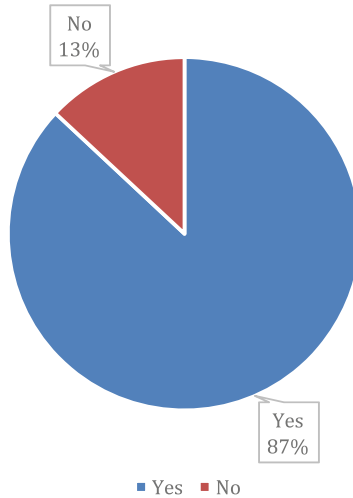


Fig. 6 Impact of changes community behavior has an impact on waste

5. Do you think that waste bank can help people manage their waste?

According to the interview results that shown on Fig. 7, 43% or 20 of the 46 participants believe that the waste bank can help people manage their waste, while 57% or 26 believe that the waste bank cannot help people manage their waste.

6. Do you think the community needs a waste bank, especially during the pandemic?

According to the interview results as presented on Fig. 8, 72% or 33 of the 46 participants believe that a waste bank is required in the community, particularly during the pandemic, while 28% or 13 believe that a waste bank is not required in the community, particularly during the pandemic.

4 Discussion

This research study discusses about the importance of waste bank in a pandemic situation based on the survey and research results. Based on the results of the interview, 78% of people are already aware of the waste bank. However, 22% of them are still unaware of the waste bank. The results also show that 46% of people are still unaware of the benefits of a waste bank, despite the fact that the number of people who are aware of the waste bank is still quite large.

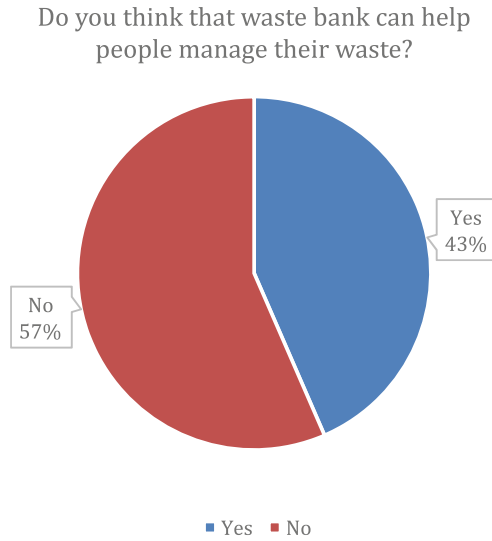


Fig. 7 Result whether the waste bank can help people manage their waste

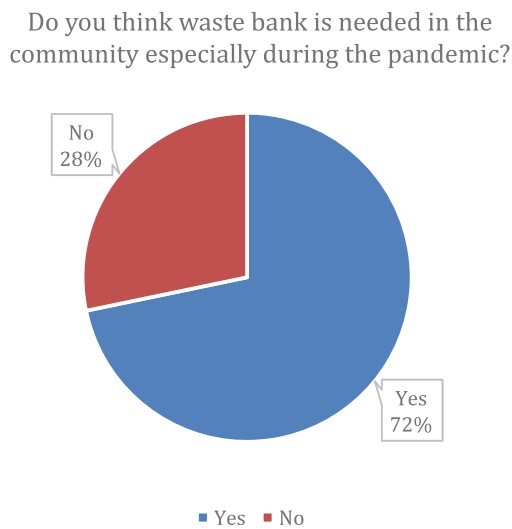


Fig. 8 Response whether the community needs a waste bank, especially during the pandemic

5 Conclusion

Waste bank is a banking system concept in which the saving is waste that has already been collected and sorted rather than money. Customers, also known as savers, are given a bankbook with which they can borrow money, which is then returned as

waste. The weighed and valued customers' waste will be sent to manufacturers or recycling agencies for processing, or it can be sent to local recycling and reuse agents [17]. Waste banks can help the community by exchanging waste for money. Based on the findings, it is possible to conclude:

1. What is the impact of changes in community behavior that affect waste in this new normal era?

The impact is people will be more diligent in collecting garbage and reducing littering.

2. What type of garbage accumulates the most?

According to the interview results, out of 46 participants, 25 people answered plastic, 2 people answered organic, 16 people answered paper, 2 people answered glass, and 1 person answered metal. As a result, plastic waste considered as the most accumulated waste.

3. Are people familiar with the waste bank?

Based on the interview result, out of 46 participants, 78% or 36 people know about waste bank and 22% or 10 people does not know about waste bank.

4. Do people know what the benefits of a waste bank are?

According to the interview results, from 46 participants 54% or 25 people know about the benefit of a waste bank and 46% or 21 people does not know about the benefit of a waste bank.

5. How can a waste bank affect people's lifestyles in managing waste?

According to the interview results, 43% or 20 of the 46 participants believe that the waste bank can help people manage their waste, while 57% or 26 participants believe that the waste bank cannot help people manage their waste. The waste bank itself provides a benefit where people can exchange their waste for money. As a result, the community will become more meticulous in collecting the waste and reducing littering.

Acknowledgements We would like to express our appreciation and gratitude with the RTTO Binus University with the grant no contract number: 017/VR.RTT/III/2022.

References

1. Dinas Lingkungan Hidup Provinsi DKI Jakarta (2021) Volume Sampah yang Terangkut per Hari Menurut Jenis Sampah di Provinsi DKI Jakarta (Ton), 2018–2019. <https://jakarta.bps.go.id/statictable/2021/09/21/306/volume-sampah-yang-terangkut-per-hari-menurut-jenis-sampah-di-provinsi-dki-jakarta-ton-2018-2019.html>
2. Dinas Lingkungan Hidup (2020) Apa Itu Bank Sampah !!! Bulelengkab.go.id
3. <https://covid19.go.id/>
4. Bulelengkab.go.id (2020) APA ITU Bank Sampah !!! | Dinas Lingkungan Hidup [online]. Available at: <https://dlh.bulelengkab.go.id/informasi/detail/artikel/apa-itu-bank-sampah-26>

5. Tim Redaksi Katadata (2020) Banjir Sampah Plastik Selama Pandemi - Analisis Data Kata-data [online]. Katadata.co.id. Available at: <https://katadata.co.id/timredaksikatadata/analisisdata/5fc719de77307/banjir-sampah-plastik-selama-pandemi>
6. Widyastuti S, Sutrisno J, Wiyarno Y (2021) Pengelolaan Sampah di Masa Pandemi. *Jurnal Penamas Adi Buana*, [online] 5(01):79–88. Available at: <http://jurnal.unipasby.ac.id/index.php/penamas/article/view/3583>
7. Rohman A (2021) Peran Bank Sampah di Masa Pandemi Covid-19—Program Studi Teknik Lingkungan [online]. Program Studi Teknik Lingkungan. Available at: <https://enviro.teknik.unej.ac.id/peran-bank-sampah-di-masa-pandemi-covid-19/>
8. Selomo M, Sampah B, Salah S, Solusi S, Sampah P, Makassar K, Bintara Birawida A, Mallongi A, Bagian M, Lingkungan K, Kesehatan F, Unhas M (n.d.) The Waste Bank is one of good solution for handling waste in Makassar City. [online] Available at: <https://media.neliti.com/media/publications/212842-bank-sampah-sebagai-salah-satu-solusi-pe.pdf>
9. Anand KB, Karade S, Sen S, Gupta RM (2020) SARS-CoV-2: camazotz's curse. *Med J Armed Forces India* 76:136–141. <https://doi.org/10.1016/j.mjafi.2020.04.008>
10. Kannan S, Gheena S, Lakshmanan G (2020) Effect of social media on human behaviour: a survey. *Int J Pharmaceut Res* 12:1596–1603
11. Fadilah A (2018) Bank waste as management of household rubber based on community. *Empow J Ilmiah Program Studi Pendidikan Luar Sekolah* 7(2):1–9. Available at: <http://e-journal.stkipsiliwangi.ac.id/index.php/empowerment/article/view/724>
12. Wulandari D, Hadi Utomo S, Narmaditya BS (2017) Waste Bank: waste management model in improving local economy. *Int J Energy Econ Policy* [online] 7(3):36–41. Available at: <https://www.econjournals.com/index.php/ijeep/article/view/4496>
13. Suardi LR, Gunawan B, Arifin M, Iskandar J (2018) A review of solid waste management in waste bank activity problems. *Int J Environ Agric Biotechnol* 3(4):1518–1526
14. Putri CA (2020) Kepala LBM Eijkman Buka-bukaan Soal New Normal Hingga Vaksin. <https://www.cnbcindonesia.com/profil/20200528112650-41-161452/kepala-lbm-eijkman-buka-bukaan-soal-new-normal-hingga-vaksin>
15. <https://www.bps.go.id/>
16. Dinas Lingkungan Hidup Prov. DKI Jakarta (2022) Bank Sampah. https://silika.jakarta.go.id/bank_sampah
17. Amanda Bahraini (2020) Waste Bank to Support Indonesia Clean-from-Waste 2025. <https://waste4change.com/blog/waste-bank-to-support-indonesia-clean-from-waste-2025/>

None Shall Pass: A Blockchain-Based Federated Identity Management System



Shlok Gilda , Tanvi Jain, and Aashish Dhalla

Abstract Authentication and authorization of a user's identity are generally done by the service providers or identity providers. However, these centralized systems limit the user's control of their own identity and are prone to massive data leaks due to their centralized nature. We propose a blockchain-based identity management system to authenticate and authorize users using attribute-based access control policies and privacy-preserving algorithms and finally returning the control of a user's identity to the user. Our proposed system would use a private blockchain, which would store the re-certification events and data access and authorization requests for users' identities in a secure, verifiable manner, thus ensuring the integrity of the data. This paper suggests a mechanism to digitize documents such as passports, driving licenses, and electricity bills, issued by any government authority or other authority in an immutable and secure manner. The data owners are responsible for authenticating and propagating the users' identities as and when needed using the OpenID Connect protocol to enable single sign-on. We use advanced cryptographic algorithms to provide pseudonyms to the users, thus ensuring their privacy. These algorithms also ensure the auditability of transactions as and when required. Our proposed system helps in mitigating some of the issues in the recent privacy debates. The project finds its applications in citizen transfers, inter-country service providence, banks, ownership transfer, etc. The generic framework can also be extended to a consortium of banks, hospitals, etc.

Keywords Blockchain · Identity management · Authorization · Authentication

S. Gilda (✉) · T. Jain · A. Dhalla
Department of Computer and Information Science and Engineering, University of Florida,
Gainesville, FL 32608, USA
e-mail: shlogilda@ufl.edu

T. Jain
e-mail: tjain@ufl.edu

A. Dhalla
e-mail: aashishdhalla@ufl.edu

1 Introduction

Identity providers or service providers act as third-party operators that authenticate and authorize the user's identity in an identity management system, and traditional systems are neither secure nor reliable. Password manager and single-sign-on provider, OneLogin was hacked, and all customers served by its US data center were affected.¹ Equifax, one of the most significant credit bureaus in the USA, was the victim of a data breach. Private records of approximately 148 million Americans were compromised in the breach, making it one of the largest cybercrimes related to identity theft [7]. Even though these digital identities are portable, they are not secure. Several industries suffer from the problems of current identity management systems. The lack of interoperability between government departments and government levels takes a toll in the form of excessive bureaucracy. This, in turn, increases processes times and costs. In the field of education, there is a massive problem of fake academic certificates.² During the registration phase, user identity information is stored in a central database (CD). As CD is compromised, user identity information can be leaked. Also, user identity and passwords are stored on CD for the authentication process, which could be compromised. Considering the flaw in such cases, we aim to utilize decentralized blockchain technology to authenticate and authorize users using privacy-preserving attributes and return control of a user's identity to the user.

Because of its decentralized structure, visibility, traceability, individual control of data in blockchain [33, 47], it has been introduced in a variety of applications, including supply chain [16, 38], health care [14, 45], and Industrial Internet of Things (IIoT) [10, 43, 46]. Miners pack the transaction into a package in a blockchain system, and a consensus algorithm is used to verify the block construction in a network. After verifying a transaction, a legitimate block is added at the end of the longest chain of verified blocks—each node on the blockchain stores a copy of the distributed ledger. Using the blockchain in the identity management system solves traditional schemes that rely on trusted third party [18, 31]. In a standard blockchain system, data are stored as plain text on the network in some schemes [2, 11, 23], violating the privacy-preserving requirement. Our proposed architecture supports attribute-based access control. The blockchain holds re-certification events of the users' identities and the data access and authorization request instead of storing the complete user identity information. These requests stored on the blockchain provide auditability and traceability of the transactions.

In the traditional identity management systems, users give away too much information as the complete identity is shared with the service providers. We want to allow the users to gain control over their identity by choosing what attributes of their identities they want to share to access the services; this can be achieved by attribute-based access control (ABAC) system [26]. ABAC is globally recognized and allows or denies user requests based on any user and object's properties and environmen-

¹ <https://www.zdnet.com/article/one-login-hit-by-data-breach-exposing-sensitive-customer-data/>.

² <https://www.fox46.com/news/easy-to-get-fake-degrees-creating-real-problems/>.

tal factors that may be more relevant to the rule in question. Our proposed solution allows us to segregate different personas of a user's identity and allow them to decide what attributes and which identities they want to use. Attribute-based access control with OpenID Connect [39] ensures that we can utilize existing technologies to ensure fine-grained access control.

We have designed a private blockchain-based system for identity management. Our system suggests a mechanism that would digitize documents such as passports, driving licenses, and electricity bills, issued by any government authority or other authority in an immutable and secure manner. Since these authorities already have access to slivers of our identity, we will not be storing these identities on the blockchain. Data owners will still be the custodians of user identity, but users will control how they want to use their identity attributes. Putting personal information on the ledger jeopardizes users' privacy and violates current privacy laws (GDPR, right to be forgotten, etc.) A user's identification traits are also dynamic. Thus, we would associate users' identities on the blockchain with the help of pseudo-anonymous identifiers. The data owners are responsible for authenticating and propagating the users' identities as and when needed. An identity provider could communicate with the consortium of authorities to validate the users' claims regarding their identity.

Our contributions can be summarized as follows:

- 1. We propose a blockchain-based identity management system that enables easy data authentication and data authorization using multiple data sources. Our proposed system emulates real-world conditions where the data owner owns identities. The problem of multiple authentications required by different service providers can be effectively reduced due to our design's federated identity management aspect.**
- 2. A private blockchain ensures the integrity of data. Even though we do not store the actual identities on the blockchain, we store data access requests and identity attribute re-certification events; this improves the transparency and traceability of the system.**
- 3. Our proposed system allows users to access different personas of their identity for availing online services based on specific use-cases. We propose an attribute-based identity management scheme that allows fine-grained access control to identity with a trust score associated with the identity.**
- 4. To ensure user security and confidentiality, we use an identity-based conditional proxy re-encryption scheme that allows transmission of encrypted identities across our system. We also use hierarchical deterministic keys to improve user privacy and increase the traceability and auditability of transactions.**

The proposed system finds its applications in citizen transfers, inter-country service providence, banks, ownership transfer, etc. The generic framework can also be extended to a consortium of banks, hospitals, etc.

The paper is organized as follows. Section 2 summarizes related work. Section 3 gives the overview of the proposed architecture and the various components utilized

in organizing and defining the end-to-end process proposed in this paper. In Sect. 4, we have described the tools and protocols used to draft the system architecture that has been used to ensure the anonymity of users and auditability of transactions. In Sect. 5, we have defined the communication pattern between user identity providers, service providers, and data owners residing on the blockchain and hence giving the complete system architecture information. The impacts of using this system architecture have been discussed in Sect. 6. Section 7 concludes our work, presenting our limitations and proposing future work.

2 Related Work

Earlier, all the user's data were handled by organizations, as the conventional identity management systems mostly depended on third parties. As third parties handle the user's data, there are chances of data being compromised or lost if third parties crumble. Moreover, the organizations can share the data themselves for money or other reasons. The main challenge in centralized systems is transferring users' confidential data from one application to another, as the centralized system is a closed system. There are various proposals given to rectify this problem of centralization. One of the solutions proposed is to use federated identity management [4, 6, 34, 41]. In federated identity management, a user with a single identity can access many applications with a single registration identity. The problem with these schemes is that the user's data are stored in plain text and controlled by a single organization. Hence, more focus is put on the privacy protection of users' data.

The user-centric identity management [5, 29] emphasizes more on the user-oriented paradigm, which allows users to select the data they want to share and, during the authentication process, can present the correct credentials. Laborade et al. [28] provide a decentralized and user-centric identity management scheme; in this case, the usability is increased as it eliminates user passwords and, with regards to privacy and sovereignty, makes the identity more trustworthy. Singh et al. [42] try to create a privacy-aware personal data storage (P-PDS) that can take security mindful choices on third parties to get to demand following user inclinations. Nevertheless, the user-centric approach is a weak model. We can take the example of Facebook Connect; here, Facebook is the sole identity provider. Some blockchain identity management schemes are proposed for a decentralized system. In other fields like the financial field, to not reveal the identity of the parties involved in the transactions, decentralized coin-mixing-based methods are suggested [36, 37, 40].

Similar work in this field has been developed by Hardjono and Shrier [24] where the researchers present the argument that the use of MIT OPAL /Enigma for privacy-preserving data sharing offers a better foundation for future Internet identities based on information about a user. In a similar technical report by Hardonjo et al. [25], the authors address the issue of retaining user anonymity within a permissioned blockchain. The paper presents the ChainAnchor architecture that adds identity and a privacy-preserving layer above the private or the public blockchain. Goyal et al. [20]

discuss the development of a new cryptosystem called key-policy attribute-based encryption (KP-ABE), where ciphertexts are labeled with sets of attributes and private keys are associated with access structures that control which ciphertexts a user can decrypt. Bethencourt et al. [8] extended this work in their paper, which describes a system wherein attributes are used to describe a user's credentials. A party encrypting data determine a policy for who can decrypt, unlike previous attribute-based encryption systems that used attributes to describe the encrypted data and built policies into users' keys.

Early work by Dunphy et al. [13] provides an early glimpse of the current strengths and limitations of applying distributed ledger technology (DLT) to identity management (IDM) using Cameron's evaluative framework. Xu et al. [44] produced the idea of blockchain-based identity management and authentication methods for mobile networks in which users create their SSI identity, and each user will have private and public keys, and the user has control over their identity. Gao et al. [17] came up with a blockchain-based privacy protection and identity authentication scheme by using different algorithms like ECDSA encryption and RSA encryption algorithms. We can reduce the storage as well by using these algorithms. This proposal gives access to users to hash their identity and decide whether to store the identity information on a blockchain or not. Faber et al. [15] gave a high-level architecture and conceptual design for blockchain-based identity management. This proposal emphasizes giving more control to users of their data. Rathor and Agarwal [1] used blockchain in a different area where they created three different modules to give the information to the people looking for a job with the data like experience and education, which helps the organization to verify its employees' records. References [1, 15, 17, 44] can rectify issues like the privacy of user's data and giving the control of user's identity back to a user. However, there is a problem with IP traceability for the IP applications as they need to trace the IP in case of any disputes. Chuxin et al. [48] proposed a blockchain-based privacy-preserving and traceability intellectual property (IP) identity management scheme, in which the user's real identity information is processed into multiple shares using improved-Shamir secret sharing that can reduce storage overhead and achieve privacy protection. Here, user information is stored on the blockchain, thus increasing storage and scalability overheads.

Most of the proposed solutions presented here either store the user's encrypted data on the blockchain or do not account for the auditability of transactions. Also, many of the proposed system architectures assume direct communication between the blockchain and the user. However, this would not be possible in real-world scenarios, leading to remarkably high API response times. Our proposed system builds on the existing federated authentication and authorization systems. This paper proposes an architecture where original data owners are still the custodian of the identity. The user can choose what identities can be used to avail of online services. Our proposed architecture supports attribute-based access control. The blockchain holds re-certification events of the users' identities and the data access and authorization request instead of storing the complete user identity information. These requests stored on the blockchain provide auditability and traceability of the transactions.

This would lead to lower resistance to adopting the system than in prior work. The attribute-based identity management system leads to a crucial criterion to minimize identity theft risk.

3 Preliminaries

The identity of an individual is composed of many facets. Every service provider does not necessarily need to know every detail about a user's identity. However, when we use federated identity providers, we give away so much unnecessary information about ourselves; this leads to massive privacy violations, as discussed before.

The system architecture proposed in this paper has four main components:

1. **Consortium of authorities on a private blockchain:** This could be a consortium of government authorities, credit bureaus, etc. It will also have a querying API server (Communication server) that will be very tightly coupled to the consortium, allowing an identity provider to query and verify the identity attributes. The communication server is also a node on the blockchain, with some additional responsibilities, including communicating with the outside world.
2. **Identity provider (IDP):** This module deals with user authentication and authorization. It would use OpenID Connect Protocol (OIDC) to implement single-sign-on functionality. It will essentially act as an OIDC provider.
3. **Service provider (SP):** This is essentially any third-part service that a user would use. They act as OIDC relying parties.
4. **User:** Users can have different privacy-preserving personas, and in our proposed system, users have the liberty to control the persona that they want to use to interact with IDP and hence giving control back to the user to decide the facet they want to use.

The consortium of authorities and the individual would be data owners, whereas the service providers would be data consumers. Our proposed system allows individuals to gain back control of their digital identities. Our system uses decentralized public identifiers and associated descriptor objects to associate user identities on the blockchain. This would ensure that we do not store any personally identifiable information (PII) on the blockchain and always refer to them using pseudo-identifiers. This is critical because a distributed ledger is immutable, meaning anything added to the ledger cannot be changed or removed. As a result, no personal information should ever be recorded in the ledger. The identity providers would communicate with a server associated with the consortium to get details about the credentials as provided by the user. The identity provider would get user consent to transfer/verify specific attributes of the users' identity to the service provider. We recommend using OpenID Connect since it is the industry standard in federated authentication.

Figure 1 gives an overview of the proposed architecture. Netflix, Bank of Timbuc-too, etc., are examples of service providers. "Identity Management" is the identity

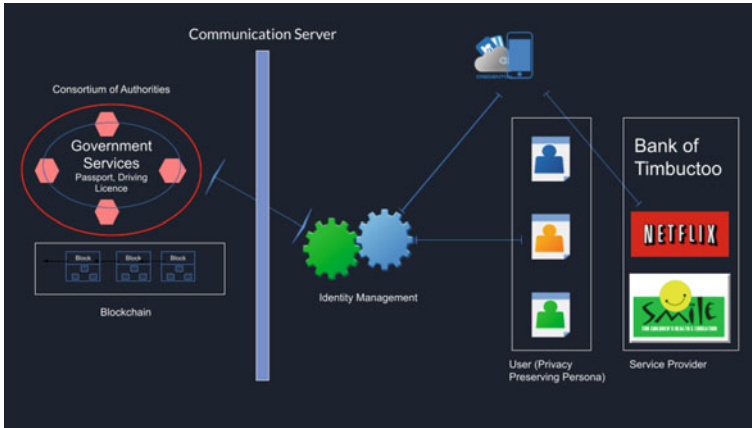


Fig. 1 Architecture overview

provider in charge of communicating with the consortiums. Our proposed system could have multiple consortiums of identity sources. This would allow the user to control which identity consortium to use depending on the specific use case. For example, a service provider like Netflix might not necessarily need a high trustworthy source of information, so the user could use their social media networks to provide details about specific identity attributes. However, the Bank of Timbuctoo would need very high attestation that the users' identity attributes are correct. In this scenario, the user could verify their identity by the government authorities consortium. This segregation of identity into various slivers gives users more control over their own identity. Additionally, users can decide the privacy-preserving persona they want to use to interact with the identity provider. The blockchain layer stores the re-certification events of the identities and data access and authorization requests, providing auditability and traceability of the transactions.

4 Tools and Protocols

This section discusses the different tools and protocols to design our identity management system. We suggest using hyperledger fabric, a permissioned blockchain for storing user transactions. A permissioned blockchain is not publicly accessible and can be accessed only by users with permissions; this provides an additional level of security over permissionless blockchain systems such as Bitcoin or Ethereum, as they require an access control layer. Next, we talk about OpenID Connect, an open and trusted authentication protocol that adds an identity layer to OAuth 2.0. Clients can utilize OIDC to verify an end-user's identity utilizing authorization server authentication. We look into the various cryptographic protocols that our design would use

to ensure confidentiality of data, transparency of transactions, and guarantee privacy, integrity, and non-repudiation. Lastly, we will look at a mechanism for generating attribute trust scores. An identity provider computes a time-varying confidence score for an asserted attribute and includes it in a security assertion returned to a service provider.

4.1 *Hyperledger Fabric*

Hyperledger fabric [3] is a distributed ledger platform that is open, tested, and enterprise ready. It is a permissioned ledger with advanced privacy restrictions, which means that only the data you want to be shared is shared among the “permissioned” (known) network participants. Fabric allows establishing decentralized trust in a network of known participants rather than an open network of anonymous participants. The transactions are confidential and only shared with parties that have permission to read those transactions. The transactions are kept private and are only shared with parties who have been granted permission to read them. A ledger resides within the scope of a channel; it can be shared across the entire network (assuming all participants are using the same shared channel), or it can be privatized to only include a limited number of participants. As a result, fabric is an ideal option for a decentralized but trusted ledger that protects user privacy.

Hyperledger fabric’s modular architecture divides the transaction processing workflow into three stages: smart contracts, also known as chaincode, consisting of the systems’ distributed logic processing and agreement, transaction sequencing, and transaction validation and commitment. This separation has many advantages:

- A reduced number of trust levels and verification that keeps the network and processing clutter-free.
- Improved network scalability.
- Better overall performance.

We suggest using hyperledger fabric to build a consortium of blockchains; this would allow similar entities to be a part of the same consortium. For example, a consortium of federal government entities (Passport authority, DMV, etc.) could be separate from a consortium of credit bureaus (Experian, Equifax, etc.). Having separate consortiums of similar entities would allow these entities to open private channels among them as and when needed, thus allowing more accessible communication between them while allowing the user to be wary of these communications.

4.2 *Cryptographic Protocols*

This section will discuss the various cryptographic protocols and primitives that our proposed system would use. These protocols ensure data privacy and integrity while

also guaranteeing non-repudiation of transactions. Our proposed system uses hierarchical deterministic wallet to generate user keys and identity-based conditional proxy re-encryption to securely store and transmit users' personal information throughout the system.

Hierarchical Deterministic Wallet Bitcoin and its derivatives use a feature known as hierarchical deterministic wallets (HD Wallets) that cause your receiving address to change after being used; this is done by creating a “master” key known as the extended private key and extended public key. This feature enhances the privacy as well as security of the users. The extended private key is the base from which all of your addresses' private keys are derived. Alternatively, in other words: The extended private key is the master key to all the private keys belonging to an account. Bitcoin improvement protocol 32 (BIP-32)³ describes the specification intended to set a standard for deterministic wallets. BIP-32 uses elliptic curve cryptography using the field and curve parameters defined by secp256k1.⁴

Figure 2 showcases how BIP-32 generates keys hierarchically. In our proposed system architecture, the user would hold the master key. The user would then distribute the derived keys as and when needed to relevant authorities/entities in the system. The user would generate two master keys in this entire process: One of them would be associated with all communications between the data owners. In contrast, the other would be associated with authentication with the identity provider. We describe the entire key generation and key distribution process in Sect. 4.5.

Since the user would have access to the master key, they would be able to monitor all transactions (in this case, data access authorization approvals). In contrast, the authorities would only be able to monitor transactions from the keys they have been given access. The changing keys have two significant advantages: increased privacy and increased security. By having multiple keys, the user would be able to segregate their identities on the different consortiums of blockchains, enhancing user privacy. Also, since the user has different keys across different data owners, an adversary would need to get multiple private keys to access all the different identities of the user.

Identity-Based Conditional Proxy Re-encryption Identity-based conditional proxy re-encryption (IBCPRE) is an extension of proxy re-encryption. It provides conditional proxy encryption and extends the proxy re-encryption concept to the identity-based public-key cryptography setup. By using conditional proxy re-encryption, a proxy can re-encrypt a ciphertext using an IBCPRE scheme. However, if a condition is given to the ciphertext and the re-encryption is satisfied, the ciphertext will be well-formed for decryption; this permits fine-grained proxy re-encryption and can be helpful for applications such as safe sharing via encrypted storage.

IBCPRE allows users to choose recipients of a message even after encrypting their messages and uploading them to the server. IBCPRE supports end-to-end as well as one-to-many encryption. The IBCPRE's conditional “tag” allows fine-grained

³ <https://github.com/bitcoin/bips/blob/master/bip-0032.mediawiki>.

⁴ <http://www.secg.org/sec2-v2.pdf>.

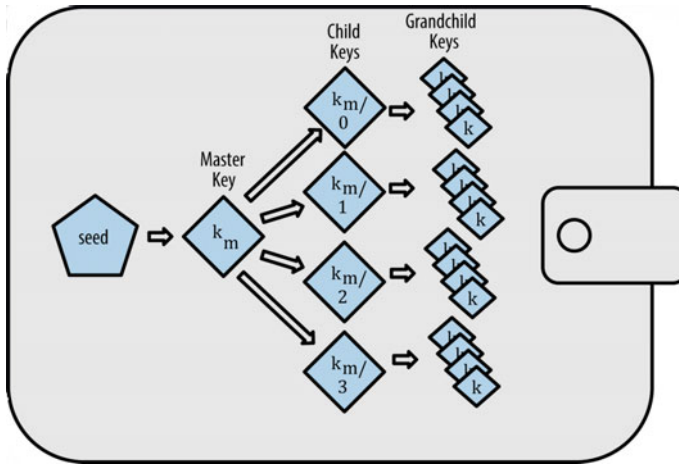


Fig. 2 Key generation scheme [27]

access to encrypted messages. The data owner can control the exact set of encrypted communications they want to share with any particular system entity by assigning different tag values to distinct encrypted messages.

Liang et al. [30] was the first to propose a single-hop identity-based conditional proxy re-encryption that was proved secure against adaptive conditions and adaptive identity chosen-ciphertext attacks in the standard model. We adopt this scheme for our proposed system and explain it further in Sect. 4.5.

4.3 Trust Score

Gokhale et al. [19] suggested the use of a time-varying trust score for an asserted attribute value. The identity provider would compute a time-varying confidence score for an asserted identity attribute and include it in a security assertion returned to a service provider. The confidence score typically “deteriorates” (i.e., decreases over time). One or more qualifying attribute verification events, on the other hand, may affect the degree to which the score deteriorates.

In our proposed system, we aim to use the identity attribute scoring framework [19] to assign trust scores to different attributes. The identity provider would get the verified attribute(s) from the data owner(s). Based on the type of data owners and attributes, the identity provider would generate a trust score and return the said trust score in a security assertion to the service provider. If multiple data owners verify the identity attribute, the trust score could be higher. The identity provider could also profile the service providers and determine the threshold values for an attributes’ trust score that the service provider deems acceptable. Based on this profiling, the

identity provider could recommend data sources that might be useful for achieving the threshold for a specific attribute. The trust score generation methodology could be proprietary to an identity provider, thus incentivizing multiple identity providers to be a part of the system.

We explain the necessity of a trust score with a few specific identity attributes:

- **Name:** Governmental data sources would be a more reliable source of information for truthful data about the attribute “Name” rather than a social media platform or a delivery service platform. However, suppose the service provider does not need a high trust score for this attribute. In that case, the user could choose social media platforms as a source of data information to satisfy the lower threshold of the service provider.
- **Address:** A delivery service provider might have more recent addresses as compared to a government document. Thus, sourcing “address” from a delivery service provider might result in a higher trust score than a governmental data agency. However, a recent re-certification of this attribute with a governmental agency might result in a better trust score. Thus, it is up to the user to choose relevant data owners to verify their identities based on the specific requirements of the service provider.

Our proposed system allows service providers to be more legally compliant. For example, Facebook requires that its users be at least 13 years of age to be able to use their platform;⁵ however, they currently have no way of verifying this. Our system would let them verify the veracity of the “minimum age” claim without compromising the user’s identity.

4.4 OpenID Connect Protocol

OpenID Connect (OIDC) is an open authentication protocol that adds an identity layer to OAuth 2.0 by profiling and extending it [39]. Clients can utilize OIDC to verify an end-user’s identity utilizing authorization server authentication. By layering OIDC on top of OAuth 2.0, a single framework is created that promises to protect APIs, mobile native apps, and browser apps in a single, unified design.

Our system would use OpenID Connect for user authentication and identity attribute authorization. OIDC allows service providers to verify the end-user’s identity based on the authentication performed by an identity provider and obtain basic profile information about the end user in an interoperable and REST-like manner. The identity provider obtains the user’s basic profile information from the data owners. After obtaining this information, it calculates the trust scores for the requested identity attributes and returns those values with the identity assertions.

⁵ <https://about.fb.com/news/2021/07/age-verification/>.

4.5 Key Generation

This section describes the various keys and secrets being used throughout our system. We will look into the key generation, ownership, and distribution process. We assume that the communication channels between all the entities are secure. We also assume that the user knows the public keys required for IBCPRE of the identity provider and data owners.

Hierarchical Deterministic Keys The user uses the general idea mentioned in BIP-32 to generate two private keys: “Data Access Key” and “Data Authorization Key.” The data access key is associated with the data owners, and the data authorization key is associated with the identity provider. Having two separate master keys, data access key and data authorization key, allow us to keep the data owners’ information separate from the identity providers. The user generates these two keys and keeps them private and safe; leaking a private key would mean a loss of privacy. A “Child Key Derivation Function” computes a child key given a parent key and an index i . Modifying i , we can get new child keys.⁶

The user could generate child private keys for individual data sources from the data access key. These are referred to as “Data Owner Keys.” These data owner keys are then registered with every data owner, either during the creation of the identity or later. The data owners have access to the keys they have and any subsequent keys that the data owners derive from their keys. The data owners can derive further keys, which can then be used to associate transactions about a user on the blockchain. That ensures some level of anonymity on the blockchain. Since these data owner keys are derived from the data access key, the user would have access to all subsequent keys that the data owner derives further. Figure 3 explains the key generation process of the data access key.

Our system allows for multiple identity providers. We can derive multiple “Identity Owner Keys” from the data authorization key. These identity owner keys can be used to signup with identity providers and be later used for authentication with the identity provider. Figure 4 depicts the key generation and ownership process. Section 5.4 explains the key distribution process.

There are a few shortcomings in the original BIP-32 protocol. Given the master public key and any child private key, an attacker can easily extract the master private key. To tackle this, Gutoski and Stebila [22] came up with a better version of HD wallets that is not vulnerable. Their proposed system can handle the leakage of up to m private keys with a master public-key size of $O(m)$. In order to improve the security of BIP-32, Das et al. [12] came up with a minor modification in the key derivation process of ECDSA. They suggested switching re-randomization in BIP-32 from additive to multiplicative to achieve tighter security without additional costs. They observed that BIP-32 gives roughly 92 bits of security based on their theorems and a conservative choice of parameters; however, the multiplicative version of ECDSA

⁶ See footnote 3.

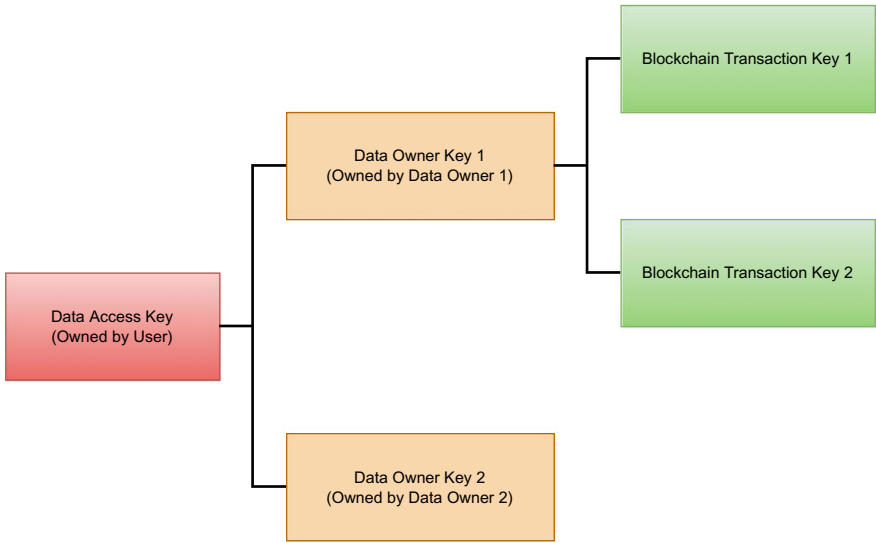


Fig. 3 Key generation process for data access key

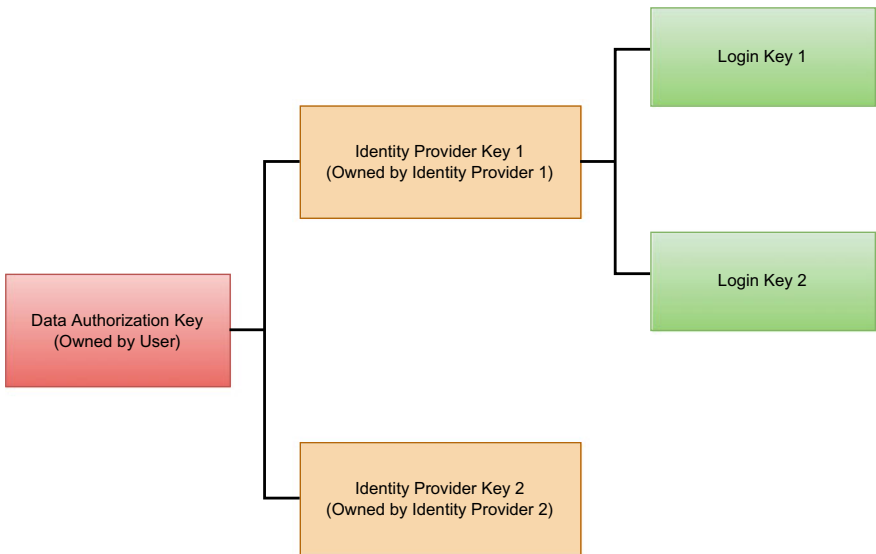


Fig. 4 Key generation process for data authorization key

gives 114 bits of security with a similarly efficient scheme. Thus, we recommend using the multiplicative version instead of the additive version for our system.

Identity-Based Conditional Proxy Re-encryption We propose to use Liang et al. [30] identity-based proxy re-encryption scheme. Their proposed scheme is col-

lusion resistant, secure against chosen-ciphertext attacks, and supports conditional re-encryption. By associating a condition with the encryption key, a sender can enforce fine-grained re-encryption over their initial ciphertexts. Only, the ciphertexts that meet the specified condition can be re-encrypted by the proxy holding the corresponding encryption key.

The user would generate a public key and master private key initially. The user would use the public key of the identity provider and the identity provider's identity to encrypt any data that they would want to store/transmit via the identity provider. Whenever the user wants to share their identities with the data owners for data verification, they would generate a re-encryption key based on the public key of the data owner and the user's identity. This ensures that the identity provider cannot access the identities when transmitting data via the identity provider. The data owners can decrypt the identities using their private keys. Liang et al. [30] explain the process of key generation and provide a security analysis for their proposed scheme. Section 5.3 explains the usage of IBCPRE, while the user tries to log in to the service provider.

5 System Architecture

This section explains the communication patterns between the user, identity provider, service provider, and the data owners residing on the blockchain. Our system uses single-sign-on facilities to gain access to online services. We explain the proposed scheme in four phases:

1. User registration with data owners.
2. User interaction with identity provider(s).
 - (a) User registration with identity provider.
 - (b) User login with identity provider.
3. User login with service provider.
4. User identity registration with identity provider.

We assume that the user's client application has already generated the "Data Access Key" and "Data Authorization Key" before proceeding further.

5.1 *User Registration with Data Owners*

This is the first stage of the process. In this step, the user registers their identity keys with the data owner(s) offline. The user uses their "Data Access Key" to derive individual data owner keys, as can be seen from Fig. 3. The user's client application keeps these details safe. These data owner keys are then registered with respective data owner(s), either during the identity creation or later. These data owner keys will

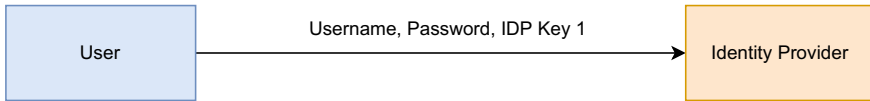


Fig. 5 User registration with IDP

serve as the pseudo-identifier for the individual on the blockchain. This would prevent other consortium members on the blockchain from co-relating these identifiers with the user since every data owner will have their separate data owner key.

5.2 User Interaction with Identity Provider(s)

This section describes user interaction with the identity provider. We explain the user registration and sign-in process.

User registration with identity provider We propose a multi-factor authentication system for our system design. Figure 5 details the user signup process with an identity provider. The user generates an identity provider key from the “Data Authorization Key,” as shown in Fig. 4. The user chooses a *username*, *password*, and the recently generated *identity provider key* as signup parameters with the identity provider. The user also signs up with a TOTP-based [32] 2FA (Two-factor authentication) authenticator app for added security.

We recommend using multi-factor authentication methods over TOTP-based authentication methods since there have been multiple incidents where the latter has proved to be not secure enough.^{7, 8} We also recommend following NIST’s password policy guidelines [9].

User login with identity provider This section describes the user login process with the identity provider(s). Figure 6 can be used as a reference to understand this process. There are multiple steps in this phase:

1. The user enters the *username* and *password*. If verified, it moves on to the TOTP-based 2FA screen. If this fails, the user is asked to re-enter the details.
2. Once username and password are verified, the user verifies the 2FA code. If approved, we move to the next step; else, we redirect the user to the failed login state.
3. Once the 2FA code is verified, the user generates a login key from the identity provider key of the current identity provider key that the user holds. The user generates a login key by modifying the index i in the hierarchical deterministic key generation process. The user sends the login key with the index i to the identity provider.

⁷ <https://www.theverge.com/2019/8/14/20805194/suprema-biostar-2-security-system-hack-breach-biometric-info-personal-data>.

⁸ <https://www.varonis.com/blog/box-mfa-bypass-totp>.

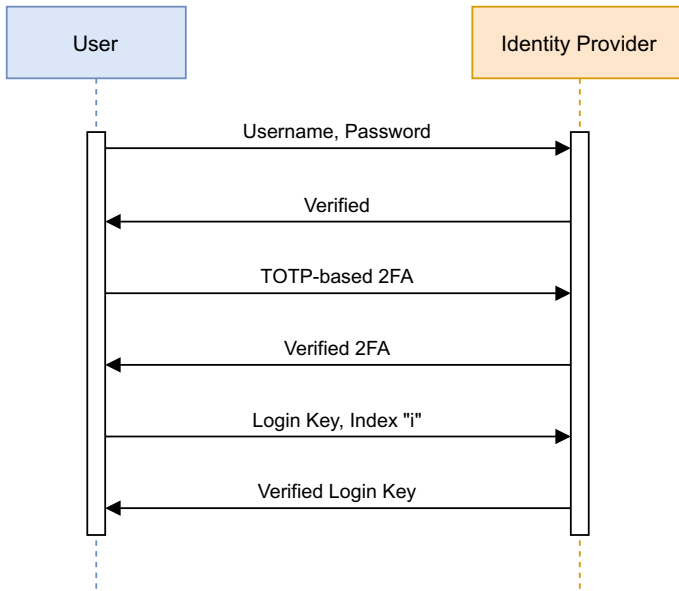


Fig. 6 User login with IDP

4. The identity provider has the identity provider key of the user. It fetches that and uses the input index i received by the user. If the generated key matches the login key that the identity provider got as input from the user, we can safely say that the user authentication was successful.

This multi-step authentication process adds friction to the sign-in process but ensures security.

5.3 User Login with Service Provider

Our system uses OpenID Connect for federated authentication and authorization. All communication between the identity provider and the consortium of data owners is through the communication server of the consortium. Figure 7 can be used as a reference.

1. A user can select the identity provider when they try to log in to a service provider. A user attempting to log in to a service using some IDP is redirected to that IDP with the required claims whenever they wish to use it.
2. The user logs in to the IDP. The user can select specific data owners that they want to use to certify certain attributes of their identity. The user can choose relevant data owners based on the service they are trying to access and the required attribute claims. If, for example, a user needs their date of birth for their social security

number, they can use their passport as the identity source since it is a better source of truth for the said attribute. However, if a service like an email provider needs a user's name, users can use their social media identities to get those attribute claims. Choosing the data owners depends on the kinds of service the user is trying to avail.

3. The user uses their secret IBCPRE key to encrypt the data identity and the data owner key. Next, they send these encrypted documents to the identity provider. We need to note that even the identity provider will not be able to access these documents at this stage since they are encrypted.
4. The user uses the data owner's public key to generate a re-encryption key to re-encrypt the data owner key and identity. These encrypted objects are sent to the data owners, where they decrypt these objects using their secret IBCPRE keys and the list of required attributes.
5. The data owner decrypts the objects, fetches the identities stored on their systems against the received data owner key and verifies information on their system with the input identity document from the user. If these documents match, the data owner sends a green signal to the identity provider about the veracity of the attributes. It also sends information about the latest re-certification event for the said attribute from the blockchain.
6. The data owner adds a new transaction on the blockchain detailing the data access request. The blockchain transaction holds timing information about the data access request, identity provider details, service provider details, and the requested attribute details. The data owner generates a new blockchain transaction key for the user using the data owner key of the user for the above transaction.
7. Once the identity provider receives the green signal from the data owners, the user generates a re-encryption key for the identity provider. The identity provider can decrypt the already stored documents. Since we are only generating re-encryption keys (IBCPRE) for the data owner and the identity provider, the documents sent to the data owner are the same as being used by the identity provider.
8. Based on the decrypted documents, the identity provider can calculate a trust score for the required attributes. The identity provider sends over the asserted attribute values and the trust score for the said attribute back to the service provider.
9. If the trust score is greater than the threshold set by the service provider for a threshold, the user is granted access to the service. Else, they will have to choose a different data owner.

5.4 User Identity Registration with Identity Provider

When the user tries to log in to the service provider via the identity provider, the user has to furnish their data documents every single time. That becomes a hassle and adds friction to the entire process. This section proposes a secure way to store encrypted documents with the identity provider. However, we do not recommend

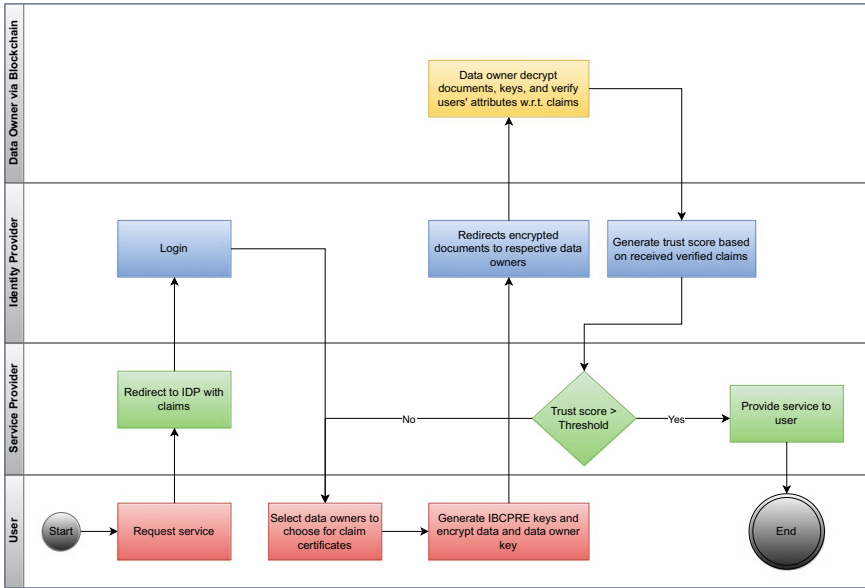


Fig. 7 User login with service provider

using this method since we have to trust the identity provider to some extent here. Despite the apparent security concerns, this method reduces the time required to verify the identity of a user and reduces the communication overhead with the data owners consortium.

1. A majority of the steps remain the same as mentioned in Sect. 5.3. The user logs in to the IDP, selects data sources, and tries to verify their identities with the respective data owners.
2. If the user’s identity document is verified, the data owner sends a green signal to the identity provider. However, no information about re-certification events is sent since this was not requested.
3. Now, the identity provider stores the verified but encrypted document in their systems. The user can use this every time they want to use these identities to avail of online services.
4. Whenever a user would try to use their identities stored at the identity provider, the user would generate a re-encryption key using the identity provider’s IBCPRE public key. The identity provider would decrypt the stored document and verify the identity attribute claims being requested by the service provider.

One apparent disadvantage to this method is that the trust score of an attribute might be slightly lower since the identity provider has no information about the re-certification events. Also, the user might not be able to track their data access requests as they could using the blockchain. Nevertheless, this does provide a possible business

opportunity for the identity providers to lure users into using their services since they might provide additional features such as tracking data access requests. Essentially, it is up to the user to decide what identity provider they want to use.

6 Discussion

In this section, we analyze the security of our proposed scheme and discuss our contributions with respect to the existing systems.

6.1 Identity Security

Firstly, the users generated the data access key and data authorization key and stored them locally. Thus, no adversary can obtain them without physical access to the user's device. Secondly, the user needs to provide their identity to the identity provider only when trying to access online services. The identities are encrypted using IBCPRE, and the user maintains the encryption keys. Thus, not even the identity provider has access to the documents unless the user generates the re-encryption keys for the identity provider. The data access key is also encrypted using the IBCPRE scheme when in transit, thus disallowing the identity provider from ever gaining access to it. Lastly, data access key and data authorization key are separate and are used for different use-cases altogether, thus, providing the much-needed barrier between data authentication and data authorization.

6.2 Data Confidentiality and Privacy

Our proposed system uses state-of-the-art cryptographic algorithms to ensure the confidentiality of data. Our IBCPRE scheme uses AES-512 to encrypt the documents/information in transit. Also, a user creates a different data owner key for every data owner, ensuring the segregation of information between the data owners. Our hierarchical deterministic key generation scheme allows users to have a separate transaction key on the blockchain every time a new transaction is committed, thus reducing the chances of data correlation if an adversary gains unauthorized access to the blockchain. The blockchain does not store any private information about a user, just metadata about data access requests and re-certification events.

6.3 Data Transparency

The blockchain stores only the identity attributes' data access requests and re-certification events. Since the transaction keys used for these transactions are derived from a single parent key, the data owner can trace all the transactions belonging to the same user. This assures traceability and auditability of transactions as and when needed. Also, the user can trace back their data access requests since they hold the master data access key; the user can trace transactions on the blockchain for all data owners.

6.4 Transaction Integrity and Non-repudiation

Firstly, using a trusted, permissioned blockchain ensures that all consortium members of a private blockchain trust each other. Since these transactions on the blockchain are immutable and permanent, they cannot be modified or deleted, thus ensuring the system's integrity. Secondly, only, the user holds the data access key and can thus initiate and finalize a data access request. This ensures non-repudiation of the identity.

6.5 Authentication and Authorization

In our proposed scheme, we follow a multi-factor authentication strategy. Section 5.2 explains how we maintain security during the login phase. The identity provider has no idea about the user's identity but only serves as a trusted proxy between the user and the service provider. It relies on the data owner to verify the identity of the users. Our proposed system suggests using attribute-based access control policies to limit data leakage to the service provider. The service provider could request as many attributes as they want; however, it is up to the user to authorize/unauthorize what attribute information needs to be propagated to the service provider. The service provider can make specific attribute requirements mandatory and limit access to complete service functionality based on the approved attributes it receives from the user. For example, suppose the user does not wish to disclose their age to a media streaming platform. The media streaming platform could disallow users from viewing adult content on their platform. Our platform establishes that even though the user has control over their identity, they would still have to face the consequences of not allowing access to the required attributes; however, at least, the user has a choice.

6.6 Scalability and Availability

Our proposed scheme stores only re-certification events and data access requests on the blockchain and not the complete identity [48]. Hence, this allows us not to overload the blockchain, thus improving the system's scalability. In the real world, data owners are regarded as data custodians. For example, the passport of a citizen is still the property of that country's government and not the user. Thus, we extend this same philosophy over the Internet and do not focus on self-sovereign identities like a few previous studies [21, 35]. The system can support large-scale user identity management and authentication by supporting multiple consortiums of data owners.

One limiting factor in the proposed scheme is the availability of data sources when needed. If the data sources are offline, the user might be blocked. We can overcome this by letting users upload their identities on the identity provider (though not recommended). The system could also be modified to allow the identity provider to generate a lower trust score if a data source is unavailable if they have the user's consent.

7 Conclusion and Future Work

This paper proposes a blockchain-based, privacy-preserving user identity management system. In our proposed identity management protocol, users can use their existing identities to access services using hierarchical deterministic keys to ensure access to their data and a proxy re-encryption scheme to store and transmit data securely. Our system uses existing single-sign-on facilities with attribute-based access control for efficient, fine-grained user authentication and authorization. Furthermore, the user can also trace their authentication and data access requests as and when needed using a secure, permissioned blockchain. Using multiple data owners allows users to maintain and control different slivers of their identity. An identity provider assigns a trust score for every asserted attribute based on the data sources and various attribute re-certification events. This trust score allows service providers to control or limit access to a particular service aspect based on their internal thresholds for attributes. Our proposed identity management scheme gives users control over their data, allowing them to access online services securely.

To that end, there is room for improvement. This study is one of the initial stages toward the implementation of a fully functional privacy-first federated identity management system. In order to understand and utilize the system, the interface will need to abstract away most of the blockchain and identity provider interactions. As a result, more fieldwork in usability testing will be required. We also need to do some performance testing to see if this system architecture is feasible in the real world. While there are cryptographic guarantees for the privacy of the transaction when it is stored, there is always the risk of some data leakage (e.g., by usage, frequent updates, etc.) that might worry policymakers. As a result, further security testing would be required. In the future, we want to resolve these concerns.

References

1. Agrawal A, Rathor S (2020) A robust verification system for recruitment process by using blockchain technology. *Int J Bifurc Chaos* 1:389
2. Almadhoun R, Kadadha M, Alhemeiri M, Alshehhi M, Salah K (2018) A user authentication scheme of iot devices using blockchain-enabled fog nodes. In: 2018 IEEE/ACS 15th international conference on computer systems and applications (AICCSA), pp 1–8. <https://doi.org/10.1109/AICCSA.2018.8612856>
3. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, Enyeart D, Ferris C, Laventman G, Manevich Y, et al (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the thirteenth EuroSys conference, pp 1–15
4. Basney J, Flanagan H, Fleury T, Gaynor J, Koranda S, Oshrin B (2019) CILogon: enabling federated identity and access management for scientific collaborations. PoS ISGC2019:031. <https://doi.org/10.22323/1.351.0031>
5. Beltran V, Martinez JA, Skarmeta AF (2017) User-centric access control for efficient security in smart cities. In: 2017 Global Internet of Things Summit (GIoTS), pp 1–6. IEEE
6. Bendiab G, Shiaele SN, Boucherkha S, Ghita BV (2019) Fcmdt: a novel fuzzy cognitive maps dynamic trust model for cloud federated identity management. *Comput Secur* 86:270–290
7. Berghel H (2020) The equifax hack revisited and repurposed. *Computer* 53(5):85–90
8. Bethencourt J, Sahai A, Waters B (2007) Ciphertext-policy attribute-based encryption. In: 2007 IEEE symposium on security and privacy (SP'07), pp 321–334. IEEE
9. Burr WE, Dodson DF, Newton EM, Perlner RA, Polk WT, Gupta S, Nabbus EA (2011) Sp 800-63-1. Electronic authentication guideline
10. Cai X, Geng S, Zhang J, Wu D, Cui Z, Zhang W, Chen J (2021) A sharding scheme-based many-objective optimization algorithm for enhancing security in blockchain-enabled industrial internet of things. *IEEE Trans Indus Inform* 17(11):7650–7658. <https://doi.org/10.1109/TII.2021.3051607>
11. Cui Z, Xue F, Zhang S, Cai X, Cao Y, Zhang W, Chen J (2020) A hybrid blockchain-based identity authentication scheme for multi-wsn. *IEEE Trans Serv Comput* 13(2):241–251. <https://doi.org/10.1109/TSC.2020.2964537>
12. Das P, Erwig A, Faust S, Loss J, Riahi S (2021) The exact security of bip32 wallets. In: Proceedings of the 2021 ACM SIGSAC conference on computer and communications security, pp 1020–1042
13. Dunphy P, Petitcolas FA (2018) A first look at identity management schemes on the blockchain. *IEEE Secur Priv* 16(4):20–29
14. Ekblaw A, Azaria A, Halamka JD, Lippman A (2016) A case study for blockchain in healthcare : “medrec” prototype for electronic health records and medical research data
15. Faber B, Michelet GC, Weidmann N, Mukkamala RR, Vatrupu R (2019) Bpdim: a blockchain-based personal data and identity management system. In: Bui T (ed) 52nd Hawaii international conference on system sciences, HICSS 2019, Grand Wailea, Maui, Hawaii, USA, January 8–11, 2019, pp 1–10. ScholarSpace/AIS Electronic Library (AISel). <http://hdl.handle.net/10125/60121>
16. Francisco K, Swanson D (2018) The supply chain has no clothes: technology adoption of blockchain for supply chain transparency. *Logistics* 2(1). <https://doi.org/10.3390/logistics2010002>, <https://www.mdpi.com/2305-6290/2/1/2>
17. Gao S, Su Q, Zhang R, Zhu J, Sui Z, Wang J (2021) A privacy-preserving identity authentication scheme based on the blockchain. *Secur Commun Netw*
18. Gao Z, Xu L, Turner G, Patel B, Diallo N, Chen L, Shi W (2018) Blockchain-based identity management with mobile device. Association for Computing Machinery, New York. <https://doi.org/10.1145/3211933.3211945>
19. Gokhale P, Malik S, Gilda SS, Rizvi SH, Poulouse R (2021) Identity attribute confidence scoring while certifying authorization claims, 7 Sep 2021, US Patent 11,115,419

20. 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
21. Grech A, Sood I, Ariño L (2021) Blockchain, self-sovereign identity and digital credentials: Promise versus praxis in education. *Front Blockchain* 4. <https://doi.org/10.3389/fbloc.2021.616779>, <https://www.frontiersin.org/article/10.3389/fbloc.2021.616779>
22. Gutoski G, Stebila D (2015) Hierarchical deterministic bitcoin wallets that tolerate key leakage. In: International conference on financial cryptography and data security, Springer, pp 497–504
23. Hammi MT, Hammi B, Bellot P, Serhrouchni A (2018) Bubbles of trust: a decentralized blockchain-based authentication system for iot. *Comput Secur* 78:126–142. <https://doi.org/10.1016/j.cose.2018.06.004>, <https://www.sciencedirect.com/science/article/pii/S0167404818300890>
24. Hardjono T, Pentland A (2019) Core identities for future transaction systems. MIT Press
25. Hardjono T, Smith N, Pentland AS (2014) Anonymous identities for permissioned blockchains
26. Hu VC, Kuhn DR, Ferraiolo DF, Voas J (2015) Attribute-based access control. *Computer* 48(2):85–88. <https://doi.org/10.1109/MC.2015.33>
27. Josh (2021) What methods are used by wallets to generate keys? <https://cryptochamp.com/what-methods-are-used-by-wallets-to-generate-keys/>
28. Laborde R, Oglaza A, Wazan S, Barrere F, Benzekri A, Chadwick DW, Venant R (2020) A user-centric identity management framework based on the w3c verifiable credentials and the fido universal authentication framework. In: 2020 IEEE 17th annual consumer communications networking conference (CCNC), pp 1–8. <https://doi.org/10.1109/CCNC46108.2020.9045440>
29. Lenz T, Krnjic V (2018) Towards domain-specific and privacy-preserving qualified eid in a user-centric identity model. In: 2018 17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE), pp 1157–1163. <https://doi.org/10.1109/TrustCom/BigDataSE.2018.00160>
30. Liang K, Liu Z, Tan X, Wong DS, Tang C (2012) A CCA-secure identity-based conditional proxy re-encryption without random oracles. In: International conference on information security and cryptography, pp 231–246. Springer
31. Liu Y, He D, Obaidat MS, Kumar N, Khan MK, Raymond Choo KK (2020) Blockchain-based identity management systems: a review. *J Netw Comput Appl* 166:102731. <https://doi.org/10.1016/j.jnca.2020.102731>, <https://www.sciencedirect.com/science/article/pii/S1084804520302058>
32. M'Raihi D, Machani S, Pei M, Rydell J (2011) Totp: time-based one-time password algorithm. Tech Rep
33. Nofer M, Gomber P, Hinz O, Schiereck D (2017) Blockchain. *Bus Inform. Syst Eng* 59:3. <https://doi.org/10.1007/s12599-017-0467-3>
34. Premarathne US, Khalil I, Tari Z, Zomaya A (2017) Cloud-based utility service framework for trust negotiations using federated identity management. *IEEE Trans Cloud Comput* 5(2):290–302. <https://doi.org/10.1109/TCC.2015.2404816>
35. Rathee T, Singh P (2022) A self-sovereign identity management system using blockchain. In: Khanna K, Estrela VV, Rodrigues JJPC (eds) *Cyber security and digital forensics*. Springer, Singapore, pp 371–379
36. Ruffing T, Moreno-Sanchez P, Kate A (2014) Coinshuffle: practical decentralized coin mixing for bitcoin. In: Kutyłowski M, Vaidya J (eds) *Computer security—ESORICS 2014*. Springer International Publishing, Cham, pp 345–364
37. Ruffing T, Moreno-Sanchez PA, Kate A (2016) P2p mixing and unlinkable bitcoin transactions anonymity of the people, by the people, and for the people
38. Saberi S, Kouhizadeh M, Sarkis J, Shen L (2019) Blockchain technology and its relationships to sustainable supply chain management. *Int J Prod Res* 57(7):2117–2135. <https://doi.org/10.1080/00207543.2018.1533261>, <https://doi.org/10.1080/00207543.2018.1533261>
39. Sakimura N, Bradley J, Jones M, De Medeiros B, Mortimore C (2014) Openid connect core 1.0. The OpenID foundation, p S3

40. Saxena A, Misra J, Dhar A (2014) Increasing anonymity in bitcoin. In: Böhme R, Brenner M, Moore T, Smith M (eds) *Financial cryptography and data security*. Springer, Berlin, pp 122–139
41. Selvanathan N, Jayakody D, Damjanovic-Behrendt V (2019) Federated identity management and interoperability for heterogeneous cloud platform ecosystems. In: *Proceedings of the 14th international conference on availability, reliability and security*
42. Singh BC, Carminati B, Ferrari E (2021) Privacy-aware personal data storage (p-pds): learning how to protect user privacy from external applications. *IEEE Trans Depend Secure Comput* 18:889–903
43. Wang W, Xu H, Alazab M, Gadekallu TR, Han Z, Su C (2021) Blockchain-based reliable and efficient certificateless signature for iiot devices. *IEEE Trans Indus Inform* 1. <https://doi.org/10.1109/TII.2021.3084753>
44. Xu J, Xue K, Tian H, Hong J, Wei DSL, Hong P (2020) An identity management and authentication scheme based on redactable blockchain for mobile networks. *IEEE Trans Ve Technol* 69(6):6688–6698. <https://doi.org/10.1109/TVT.2020.2986041>
45. Yaqoob I, Salah K, Jayaraman R, Al-Hammadi Y (2021) Blockchain for healthcare data management: opportunities, challenges, and future recommendations. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-020-05519-w>
46. Yu K, Tan L, Aloqaily M, Yang H, Jararweh Y (2021) Blockchain-enhanced data sharing with traceable and direct revocation in iiot. *IEEE Trans Indus Inform* 17(11):7669–7678. <https://doi.org/10.1109/TII.2021.3049141>
47. Zheng Z, Xie S, Dai HN, Chen X, Wang H (2018) Blockchain challenges and opportunities: a survey. *Int J Web Grid Serv* 14:352. <https://doi.org/10.1504/IJWGS.2018.095647>
48. Zhuang C, Dai Q, Zhang Y (2022) Bcppt: a blockchain-based privacy-preserving and traceability identity management scheme for intellectual property. *Peer-to-Peer Netw Appl* 15:1–15. <https://doi.org/10.1007/s12083-021-01277-1>

Efficient Big Data Clustering Using Adhoc Fuzzy C Means and Auto-Encoder CNN



Shreehari Harthi Sreedhara, Vineet Kumar, and Subia Salma

Abstract Clustering, a well-known unsupervised machine learning technique is effective in handling massive amount of data for a variety of applications. Clustering is the process of grouping identical data into clusters that belong to a specific category. Big data consists of massive amounts of data generated every second; hence, clustering this data is quite challenging. Deep learning techniques are used to analyze large amounts of data, which necessitates the use of a large number of samples for training, resulting in a time-consuming and inefficient process. The adhoc fuzzy *C* means method can be used to overcome this limitation. In this research study, a fusion approach known as fusion clustering is built by combining auto-encoder features with adhoc fuzzy *C* means (AFCM) technique to improve the fusion approach. The auto-encoder, in combination with the CNN approach, is used for feature engineering to improve the performance, resulting in increased computational speed. The efficiency of the proposed model is improved by evaluating each performance metrics in detail by using three different data types: MNIST, Fashion-MNIST dataset, and USPS. The comparison of each dataset demonstrates that the proposed adhoc fuzzy *C* means (AFCM) model outperforms other state-of-the-art techniques.

Keywords Adhoc fuzzy *C* means clustering · Fashion-MNIST dataset · USPS · Auto-encoder · Adhoc fuzzy *C* means (AFCM)

1 Introduction

Recently, a huge amount of data is generated from various domains such as health care, social media websites, bank transactions, and so on. The data [1] generated

S. H. Sreedhara

Department of ECE, S J C Institute of Technology, Chikkaballapur, Karnataka, India

e-mail: shreehari.harthi@gmail.com

V. Kumar · S. Salma (✉)

Ramaiah Institute of Technology, Bengaluru, Karnataka, India

e-mail: subiasalma89@gmail.com

consists of confidential data as well as other information related to business intelligence, weather forecasting, and research data collected through various methods. As digitization influences every aspect of human lives, massive amounts of data are generated and stored. Exploring the petabytes of data [2] generated by Walmart and Facebook is a time-consuming task. Big data refers to the enormous amount of useful data generated; exploration and analysis of this big data are highly required to extract the necessary information for business intelligence and related research.

As a vast amount of data is present on the Internet and multiplies exponentially, big data is distinguished into three types, [3] they are structured, semi-structured, and unstructured data. The major chunk of data generated is unstructured data, and this data requires new methods to be handled. Big data possess $3V_s$, they are volume, velocity, and variety. The volume features the growth of data exponentially as the technology is improvising each day, it results in generating an enormous amount of data. According to the research, the generated data has doubled over 4 years. Velocity refers to how fast the data could be generated to flow from one direction to another, and hence, data progresses very quickly. Variety stands for the collection of structured as well as unstructured data, which is extracted from various sources. The data is then generated [4] in a real-time scenario. This defines the difficulties of utilizing the data generated from different firms with the traditionally generated data.

The data generated by various means; like health care, scientific research, and other fields are grouped for analysis, this process is termed as clustering. Consider an example [5], when a person is suffering from cancer with certain symptoms is grouped into one cluster and further, the other person showing tuberculosis symptoms cannot be grouped into the same cluster. An unsupervised technique [6] for classifying and analyzing the data generated by machine learning algorithms are termed as clustering. Clustering technique segments the data into different groups in a specific way will result in the clustering of similar data. The clustering mechanism is stated by using Eq. 1; considering CS as a cluster segment, CS1 and CS2 are different clusters, which is considered as a machine learning approach (Fig. 1).

$$CS_1 \xi CS_2 \xi CS_3 \dots \xi CS_n = \Omega \quad (1)$$

Generally, the clustering mechanism is divided based on two features known as single machine clustering and multiple machine clustering; recently, multiple machine clustering is widely used in different applications. In a single machine clustering algorithm, the data mining technique aims at combining the data in relevance according to their generic constraints. However, this clustering mechanism is focused on segmentation process, which segments the dataset in a single partition via its distance for clustering based on their similarities. The limitation of this segmentation approach is that it needs a constraint that is already defined and non-deterministic. Single machine clustering is termed as a bottom-up approach, whereas hierarchical clustering is termed as a top-down approach. Hierarchical clustering is divided into two types, known as divisive and agglomerative; agglomerative is said to be a bottom-up approach that groups the clusters into a single object and

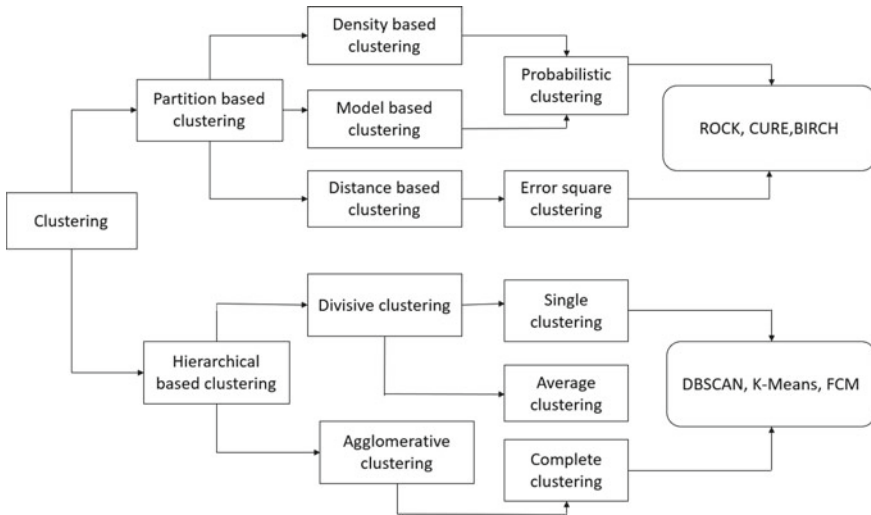


Fig. 1 Different clustering mechanisms

then combine into the atomic cluster and further on synchronizes into a larger one, however, in hierarchical clustering, where each clustering mechanism is segmented into different clusters until termination is considered. The segmentation approach of clustering goes in the opposite sense of hierarchical clustering. Data is allocated to a particular cluster without subsequent structure by considering the Euclidean distance. Euclidean distance calculates the minimum distance, which is experiential amid the cluster and points allocated. This clustering mechanism deals with several pros and cons; such as the clustering approach has benefits of implementation, whereas its drawback results in dealing with huge amounts of data. Further, the dimensionality reduction approach provides dataset reduction and reduction of cost but failed in generating the effective solutions in the scenario of a dimensional dataset and has several limitations. Further, parallely classifying this technique helps to minimize the execution time and delivers usability. To address several complexities from this mechanism like that of the MapReduce is designed for offering a highly scalable nature which in turn gives an appropriate response, which does not necessarily generalize the functionalities, and implementing each task is quite tedious and indeed a requirement of more resources.

1.1 Motivation and Contribution of Research Work

The total weights associated with measuring the weighted distance in accordance with the prerequisite in the clustering technique are assigned arbitrarily by the user. Fuzzy and neural networks focus on utilizing the knowledge process; previous works

indicate that very few deep learning studies have been conducted, and an auto-encoder-based CNN mechanism is not used. To improve the value of performance metrics, an improved fuzzy C means mechanism by including the auto-encoder CNN with fuzzy C means is used. The contribution to this research is stated below:

- In this research, an efficient fuzzy C means approach is modeled to improve the accuracy in big data clustering.
- In the process of enhancing the fuzzy C means mechanism comprising a typical framework, the first phase deals with the auto-encoder-based CNN mechanism, the second phase deals with the clustering mechanism, and the third phase with designing the multidimensional feature for feature engineering.
- Effective way for enhancing fuzzy C means for dual CNN and execution of big data.
- Enhancing fuzzy C means is evaluated by establishing data such as MNIST, Fashion-MNIST, and USPS by comparing it with the existing model.
- To enhance the FCM to perform better than the existing model.

2 Related Work

Data clustering is a primary concern these days. This involves segmenting the objects into different clusters, where several algorithms are proposed for clustering the data, ranging from simple to complex algorithms. A thorough review is carried out based on the existing methodologies for clustering purpose. An approach is modeled in [1, 7, 8] for the process of clustering and outlier detection, where dendrograms are used for various taxonomical applications [9], in [2] a clustering-based segmentation approach has been used for performing time-series clustering and segmentation based on time-series data to track the patient.

At first, VAT is used to express clustering in order to alter the matrix for clustering and displaying the block diagonally. This approach uses VAT for security purposes and additionally K -means is a technique used to cluster the data efficiently. The k -means technique is a simplistic approach applied in various domains and is used as a standard for clustering. Generally, the K -means algorithm has various drawbacks as it is not properly initialized and also functions as a topology-based algorithm.

Recently, for clustering the big data, a learning mechanism is considered as a supervised approach to achieve results satisfactorily, however, this approach has certain limitations to achieve the appropriate result for the data and it also affects the accuracy. Many fuzzy techniques are developed to handle the uncertainty problem in clustering. A hierarchical approach is developed to combine the neural network along with a fuzzy methodology for performing an effective clustering. In [10], a fusion approach is built by integrating fuzzy and CNN for classification and clustering purposes. In the first approach, CNN is deployed to automate the process of feature extraction; in the second phase, the adhoc fuzzy C means approach is used for clustering purpose in an appropriate feature space. A neural network-based approach [3] is developed for performing classification and clustering mechanisms.

A set theory methodology is used for performing feature extraction, where the associated features are extracted to serve as input, and then fed to a neural network for classification and clustering purposes. This approach discussed here performs data handling effectively. These techniques are based on a supervised learning approach for collecting enormous amounts of data for training which is time-consuming.

A semi-supervised clustering approach for handling clustering and classification. They have developed a restrictive approach for labeling the data at the base level. In the beginning label, prediction for clustering algorithm and to train the neural network initially to predict the label. Label prediction [4] helps in distinguishing features, to introduce by localizing the features by initializing a pre-trained network which benefits the clustering and classification approach. To develop a model for the learning approach known as MT-model [5, 6, 11] which uses a semi-supervised learning technique. This model estimates an average weight of the MT-model which finds an average to format the teacher model. This model was built for a large dataset to develop an effective model for the deep neural network which introduces a self-ensemble mechanism for prediction of the dataset, this situation fails to achieve an accurate model for training purposes through various epoch values. This approach performs very well on the general dataset, this situation fails to achieve better accuracy for a dataset that is noisy and uncertain.

By considering this existing methodology, this clustering approach [12] has an issue of computational time, which deals with clustering while considering the performance metrics. There exists a research gap which is observed between the FCM to possess a great potential to compare the existing approach like that of k -means which is a traditional approach. The existing system follows a clustering mechanism that groups into several categories wherein FCM is a clustering mechanism. This FCM approach fails computations and performance metrics.

3 Proposed Methodology

The proposed model is based on a neural network, where big data is represented in different types to generalize the metrics, which is known as a multidimensional array. Big data is categorized in various forms such as high dimensionality, multidimensional aspects, and tabular and large-scale characters. The proposed model for big data is discussed here, $A \in F^{\alpha_1 \times \alpha_2 \times \dots \times \alpha_n}$ as a multidimensional array having a size $\alpha_1 \times \alpha_2 \times \dots \times \alpha_n$, big data is segmented into three categories such as unstructured data, structured, and semi-structured.

3.1 Initialization

Generally, the objects are assigned to each cluster based on a process known as clustering; however, this approach is neither feasible nor required. By assigning

each object to multiple clusters, the fuzzy approach overcomes this challenge. The fuzzy C means algorithm is modified in the proposed model.

3.2 Clustering and Classification of Adhoc Fuzzy C Means Algorithm

The adhoc FCM approach functions by allocating all data points to an adjacent cluster known as cluster centroid, which is calculated based on the cluster distance. Adhoc FCM methodology algorithm is shown in Table 1. In fuzzy clustering, the centroid of a cluster is considered as the mean of all points, which is weighted by the degree belonging to that cluster. The primary goal of the proposed adhoc fuzzy approach is it provides exceptional results for data that is overlapped and allocates the data point to multiple clusters. Several computations are required along with accuracy and the Euclidean distance is used for unequal weight measurements. To improve the performance of the fuzzy c means algorithm, it executes first and then assigns a high importance degree to those samples that have a high membership degree to their clusters, and then, they are used for training. The auto-encoder-based CNN mechanism simplifies the clustering process. The dataset is given as $A = \{a_1 \times a_2 \times \dots \times a_n\}$ integrated with set $B = \{b_1 \times b_2 \times \dots \times b_N\}$. Here, fuzzy C means is implemented. The auto-encoder is employed here for training purposes: the membership set $M = \{|M_{xy}| \ 1 \leq x \leq h, \ 1 \leq y \leq N\}$

$$\sum_{x=1}^h \sum_{y=1}^g M_{xy}^o U_y - a_x^2$$

$$\sum_{y=1}^h M_{xy} = 1, \quad M_{xy} \geq 0 \tag{2}$$

A fuzzy C means approach is then modified to avoid the membership units equalizing to 1. The below equation shows that

$$R_o(T, V) = \sum_{x=1}^h \Omega_a \sum_{x=1}^g (1 - v_{xR}^o)^o + \sum_{x=1}^h \sum_{x=1}^g M_{xy}^n U_y - U_a^2 \tag{3}$$

To optimize this equation, this updates the membership matrix to cluster the centers as

$$a_x = \sum_{y=1}^g M_{xy}^o U_l / \sum_{y=1}^g M_{xy} \tag{4}$$

Table 1 Adhoc_FCM methodology

Input: dataset, Max
Output: optimized cluster member and membership vector
Step 1: Initialization of membership_matrix
Step 2: for $k = 1$ to e do
Step 3: cluster center updation
Step 4: updating the fuzzy constant
Step 5: end for
Step 6: for $= 1$ to e do
Step 7: for $l = 1$ to e do
Step 8: cluster center membership
Step 9: end for loop (Step 8)
Step 10: end for loop (Step 7)
Step 11: end for loop (Step 3)

Membership matrix

$$M_{xy} = \left(1 + \left(\frac{\mu_{xy}}{\Omega_x} \right)^{-1/(o-1)} \right)^{-1} \tag{5}$$

In this Equation, μ_{xy} calculates the distance between the cluster and the matrix used for clustering.

3.3 System Architecture

In this, adhoc fuzzy C means is considered as a multidimensional array to incorporate the correlation facilities on different models. However, to deploy an adhoc fuzzy C means, a CNN model is developed for training purpose. An auto-encoder-based CNN model is then designed and the adhoc fuzzy C means algorithm is developed to achieve the desired clustering outcome as in Table 2.

3.3.1 Mathematical model

The mathematical model here performs computations by considering CNN as the base module to pre-process various parameter constraints that require high computations, hence consuming a lot of time. The proposed approach is further extended to design a model that is highly optimized by reducing its computational complexity in order to minimize the time by considering the constraints. The proposed optimal model is fed with input as $A \in F^{\alpha_1 \times \alpha_2 \times \dots \times \alpha_1}$.

Table 2 Auto-encoder-based CNN model

Input: dataset, $\hat{\Gamma}$, x, y
Output: cluster optimized matrix
Step 1: Initialization of cluster_matrix
Step 2: for $x = 1$ to $\hat{\Gamma}$ do
Step 3: for $x = 1$ to y
Step 4: update the cluster center $\Omega_x = \sum_{y=1}^g M_{xy}^y \beta_{uv(xy)} / \sum_{y=1}^o M_{xy}^o$
Step 5: for $x = 1$ to μ do
Step 6: for $x = 1$ to g do
Step7: $w_{kl} = \left(\left(1 + \left(\frac{\beta_{uv(xy)}}{\Omega_j} \right)^{-1/(o-1)} \right)^{-1} \right)$
Step 8: end of for loop (step 6)
Step 9: end of for loop (step 5)
Step 10: end of for loop (step 2)

$$\text{HL}_{y_1 \dots y_g} = \text{Auto_enc}(\theta) \left(\sum_{x_1 \dots x_g}^{x_1 \dots x_g} s_{x_1 \dots x_g}^{(1)} + Y_{\rho x_1 \dots x_g}^{(1)} Z_{x_1 \dots x_g} \right) \quad (6)$$

$$\text{OL}_{x_1 \dots x_g} = \text{Auto_enc}(\theta) \left(\sum_{y_1 \dots y_o}^{y_1 \dots y_g} s_{x_1 \dots x_g}^{(1)} + Y_{\Gamma y_1 \dots y_o}^{(1)} \text{HL}_{y_1 \dots y_o} \right) \quad (7)$$

In the equation above, x_1 indicates the dimension, y_1 indicates the hidden layer, and Auto_enc indicates auto-encoder. Further, a sigmoidal function is used in the hidden layer and outer layer.

3.4 Auto-Encoder CNN

This Auto-encoder CNN is modeled to achieve a higher computation and reduce the time without deteriorating the performance. This optimized CNN is built with two hidden layers. By optimizing ANN inputs, the $A \in F^{\alpha_1 \times \alpha_2 \times \dots \times \alpha_1}$ is extended for both the hidden layers by using this equation.

$$\text{HL}_{1y_1 \dots y_o} = \text{Auto_enc}(\theta) \left(\sum_{x_1 \dots x_g}^{x_1 \dots x_g} H_{\rho r_1 \dots r_g}^{(1)} F_{r_1 \dots r_g} + k_{y_1 \dots y_o}^{(1)} \right) \quad (8)$$

$$\text{HL}_{2c_1 \dots c_F} = \text{Auto_enc}(\theta) \left(\sum_{y_1 \dots y_g}^{y_1 \dots y_g} H_{\rho y_1 \dots y_g}^{(1)} \text{HL}_{y_1 \dots y_g} + k_{y_1 \dots y_o}^{(1)} \right) \quad (9)$$

Here, $HL_1 \in F^{y_1 \times y_2 \times \dots \times y_1}$ indicates second layer and $HL_2 \in F^{C_1 \times C_2 \dots \times C_o}$ in both, the equation $\text{Auto_enc}(\theta)$ indicates the auto-encoder.

$$OL_{x_1 \dots x_o} = \text{Auto_enc}(\theta) \left(\sum_{y_1 \dots y_o}^{C_1 \dots C_F} k_{x_1 \dots x_o}^{(2)} + Y_{\beta_{x_1 \dots x_o}}^{(1)} HL_{2C_1 \dots C_F} \right) \quad (10)$$

An optimization ANN to train the model is used to encode the activation function as

$$P'(k) = \begin{cases} 1 & k > 0 \\ 0 & k \leq 0 \end{cases} \quad (11)$$

These constraints train and reconstruct the function as

$$Y_{ATAE}(\theta) = 0.5(v - k)^F E(v - k) \quad (12)$$

In the equation above, k and v are vectors; E indicates this co-efficient. To reconstruct the function of the training sample given as

$$Y_{GRF}(\theta) = \left[1/2 \sum_{x=1}^c 0.5 \in (Y^{(1)2} + Y^{(2)2} + Y^{(3)2})^2 + 0.5(v - k)^F E(v - k) \right] \quad (13)$$

By applying, backpropagation to compute PH

$$K = K - P \left(\frac{1}{o} \sum_{x=1}^o \tau L + PL_x \right) \quad (14)$$

By applying backpropagation for Δk

$$k = k - P \left(\frac{1}{o} \sum_{x=1}^o \delta k_x \right) \quad (15)$$

In Eq. (16), P indicate the learning rate which is used in determining the derivative in each case where forward propagation is used for i/o computation.

$$\delta_x^{(4)} = \left(\sum_{y=1}^{x_1 \times \dots \times x_o} a_{xy} (k_x^{(3)} - v_{kx}) \right) \cdot P'(w_x^{(4)}) \quad (16)$$

$$\delta_{c_1 c_2 \dots c_F}^{(3)} = \left(\sum_{y=1}^{x_1 \times \dots \times x_o} k_{xy_1 \dots c_F}^{(3)} \cdot \delta_x^{(4)} \right) \cdot P'(w_x^{(4)}) \quad (17)$$

In the above equation, $\delta_{y_1, y_2 \dots y_F}^{(4)}$ and $\tau_{y_1, y_2 \dots y_O}^{(3)}$ are *I/O* values, and error estimation for each neuron is carried out through the below equation.

$$\delta_{c_1 c_2 \dots c_F}^{(2)} = \frac{\partial Y_{\text{Auto-enc}}^{(4)}(\tau)}{\partial w_{y_1 \dots y_g}^{(3)}} \quad (18)$$

The partial derivative of $\frac{\partial w^{(4)}}{\partial k^{(N)}}$ by considering this $N = 1, 2, 3$

$$O_{c_1 c_2 \dots c_F}^{(3)} = \frac{\partial b_{x_1 \dots x_O}^{(4)}}{\partial k_{y_1 \dots y_S}^{(3)}} \quad (19)$$

$$O_{c_1 c_2 \dots c_F}^{(2)} = \frac{\partial b_{x_1 \dots x_O}^{(3)}}{\partial k_{y_1 \dots y_S}^{(2)}} \quad (20)$$

$$O_{c_1 c_2 \dots c_F}^{(1)} = \frac{\partial b_{x_1 \dots x_O}^{(3)}}{\partial k_{y_1 \dots y_S}^{(2)}} \quad (21)$$

By computing the derivatives of δk and Δk

$$\Delta k^{(c)} = \frac{\partial y_{\text{Auto-enc}}^{(c)}}{\partial w^{(N+1)}} \cdot \frac{\partial y^{(N+1)}}{\partial k^{(N+1)}} \quad (22)$$

$$\Delta k^{(c)} = \zeta^{(c+1)} \quad (23)$$

The algorithm on FCM with the ANN model is shown in Table 3

4 Performance Evaluation

To evaluate the proposed model for clustering on the dataset, a comparison is carried out. Including cluster size in the algorithm for calculating the membership values after each iteration, to enhance the performance of *adhoc_FCM* in comparison with the existing model. To evaluate this mechanism for the proposed system configuration of the i5 processor packed with 2 GB Nvidia graphics and 8 GB RAM; further, Python is used as the programming language along with various machine learning libraries.

4.1 Dataset Analysis

A detailed analysis of the dataset is carried out, and a comparison is carried out on three datasets MNIST [13], Fashion-MNIST [14], and USPS [15]; they are used for

Table 3 Adhoc FCM with the ANN model

Input: $\hat{\Gamma}$ dataset
Step 1: for $edc = 1$ to m do
Step 2: for $y_N = 1$ to Y_N does Calculate forward propagation using fuzzy C means end for loop
Step 3: for $c_t = 1$ to C_t do Calculate FP or second layer using the FP of the first layer End for loop
Step 4: for $xg = 1$ to Xg
Step 5: if $(Y_{\text{Auto-encoder}}(P) > \text{threshold})$
Step 6: for $c = 1$ to Xg do Using the training sample to formulate $\delta_x^{(4)}$ End or loop
Step 7: for $c_t = 1$ to c_t ($s = 1, \dots, T$) do Use a global training sample to compute the End for loop
Step 8: for $y_1 = 1$ to O do Use a parameter to compute $\delta_{x_1 x_2 \dots x_O}^{(2)}$ End for loop
Step 9: for $x_p = 1$ to Xp ($g = 1, \dots, G$) do Compute Δd For $ct = 1$ to C_t ($t = 1, \dots, T$) do Compute $\Delta Y^{(n)}$ End for loop
Step 10: for $c_t = 1$ to C_t ($t = 1, \dots, T$) do Compute Δd $y_1 = 1, \dots, y_n do$ Compute $\Delta Y^{(n)}$ End for loop End for loop
Step 11: for $y_N = 1, \dots, y_N$ do Calculate Δd for $b_o = 1$ to b_p ($g = 1, \dots, G$) do Compute ΔY End for loop End for loop
Step 12: update parameters constraints $Y = Y - \alpha \Delta \dot{Y}$ $k = k - (\Delta k / N) \times \alpha$ End if statement End for loop

clustering. Fashion-MNIST is a clothing dataset. The proposed *adhoc_FCM* model evaluates these parameters. The data point can be a member of more than one cluster when using the fuzzy *c* means algorithm, where the likelihood or probability of each data point belonging to a cluster is assigned.

MNIST dataset: MNIST is a large dataset of handwritten digits to train and test to evaluate a machine learning approach.

USPS dataset: this dataset is a handwritten dataset consisting of 7298 training data 2009 testing datasets.

4.2 Comparison Mechanism

Adhoc_FCM: This model is used for performing a comparative analysis; this clustering approach uses the matrix for clustering and update process.

K-means is a data clustering mechanism, where data belongs to a specific cluster.

SEC: This algorithm is used on the manifold ranking algorithm.

MBKM algorithm is to improve the *k*-means algorithm in a batch-wise process to reduce the computation.

DEC algorithm focuses on deep learning concept, to cluster this model based on the specific distribution that discards the decoder.

IDEC: This is a deep clustering model, specifically designed to reconstruct a method to regularize this auto-encoder mechanism.

4.3 Performance Metrics

A. Mutual information (MI):

Generally, MI is defined to estimate the dependency between the variables. MI value lies between 0 to 1, 0 determines no MI, and 1 determines correlation property. A high value of MI indicates a better clustering technique.

$$MI = (J(I) + J(G)) \cdot (J(I, G))^{-1} \quad (24)$$

B. Rand Index (RI):

The measurement of similarity is compared along two different data clustering mechanisms. RI value has a range between 0 and 1. 0 determines two different clustering for a specific value and 1 determines the clustering mechanism. Higher RI determines the high efficacy of the proposed model.

$$RI = (RI - TN) / (\max(RI) - \text{Eff}(RI))^{-1}$$

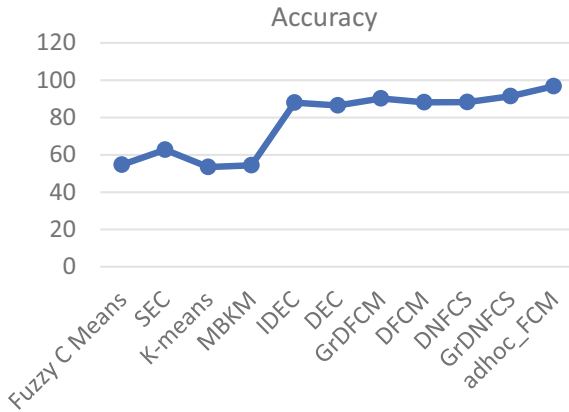


Fig. 2 Comparative analysis of accuracy for various methodologies for MNIST dataset

C. Accuracy:

This is a major feature to evaluate this model, this equation below determines the accuracy of clustering.

$$CA = g \left(\sum_{x=1}^g 1(G_k = \max(k_x)) \right)^{-1}$$

4.3.1 MNIST Dataset

Figures 2 and 3 depict the comparison of various mechanisms against the proposed mechanism for the MNIST dataset. A comparison [13] is carried out and analysis is performed on different methods based on different performance metrics. This adhoc fuzzy C means algorithm gives less accuracy of approximately 54.68%, and other methods are used to determine *k*-means and MKBM fails here with an accuracy of 53.48 and 54.43%. This adhoc system gives better performance with an accuracy of 96.82% when compared to the existing system. In comparison with other models, the metrics are determined.

4.3.2 USPS and Fashion MINST Dataset

Figures 4, 5, 6 and 7 depict the comparison of various mechanisms against proposed mechanisms for the Fashion-MNIST dataset and USPS dataset. By carefully evaluating the adhoc fuzzy *C* means is executed by comparing various metrics like accuracy, RI, and MI as well. Existing strategies like DFCM and DCFNS with good results give accuracy of 75.36 and 75.8% accuracy, respectively. Moreover,

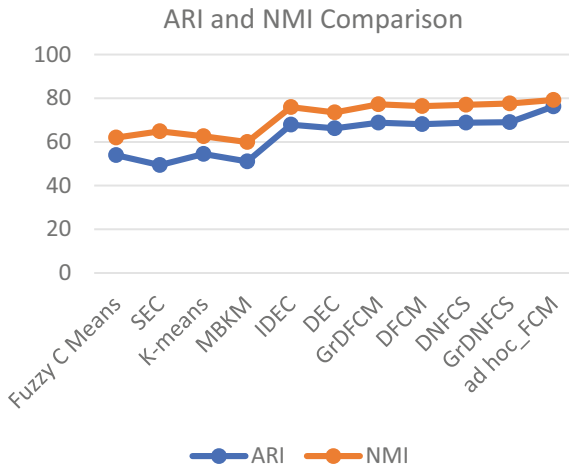


Fig. 3 ARI and NMI comparison for MNIST dataset

existing approaches GRDNFCS gives an accuracy of 76.52 whereas good accuracy of 95.12. Here, fuzzy *C* means approach achieves 85.01% accuracy. The Fashion-MNIST dataset performs comparative analysis, it is a complex dataset, comparison of different methodologies with the proposed model. This method gives a better accuracy comparison in terms of the FCM existing model. This existing model gives a better performance the model and at least improves the DFCM model. MI is considered a performance comparison metric, for fuzzy *C* means gives 51.59% and *k*-means gives 51.64%. Moreover, this model gives 66.09% to improve fuzzy *C* means achieves 67.4%.

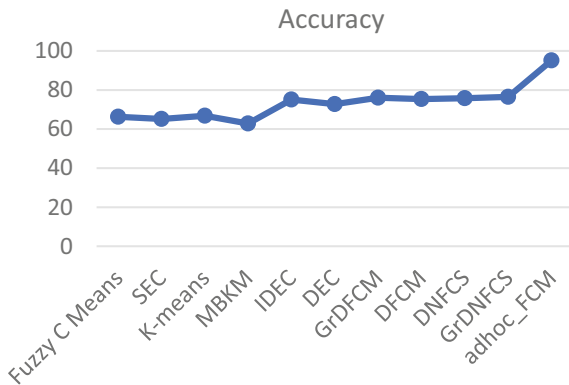


Fig. 4 Comparative analysis of accuracy for various methodologies for USPS dataset

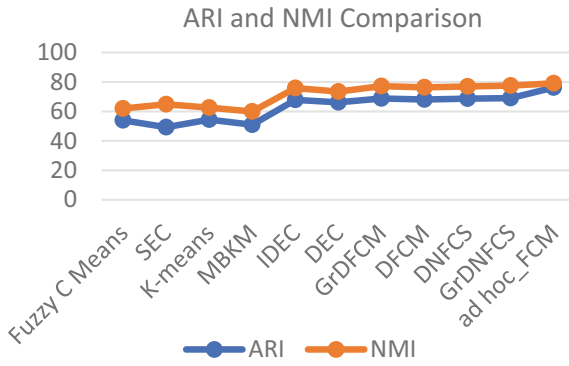


Fig. 5 ARI and NMI comparison for USPS dataset

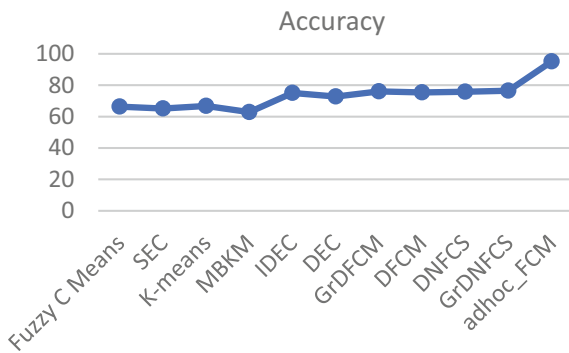


Fig. 6 Comparative analysis of accuracy for various methodologies for Fashion-MNIST dataset

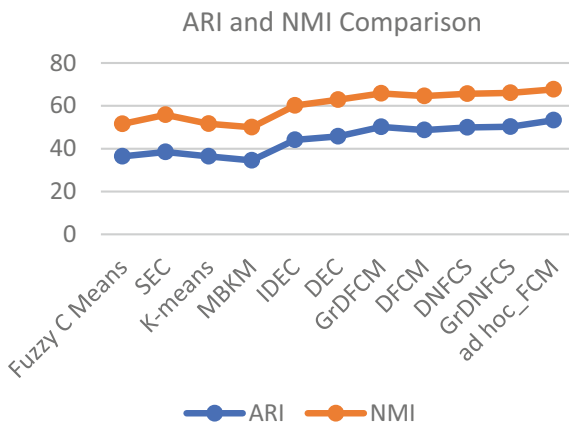


Fig. 7 ARI and NMI comparison for Fashion-MNIST dataset

5 Conclusion

In our proposed model, an adhoc model of fuzzy C means named is designed for big data clustering, generally, a fuzzy C means gives a good possibility for clustering. The time required for computation is high and performance is less giving less accuracy. Adhoc fuzzy C means consists of two different frameworks, first pass is an adhoc fuzzy C means and the second is an auto-encoder mechanism that is based on the CNN network. In addition to this, a comparison is accomplished to achieve performance metrics like RI, MI, and accuracy. In these metrics, adhoc method performs better when compared to various state-of-art methods. Clustering acts as an apprentice in machine learning prospects to analyze the data through adhoc fuzzy C means mechanism that progresses slowly when compared to another model. In the future, the study is extended to investigating different real-time data clustering mechanisms.

References

1. Bezdek J (1981) Pattern recognition with objective function algorithms. Plenum, New York, NY, USA
2. Available. <https://doi.org/10.1007/s10994-009-5103-0>
3. Agrawal D, Bernstein P, Bertino E, Davidson S, Dayal U, Franklin M, Gehrke J, Haas L, Halevy A, Han J, Jagadish HV, Labrinidis A, Madden S, Papakonstantinou Y, Patel J, Ramakrishnan R, Ross K, Shahabi C, Dan S, Vaithyanathan S, Widom J (2011) Challenges and opportunities with big data, cyber center technical reports, Purdue University
4. Ahn B (2012) Neuron machine: Parallel and pipelined digital neurocomputing architecture. In: IEEE international conference on computational intelligence and cybernetics (CyberneticsCom) 2012:143–147
5. Atluri G, Karpatne A, Kumar V (2018) Spatio-temporal data mining: a survey of problems and methods. ACM Comput Surv. <https://doi.org/10.1145/3161602>
6. Deng Y, Ren Z, Kong Y, Bao F, Dai Q (2017) A hierarchical fused fuzzy deep neural network for data classification. IEEE Trans Fuzzy Syst 25(4):1006–1012
7. Jain AK (2008) “Data clustering: 50 years beyond k-means,” in machine learning and knowledge discovery in databases. Berlin, Germany: Springer, pp 3–4
8. Yang Y, Ma Z, Yang Y, Nie F, Shen HT (2015) Multitask spectral clustering by exploring intertask correlation. IEEE Trans Cybern 45(5):1069–1080
9. Havens TC, Bezdek JC, Keller JM, Popescu M, Huband JM (2009) Is VAT single linkage in disguise? Ann Math Artif Intell 55(3–4):237–251
10. Hore P et al (2009) A scalable framework for segmenting magnetic resonance images. J Signal Process Syst 54(1–3):183–203
11. Riaz S, Arshad A, Jiao L (2018) Fuzzy rough C -mean based unsupervised CNN clustering for large-scale image data. Appl Sci 8(10):1869
12. Wu H, Prasad S (2018) ‘Semi-supervised deep learning using pseudo labels for hyperspectral image classification.’ IEEE Trans Image Process 27(3):1259–1270
13. Rajesh T, Malar RSM (2013) Rough set theory and feed forward neural network based brain tumor detection in magnetic resonance images. In: International Conference on Advanced Nanomaterials & Emerging Engineering Technologies, pp 240–244
14. <https://www.kaggle.com/bistaumanga/usps-dataset>
15. <https://nlp.stanford.edu/IR-book/html/htmledition/evaluation-of-clustering-1.html>

Covid Analysis Using Machine Learning Technique



Rohit Gussain, Nitin Kumar, Minshul Sharma, and Pooja Dehraj

Abstract The recent pandemic, covid-19 has largely affected people's lives, health, and productivity. The first case of Covid-19 was recorded on December 31, 2019, in Wuhan, China. Since then, the number of cases has increased exponentially, and subsequently, numerous precautions have been taken to prevent and cure the virus. By May 26, 2021, totally, 168 million cases were reported worldwide, with 3.49 million deaths, and the pandemic is currently underway, with people continuing to get affected and fighting for their lives from this deadly virus. The World Health Organization (WHO) has also released various precautions and vaccines to combat the pandemic, but these are insufficient to reduce the number of infected cases or save people's lives. The proposed research study discusses about the utilization of artificial intelligence (AI), machine learning (ML), and data science techniques for gaining a better understanding of covid-19 virus. This technological advancement can easily make proper judgments about covid-19, as well as the predictions on confirmed & recovered cases and deaths were made by using this technology. The datasets also include previous and current information about covid-19. The proposed research study also discusses about a tool called "Prophet." Prophet is a Facebook open-source tool, which uses the Sklearn model API. The proposed study initially creates a prophet instance and then use its fit and predict methods.

Keywords Covid-19 virus · China · Artificial intelligence · Data analysis

1 Introduction

Due to the recent covid-19 pandemic situation, people have lost their lives, health, and employment. In order to cope up with the pandemic situation, this research study

R. Gussain (✉) · N. Kumar · M. Sharma · P. Dehraj
Department of Computer Science and Engineering, Noida Institute of Engineering and Technology, Noida, India
e-mail: jhurohit@gmail.com

P. Dehraj
e-mail: drpooja.cse@niet.co.in

intends to describe how AI technology can be used to fight pandemic. Not completely, but through analysis, learning models, and algorithms, more useful knowledge can be gained on the total death cases, recovered cases, and the nature of covid-19 virus and its evolving mutants. Generally, artificial intelligence (AI) is referred as the imitation of human intelligence in machine, which is programmed to mimic human actions, and machine learning is considered as a method to perform some explorative data analysis, preprocessing, and visualization, and further data can also be forecast by using machine learning techniques. This research study proposes a novel way to extract informative data from raw data along with graphical representation and additionally prophet, an open-source tool is used to forecast the intended data. It basically follows the Sklearn model API. The utilized data will be usually collected from different platforms, and by using this data, some helpful and graphical figures of covid cases are shown in Figs. 1, 2 and 3 [1–3].

2 Literature Survey

The novel coronavirus (Covid-19) is a contagious disease that spreads primarily through nasal droplets while sniffing and through the mouth and salivation while sneezing. This disease was first reported on December 2019 in Wuhan, China. Covid-19 has become a global pandemic affecting the entire world. Numerous covid-19 predictive models have been developed by academic researchers across the globe to help them make important decisions and implement appropriate control measures. Standard methods fail to precisely predict the global effects of the pandemic due to a lack of availability of precise covid-19 records and vulnerability. To handle a similar pandemic situation in the near future, this research study presents an artificial intelligence-based (AI) meta-examination model to anticipate the pattern of covid-19 effects across the globe. Strong AI calculations, to be specific support vector machine (SVM), Naïve Bayes (NB), and linear regression (LR), have been applied on the dataset of ongoing series by holding the world record for cases announced as confirmed cases, recovered cases, deaths, and dynamic instances of the covid-19 epidemic. Measurable investigations were also carried out to provide different facts about the observed side effects of covid-19, the list of the main 20 nations impacted by the covid, and the number of cases worldwide. Among the three AI methods considered, Naive Bayes has demonstrated promising results in predicting future covid-19 patterns with lower mean absolute error (MAE) and mean square error (MSE). The lower upsides of MAE and MSE clearly demonstrate the adequacy of the Naive Bayes relapse procedure. In any case, the global impact of this pandemic remains confounded. This study uncovers various patterns and future developments of the global pandemic for individuals and states to respond proactively. This article establishes the standard for demonstrating AI's ability to anticipate the rise in covid-19 cases [4].

A pandemic without immunogenicity will result in a significant effect on human existence as well as financial and monetary frameworks throughout each



Fig. 1 Total recovered cases from covid-19 [4]



Fig. 2 Total death cases from covid-19 [4]

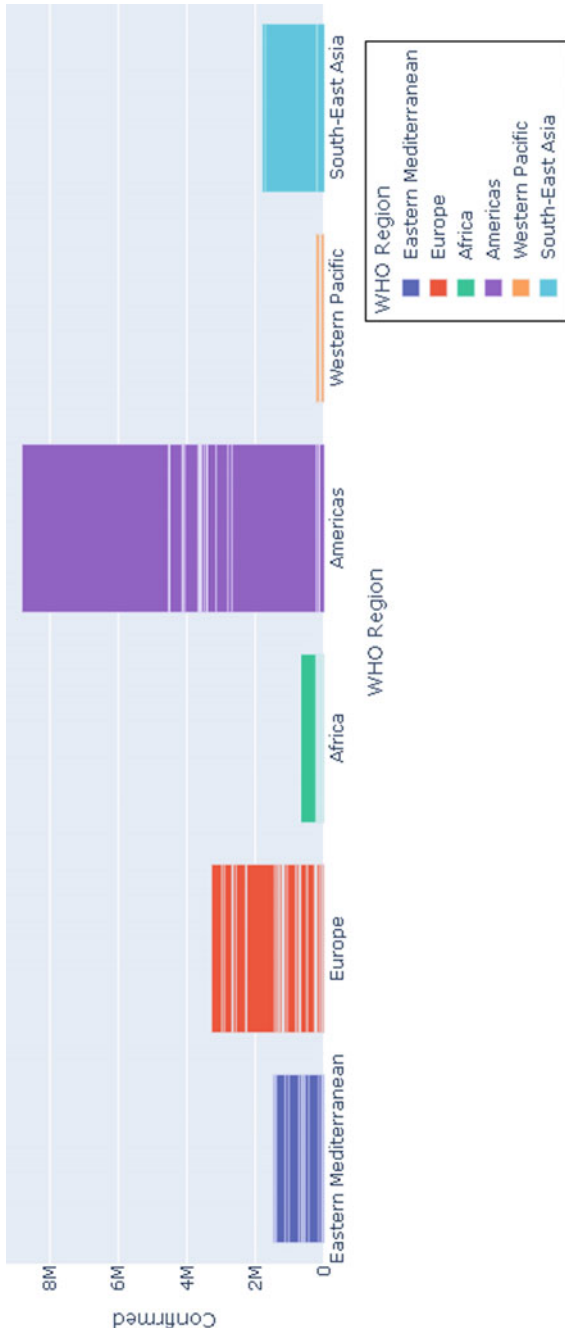


Fig. 3 Total confirmed cases from covid-19 [4]

country across the globe. The AI computations for back-handling were planned by completely unique predictive models, and their presentation was then determined and assessed. Random forest regressor and random forest classifier have outperformed the existing KNN + NCA, SVM, decision tree classifier, Gaussian and Naive Bayes classifier, multilinear regression, XGBoost classifier, and linear regression techniques [5].

It is critical to evaluate disease patterns and track contaminations through an information-driven setup that responds to pandemic in order to lay out socially strong and sustainable urban communities during pandemic situations in order to meet the requirements of protective applications that work both locally and globally. With the ongoing covid-19 pandemic, digital proximity tracing technology (DPTT) has gained a significant attention for risk minimization, regulation, recovery, and effectiveness. In this article, DPTTs and data-driven epidemic intelligence strategies (DDEISs) are contrasted, and a novel solution has been proposed. In terms of reducing the risk of disease, ensuring general well-being, security, and normalcy, a social and mechanical arrangement can be provided by incorporating the primary capacity of DDEIS. Human behavior is considered while assessing its limitations and benefits for individual arrangement and decision-making. The SEIR (Ease of Exposure-Exposure-Contagion-Recovery) epidemiological model provides preliminary data on client preferences within a DPTT. The impact of proposed model on the transmission elements of covid-19 is evaluated, and the results are presented [6].

SARSCoV2, the new covid variant that causes covid-19, has leveraged a potential setback across the globe, where patients encounter a variety of complications that necessitate medical experts to investigate innovative disease identification, treatment, and therapy plans. AI-based advancements have had a significant impact in dealing with unexplained symptoms, and various organizations have quickly accepted and modified these innovations to deal with the challenges posed by the covid-19 pandemic [7].

Covid-19 is an infectious disease that causes severe lung damage, which may even result in fatal conditions. When World Health Organization (WHO) has declared covid-19 as a pandemic, all countries have ordered to control and shut down all areas. The main reason for this lockdown is to address five distinct tasks, for example, 1. Anticipating the spread of covid-19 in locales. 2. Examine growth rates and types of declines in all countries. 3. Speculating on how the pandemic will end [8, 9].

Further, researchers propose a smart support vector regression strategy to investigate five unique coronavirus symptoms. Instead of preferring a simple review line, support vectors are used to achieve higher precision. With the available datasets, the proposed technique is tested and compared to other known review models. The promising outcomes demonstrate its superiority in terms of effectiveness and precision [10].

An analysis of studies published in 2020 related to covid-19 theme by searching Science Direct, Springer, Hindawi, and MDPI with covid-19-related terms and catchphrases. The total number of articles obtained was 16,306, but only 14 investigations on these subjects were suitable for this survey. The research findings suggest that AI could play a significant role in covid-19 investigation, hypothesis, and

preference. Overall, AI could be useful in medical services supplier program and investigating covid-19 cases. With a test precision of 92.9%, administered learning has outperformed other unaided learning calculations. Later on, regular administered examination can be used for achieving more accuracy [11].

As indicated by logical information provided by the European Center to Disease Prevention and Control, there is a fast spread of new covid variant (Covid-19) among general public, resulting in a significant rise in fatal rate. Be that as it may, the quantity of test packs accessible for covid-19 stays restricted despite the fact that the quantity of cases tends to increase every day. The implementation of a highly automated discovery framework is critical for analyzing and preventing the spread of covid-19. Chest X-rays are used to differentiate corona infection by using three major complex brain network models, specifically Inception ResNetV2, InceptionV3, and ResNet50. The ResNet50 model provides the most notable characterization precision and execution among existing frameworks. When compared to existing models, a structure based on the proposed CNN model provides enhanced clarity, responsiveness, and precision. A cumulative cross-approval is used to examine the existing models and compare to the proposed model by using disarray lattices and ROC analysis [12].

Coronavirus (COVID-19) has had a worldwide effect, and analysts across the globe are working constantly to distinguish and predict related designs. Many clinical preliminaries are in progress to track down a solution for this sickness. The unsafe and unrestrained globalization of covid-19 has prompted an extraordinary worldwide closures and leveraged a significant weight on the advanced frameworks. World Health Organization (WHO) has suggested a fast audit of accessible information to comprehend the consideration and steps required for combating covid-19. In [13], authors have recorded different AI techniques that have been utilized in the past to efficiently handle the pandemic/endemic situations like Ebola, H1N1 influenza, Zika, and norovirus. The authors have also discussed about the examination of covid-19 affected patient information to further predict the spread of covid-19 disease.

The covid-19 pandemic is the world's most serious health emergency in the last eight months. There is no predefined deadline for the end. Across the globe, more than 31 million people are contaminated as of September 18, 2020. Anticipating the covid-19 pattern has become a difficult task. Covid-19 data for the United States, Germany, and the rest of the world were obtained from the World Health Organization for this review, which range from 2020-01-20 to 2020-09-18. The informational index is compiled with week by week tested cases and 35 weeks after the cases have been confirmed. The circulation of information was then assessed by using the most recent week by week covid-19 data. Further, its boundaries were determined by the factual dispersion. Similarly, an AI-powered time series model has been proposed to detect the scourge bend and predict the pandemic pattern. AI techniques included here are linear regression, multilayer perceptron, random forest, and support vector machine (SVM). The techniques are displayed against the figures of RMSE, APE, and MAPE, and it is discovered that SVM has a dominant pattern. The global plague is expected to spread until the end of January 2021, with an estimated 80 million people infected [14].

3 Proposed Methodology

A. Data Loading

It is a method used to load data from file, folder, or any application to dataset. Data loading can be done by using a Python library called NumPy, which convert data into data frame type [7].

B. Exploratory Data Analysis

It is a crucial technique used for performing activity preliminary investigations based on information in order to determine patterns, spot irregularities, test hypothesis, and envisage assumptions with the assistance of outline statistics and graphical representations.

C. Data Preprocessing

It is a data mining technique, which involves a data transformation of raw data into an understandable term [7].

D. Data Splitting

In data splitting, the data should be divided into two parts such as training data and testing data.

E. Model

A machine learning model is a file which has been trained to recognize different patterns. A model can be developed by training different datasets into an algorithm from which it can easily learn the required information [15].

3.1 Data Mining and Data Transformation

In covid analysis, first, we used different sources to collect data. Data collection is the most important part for any predictive modeling. How much the data is qualitative and quantitative the prediction accuracy is more good. So, here are the five different stages of data mining process we used.

A. Data Cleansing

Here, different techniques are used to remove the irrelevant and inaccurate data from the input source. Data cleaning is the process in which the inaccurate data will be detected and cleaned. Additionally, this process will also remove, replace, and modify the artifacts or coarse data with globally accepted values or predictable or mean values.

B. Data Integration

The data integration cycle includes synchronizing a smart information arrangement framework. The source may include a variety of informational collections,

Web sites, or level documents. The standard execution of information incorporation creates the enterprise data warehouse (EDW), and then discusses two ideas—tight and free coupling, but the complexities should not be disregarded.

C. Data Transformation

This step necessitates the transformation of data within designs, typically from the source framework to the predicted framework. Normalization, aggregation, smoothing, generalization, and trait development are some of the techniques used in this process.

D. Concept Hierarchies

They minimize information by replacing the low-level ideas with high-level ideas. The complex information with various degrees of reflection is characterized by reasonable ordered progressions. Binning, bunch analysis, histogram survey, and other techniques are used in this process.

E. Pattern Evaluation and Data Presentation

Assuming the information as a primary source, the user and clients should make the best use of it. Following the preceding segments, the information will be presented as charts and graphs, which is then comprehended with minimal factual information.

3.2 *Machine Learning Training*

Pattern recognition is a process of understanding the information processing activities and provide decisions similar to the human brain (Artificial intelligence (AI) and machine learning (ML) techniques are designed and developed to mimic human activities). However, in computer science, pattern recognition is a technology that analyzes incoming data and information stored in a database. Moreover, pattern recognition is considered as a subset of machine learning since it employs machine learning algorithms to identify patterns.

Pattern recognition and machine learning techniques can be used to detect data features that reveal the original information about a specific dataset or system and is then distinguished by the four characteristics listed below:

- Reading from the input data
- Automatically detect patterns even if it is slightly visible
- Detect common patterns
- Recognition of different shapes and angles.

In other words, pattern recognition and machine learning are two sides of the same coin.

4 Algorithm

Here, are the two machine algorithms which are used to forecast covid cases, deaths and recovered cases.

Linear Regression

It is a supervised learning-based machine learning algorithm. Linear regression is capable of performing only the regression tasks. Linear regression is useful for determining the relationship between two or more continuous variables, one of which is a predictor or independent variable and the other is a response or deterministic variable. The main idea of this algorithm is to find the best fit line among the data in order to minimize the total prediction error. The distance between the point and regression line is defined as error.

Random Forest Regressor

It is a supervised learning-based machine learning algorithm. Random forest regression is capable of performing both classification and regression. A random forest is a collection of decision trees. This input data are routed through various decision trees. It employs a different number of decision trees during training and outputs the class of each tree, which is a phase (for classification) mode or a prediction of the definition (for regression).

5 Graphical Representation of Covid Data

The following graph is visualized after performing an exploratory data analysis, data preprocessing, and data visualization process on the given input data [5], from which the total death, recovered, and confirmed cases can be easily predicted for different regions around the world.

Using this visualization technique, the total recovered cases can be easily obtained as a data categorized by country, as shown in Fig. 1.

Using this visualization technique, the total death cases can be easily obtained as a data categorized by different regions across the globe, as shown in Fig. 2.

The given graph plots present the no. of confirmed, recovered, and death cases in different regions. Each bar is also distributed into larger and smaller sections based on which the region with maximum cases is shown in Fig. 3.

Using above heat map diagram, the most correlated features with no. of confirmed cases in world can be figured as in Fig. 4. Below-mentioned are some of the most correlated features.

- Deaths (93%)
- Active (93%)
- Recovered (91%)
- New cases (91%)

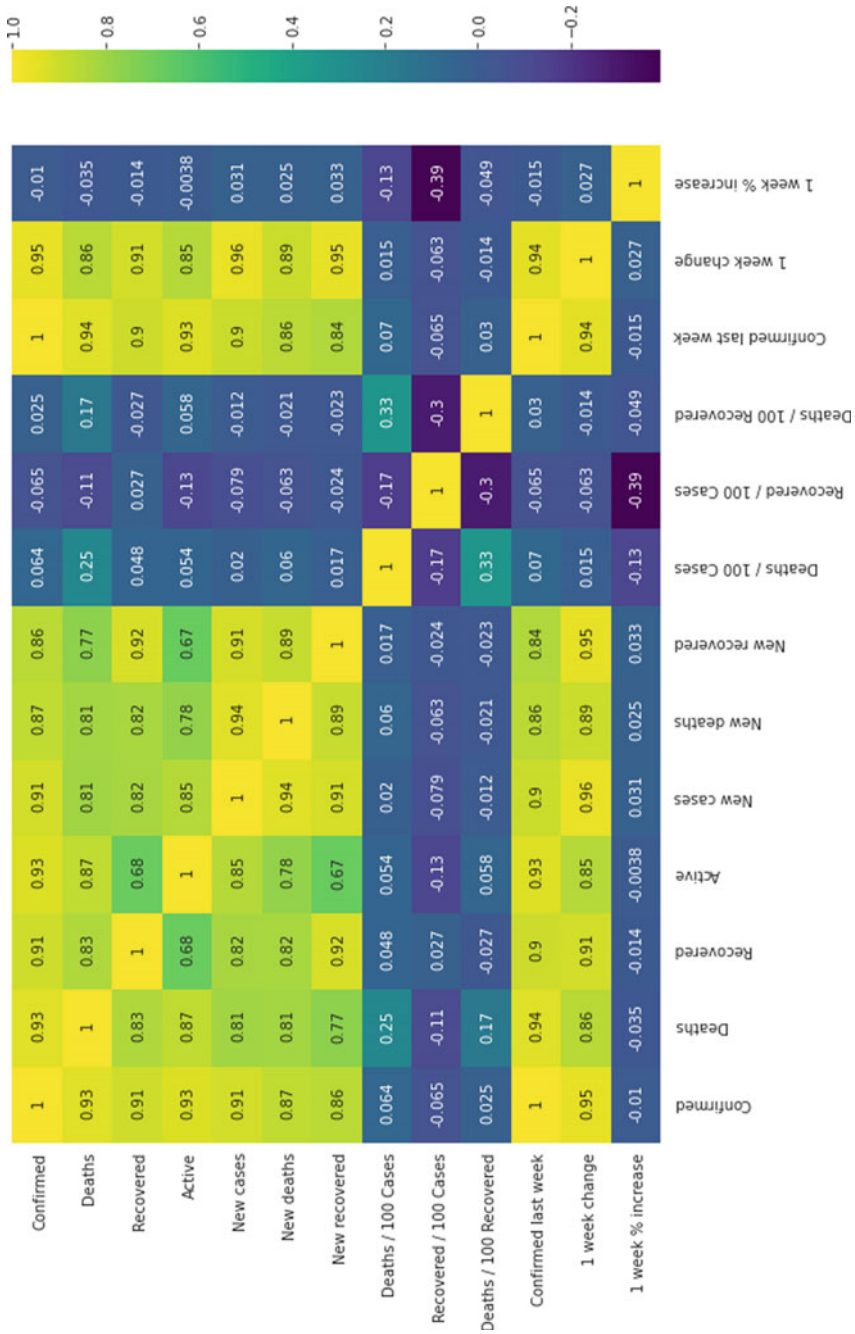


Fig. 4 Heat map diagram of covid-19 cases [4]

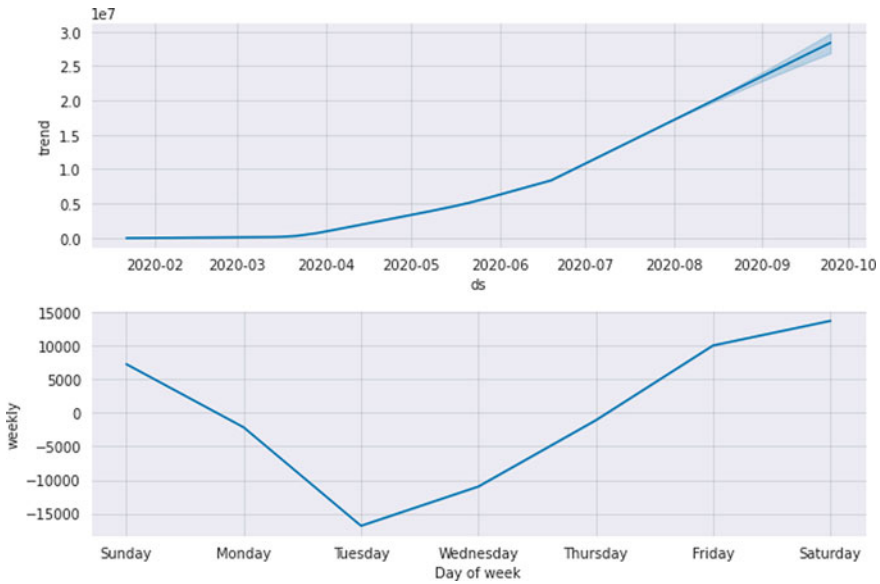


Fig. 5 Weekly graph with the amount of cases [4]

- New deaths (87%)
- New recovered (86%)
- Confirmed last week (100%)
- 1 week change (95%).

Here, the covid cases are predicted using prophet, which adheres to the Sklearn model API by using the informative data in data frame.

As there is an increased number of recovered cases, it is observed that the weekdays have the most recovered cases, whereas weekends have the fewest recovered cases, as illustrated in Fig. 5.

It is also observed that the deaths were flattening at the start of 2020, but as time passed, the death cases began to increase. Saturday has the highest number of deaths, while weekdays have the lowest number of deaths.

This graph depicts the seasonality trend of the cases on a yearly and weekly basis. From Gregorian calendar month of 2020 to October 2020, there is a gradual increase in cases. The weekly graph shows that, Saturday has the most cases with Tuesday having the low number of cases as in Fig. 6.

6 Covid-19 Coronavirus Tracker

The tracker confirms the information about covid-19 about based on the number of confirmed, death, and recovered cases by country and region. The trend shows the

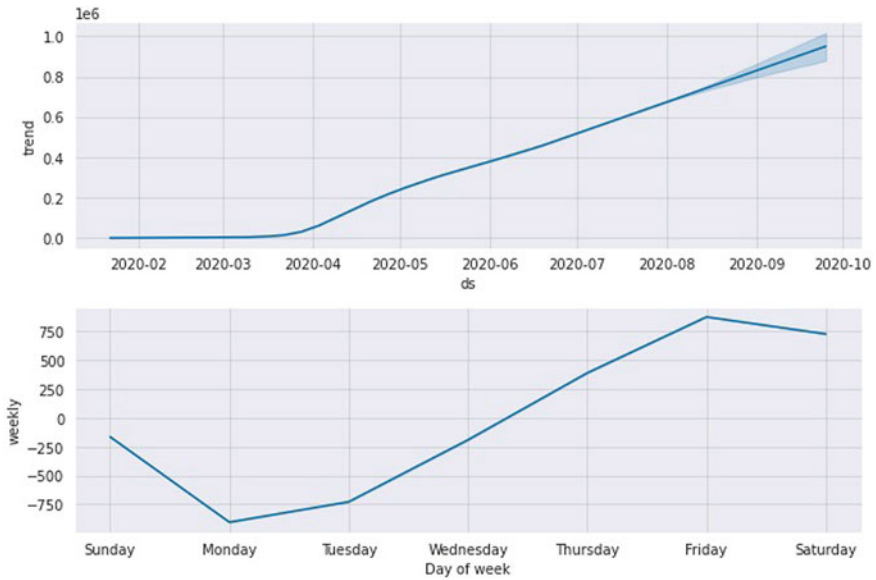


Fig. 6 Weekly amount of death cases [4]

number of confirmed, death, and recovered cases by country, and the worldwide map shows the number of confirmed, death, and recovered cases. The data are visualized by using the covid-19 map generated from the Johns Hopkins University (JHU) Coronavirus Resource Center and the World Health Organization’s reports.

This tracker is updated as per WHO norms regularly, when the new data are released [6]. Here, the reported cases are visualized in a panel of an analytics dashboard to figure out the rate of recovered, deaths, and confirmed cases in various regions and countries worldwide. This aligns with the goal of determining the current rage over the amount of time consumed in each phase. The total number of confirmed cases is increased rapidly; especially on day 24, a large number with 15,000 variations were discovered. This could be due to a change in how China expedites the number of confirmed cases worldwide (Fig. 7).



Fig. 7 Covid-19 coronavirus tracker trend. [6]

7 Conclusion

As corona is rapidly spreading throughout the world, many people have lost their lives, health, and employment. To combat the further spread of corona virus variants, the proposed study has successfully employed the machine learning techniques and Fbprophet to predict future cases by analyzing the existing covid data based on various terms. The proposed study has used the exploratory data analysis, data preprocessing, and data visualization techniques on the data collected from various platforms, and finally, the obtained informative data from which we figured out some facts in a graphical view are efficiently used to easily check the confirmed, death, recovered, and other important facts of coronavirus in different regions. Further, the proposed study has successfully predicted the death and recovered cases in various regions by using Fbprophet, an open-source tool. Raw data can be easily analyzed by using machine learning and data science techniques, and the recovered, dead, and confirmed cases can be detected all over the world by recognizing the significant features.

References

1. Zoabi Y, Deri-Rozov S, Shomron N (2021) Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Digital Med* 4(1):1–5
2. Kleitman S, Fullerton DJ, Zhang LM, Blanchard MD, Lee J, Stankov L, Thompson V (2021) To comply or not comply? A latent profile analysis of behaviours and attitudes during the COVID-19 pandemic. *PLoS one* 16(7):e0255268
3. Dashraath P, Wong JL, Lim MX, Lim LM, Li S, Biswas A, Choolani M, Mattar C, Su LL (2020) Coronavirus disease 2019 (COVID-19) pandemic and pregnancy. *Am J Obstet Gynecol* 222(6):521–531
4. Tiwari D, Bhati BS, Al-Turjman F, Nagpal B (2022) Pandemic coronavirus disease (Covid-19): world effects analysis and prediction using machine-learning techniques. *Expert Syst* 39(3):e12714
5. Prakash KB, Imambi SS, Ismail M, Kumar TP, Pawan YN (2020) Analysis, prediction and evaluation of covid-19 datasets using machine learning algorithms. *Int J* 8(5)
6. Sungheetha A (2021) COVID-19 Risk minimization decision making strategy using data-driven model. *J Inf Technol* 3(01):57–66
7. Syeda HB, Syed M, Sexton KW, Syed S, Begum S, Syed F, Yu F Jr (2021) Role of machine learning techniques to tackle the COVID-19 crisis: systematic review. *JMIR Med Inform* 9(1):e23811
8. <https://www.kff.org/coronavirus-covid-19/fact-sheet/coronavirus-tracker/>
9. <https://www.analyticsvidhya.com/blog/2020/08/exploratory-data-analysiseda-from-scratch-in-python/>
10. Yadav M, Perumal M, Srinivas M (2020) Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos Solitons Fractals* 139:110050
11. Kwekha-Rashid AS, Abduljabbar HN, Alhayani B (2021) Coronavirus disease (COVID-19) cases analysis using machine-learning applications. *Appl Nanosci* 1–13
12. Dhaya R (2020) Deep net model for detection of covid-19 using radiographs based on roc analysis. *J Innov Image Process (JIIP)* 2(03):135–140
13. Agrawal R, Gupta N (2021) Analysis of COVID-19 data using machine learning techniques. *Data analytics and management*. Springer, Singapore, pp 595–603

14. Ballı S (2021) Data analysis of covid-19 pandemic and short-term cumulative case forecasting using machine learning time series methods. *Chaos Solitons Fractals* 142:110512
15. <https://github.com/>. <https://github.com/i-am-rohit/covid-analysis-using-machine-learning/blob/main/corona.ipynb>. [Accessed on 31 Jul 2021]

Skin Lesion Classification Using Machine Learning



J. Vindhya, C. Pooja, Manisha H. Dongre, S. Gowrishankar, A. H. Srinivasa, and A. Veena

Abstract The proliferation of abnormal cells that are capable of invading or spreading to other parts of the body causes skin cancer. Early and accurate detection of skin cancer can be treated successfully. Therefore, we look into the different machine learning models that are implemented to classify the skin lesion images accurately and in a short amount of time. It is classified into benign or malignant skin lesions. Machine learning techniques have great potential in classifying the medical images for detecting the disease. The main features for this task include color, texture, and shape which differentiate malignant from benign lesions. Considering these serious issues, a variety of initial recognition methods for malignant lesions has been developed by researchers. This paper provides a comprehensive evaluation of different machine learning algorithms for skin cancer detection.

Keywords Feature extraction · Skin cancer detection · Activation function · Skin lesion · Dermoscopy image · Convolution neural network · Image segmentation

J. Vindhya (✉) · C. Pooja · M. H. Dongre · S. Gowrishankar · A. H. Srinivasa · A. Veena
Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology,
Bengaluru, Karnataka 560056, India
e-mail: vindhya0919@gmail.com

C. Pooja
e-mail: poojapoorvi81171@gmail.com

M. H. Dongre
e-mail: manishadongre2000@gmail.com

S. Gowrishankar
e-mail: gowrishankarnath@acm.org

A. H. Srinivasa
e-mail: srinivasaah.cs@drait.edu.in

A. Veena
e-mail: veenaal@acm.org

1 Introduction

Skin cancer is the most frequent type of cancer, which amounts for more than 40% of all cancer cases globally [1]. Out of these, the most dangerous is melanoma. Despite accounting for approximately 4% of all skin malignancies, over, 75% of deaths related to skin cancer is caused by melanoma [2]. Although the exact reason of all melanomas is unknown, ultraviolet (UV) exposure from the sun and tanning lamps raises the risk of melanoma. Therefore, detection and diagnosis at an early age are critical for survival chances of affected patients [3].

Tumors form when healthy cells begin to grow uncontrolled. It can be benign or cancerous tumor. A malignant tumor is one that can grow and spread throughout the body [3]. Therefore, it is necessary to classify and treat them in order to avoid spreading. Identifying benign versus malignant skin lesions accurately is key to preventing missed malignant melanoma as well as unnecessary excisions [4]. In this paper, we will be considering the different models used to classify malignant skin lesions from the benign ones. The first section covers the convolution neural networks (CNNs) models used on different dataset, with varying parameters. The second section deals with the support vector machine (SVM) algorithms for classification of skin lesions. The third section talks about the various other supervised learning algorithms used for classification and the final section discussing the other machine learning models. The various feature extracting models and their combinations give better results. The accuracy and performance metrics are compared and discussed.

2 Models Utilized for Classification

2.1 Convolution Neural Network

Convolution neural network (CNN) model is seen efficient and a go-to approach for image classification problems as it can automatically detect the important features of the image and classify it. The authors in the paper [5] use this model for the classification of input image as cancerous and non-cancerous lesions as shown in Fig. 1. They have considered ISIC 2017 dataset which contains 7 different types of skin lesions for classification including actinic keratosis and basal cell carcinoma. The data are augmented first to prevent overfitting. The convolution layer is used for feature extraction using filters (kernel). The features obtained from the convolution layer are fed through non-linear activation functions. Rectified linear unit is the activation function used here. The pooling layer decreases the input volume through reducing the number of parameters and sends it to the layer which is completely connected for the classification purpose in which the softmax activation function was used. Overfitting problem was solved by the use of dropout function. The initial learning rate was assigned as 0.0001, and for every 10 epochs, 0.1 was multiplied

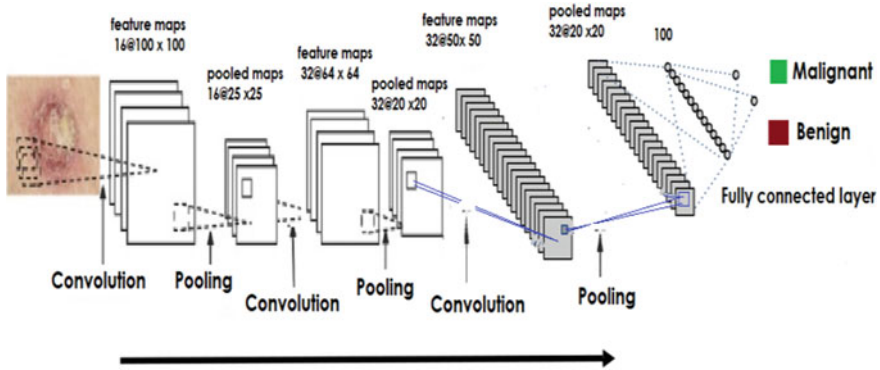


Fig. 1 CNN architecture [5]

to the learning rate until it reached 100 epochs. In case of no decrease in the loss function for seven successive epochs, early stopping was implemented in the model, and the learning rate was lowered. The accuracy kept increasing with the decreasing loss for consecutive epochs. The optimization algorithm used was Adam optimizer, and the training accuracy of 0.78 was obtained. The test loss was 0.621 with the test accuracy of 0.768.

This paper [6] proposed the use of CNN model for classification. Since the neural network is trained in an iterative manner, the loss function should be minimized to make it more accurate. To do this, the model needs to be trained by varying the activation functions and optimizers for each epoch of the input image. Later, the model is tested using testing dataset and yields the output as either cancer or rashes image. The model was trained with ReLU, sigmoid, tanh activation functions individually and at last with all three activation functions combined. The accuracy of 80.2% was obtained, while all the three activation functions were integrated whereas 50% accuracy was obtained when only ReLU was used, and the model was less than 50% accurate when tanh activation function was used, and the model exhibited highest accuracy when all three functions were combined, and the ability of the model to predict the correct output was more when the three functions were combined rather than using them individually.

Hasan [7] uses convolution neural network (CNN) to categorize the images into benign and malignant in this project. There are many symptomatic checklists available; for this project, they have chosen ABCDE checklist—(A) Asymmetry—the portion of the cell which is affected will turn into a tumor and that portion will not have any similarity with the other half, (B) Border—the mole’s border will be uneven. (C) Color—shade of the affected area isn’t uniform. (D) Diameter. The cell width will be 6 mm and above. (E) Evolution—difference between the previously stated changes and the present state. In this project, they have selected the data set from ISIC dermoscopic archive. The images used to train the model are “Benign” and “Malignant.” Steps that are followed to train the model are as follows: preprocessing the data: Size of the image will be huge, and there are 3 channels of images

(Red, green, and blue). Due to lack of computational capacity, the size of the image is reduced, and only, one color channel is used, i.e., gray scale. Then, these preprocessed data are fed to the convolution layer, where the image is passed through filter. For example, if the image is 6×6 grayscale image and the filter will be 3×3 . A 3×3 matrix is selected from the image matrix and will be multiplied with the filter matrix and a 4×4 matrix is produced as output. In general, $(x - y + 1) \times (x - y + 1)$ are the result by convolving an input of $x \times x$ with a $y \times y$ filter. An example is shown in the Fig. 2. In this operation, the image will be shrunk. Compared to the pixel at the center of the image, the pixels at the corner are utilized only once. So, to overcome this, they have added extra border to the original image making it 8×8 matrix. Now, the convolution operation will result in original image of a size 6×6 matrix. The model is trained up to 200 times; as there is no noticeable change in loss, they concluded that as the iteration increases, accuracy of the model also increases.

Xiangfeng Dai [8] proposes an on-device interference app with a pre-trained classification model which is stored on a device. In this approach, the characteristics of each category are learnt by the convolution neural network without supervision. The image taken from the user’s device is fed into the trained CNN model for classification into different categories. A sequence of convolution layer is followed by pooling, and sub-sampling layers, process the input in the CNN architecture. Then, there are fully linked layers that allocate the input image’s probability based on the predefined attributes. After the convolution layer, the max pooling layer, MaxPool2D is used to prevent overfitting. After these layers, the multidimensional array is reduced to a

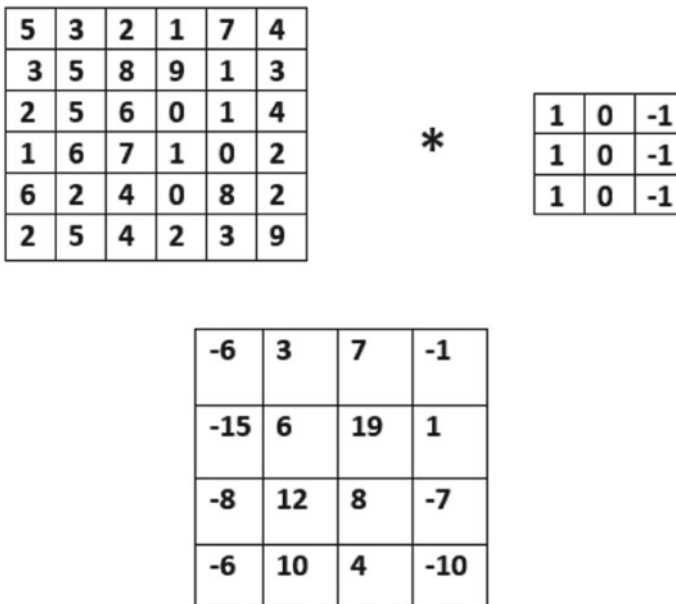


Fig. 2 6×6 image with 3×3 filter resulting in 4×4 image [7]

two-dimensional array using flattening and fully connected layers. Softmax function is used in output layer responsible for calculating the probability of each skin lesion class. It gives an accuracy of 75.2% with a loss of 0.71.

This paper [9] proposes a multiclass skin cancer classification method which classifies the HAM10000 skin cancer dataset into seven cancer classes. To make the photos compatible with the MobileNet training model, the resolution was reduced to 224×224 . They used a highly accurate pre-trained MobileNet conventional neural network. MobileNet usually has 93 layers, but they dropped the last 5 layers, and softmax activation function for prediction was used. Except for the last 25 layers, which were trained, the weights of all layers were kept constant. They used 70/30 split of dataset for testing and training with 30 epochs. Here, precision, accuracy, and recall are used for evaluation of the model. For each of the seven classes, the weighted average precision was 0.90. The precision calculated for the melanocytic nevi type of skin cancer was the best at 97. When compared with other studies, this model got the highest categorical accuracy of 97.07. This model is lightweight and predicts within 2–3 s. It was observed that the accuracy increases and loss decreases as the number of iterations are increasing.

2.2 Support Vector Machine (SVM)

Monica [1] considered the ISIC 2019 dataset for the classification which includes 8 varieties of classes combined into 800 images in which 200 images are allotted for each class. The image quality is improved by eliminating the undesired parts of the image in the data preprocessing step. The unwanted hair in skin lesion is erased using the dull razor method. The region of interest is separated using segmentation. Color-based k -means clustering is used in this model. The two methods used in this model are ABCD and GLCM where the features are extracted from the images, and the results are stored in an Excel sheet. Features like diameter, standard vector, asymmetry index, autocorrelation, homogeneity, and contrast are yielded for further classification purposes. Here, SVM method is used for multiclass classification. The dataset is divided into training dataset that includes 70% of images, and the rest is testing dataset. The accuracy of 96.25% and precision of 96.32% were achieved.

In this paper [10], the authors propose a method that uses image processing and machine learning approaches to classify benign lesions from malignant lesions using dermoscopic images. The dataset is from the ISIC Web site and contains the ISBI-2016 challenge, which includes RGB dermoscopic images, labels, and segmentation ground truths. The image undergoes initial processing to make them more meaningful. The images are resized; noise is removed; RGB is converted into gray scale, and hair is removed. Then, it undergoes image segmentation to capture the affected region. For this, they have used the OTUS threshold algorithm. Then, the pre-processed images undergo feature extraction. The features extracted are texture, shape, and color. The dataset is imbalanced. Nearly, 80% of the data in the ISIC dataset is classified as “benign,” whereas 20% is classified as “malignant.” So, the data

balancing is done using the synthetic minority over sampling (SMOTE) technique. As a result of this procedure, data are 50% benign and 50% cancerous. The characteristics are then normalized in a range of 0 to 7 and scaled using the min–max mapping algorithm. They are classified as benign or malignant, for which they have considered 3 classifiers including random forest, quadratic discriminant, and SVM (Medium Gaussian). As the random forest classifier gives remarkable results, it is considered. The accuracy of this model is 93.89% is the highest when compared to SVM with 88.17% accuracy and quadratic discriminant with 90.84%.

Vidya and Maya [11] have chosen the dataset from ISIC dataset. The dataset was projected to preprocessing, segmentation, and feature extraction before undergoing classification. The features were extracted from the dataset using 3 different methods which are ABCD, GLCM, and HOG. In the ABCD method, the asymmetry, border irregularity, color, and diameter of the lesion are extracted. Textural analysis is done in the gray-level co-occurrence matrix (GLCM) method that gives the features like contrast, homogeneity, and shade. Histogram of oriented gradients (HOGs) extract the information about shape and edge of the lesion. The different classification of model along with the hybrid feature extraction led to the higher accuracy of the model. It is noticed that SVM classifier is the most accurate with an accuracy of 97.8% and AUC of 0.94. The accurate result is obtained after augmentation performance.

2.3 Supervised Machine Learning Algorithms

For skin cancer image classification, Ali [12] utilized the algorithm multiple feature construction with ensemble classification (MFCEC) with genetic programming (GP). Two real-world skin lesions image datasets, Dermofit Image Library and PH2, are used in the proposed model. The dataset is partitioned, for training, into nine folds for and one-fold for testing using tenfold cross-validation. Feature extraction is done by converting the images into feature vectors before the training data is separated from the test data. Each image will have 5 feature vectors as shown in Fig. 3, namely LBP extracted from RGB color channel images (LC), LBP extracted from gray image (LG), global color variation features (CV), lesion geometrical shape features (S), and frequency-based wavelet features (W). The GP then uses the training set to create many features in a single GP individual. Each tree in a GP individual is considered as one constructed feature (CF). These CFs are sent to SVM, J48, and RF algorithms as input, as these algorithms have shown the best performance among other models. Finally, 3 trained classification models were obtained. The test accuracy is chosen from among these three models with the highest accuracy. The accuracy of the model is 98.03 on PH2 dataset and 85.20 on Dermofit image library.

Qian and Zhou [13] have classified the skin lesions as melanoma, nevus and BCC. First, the noise in the image is removed by preprocessing. This is done by using the algorithm of Adam Huang (for hair removal), and other noises are removed using e Gaussian filter and Wiener filter. Then, OTUS method is used to obtain the lesion area. Then, the preprocessed images undergo feature extraction. This is done on

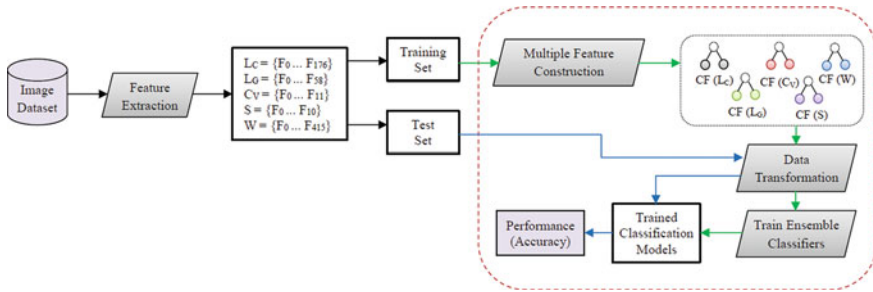


Fig. 3 Workflow of MFCEC method [12]

the basis of ABCD rule, and it was found that the texture features combined with color features are more suitable as classification basis. Using KNN classification and MDCA method in the feature fusion phase, the model gives the most accurate value and reduces the total time consumption. The proposed model has an accuracy of 75.86%.

Nour [14] developed a model for classifying the lesion into benign and malignant by using the HAM10000 dataset. Three types were chosen from the dataset, basal cell carcinoma, dermatofibroma under cancerous and benign keratosis-like lesions as a non-cancer type. The data augmentation is carried out to avoid overfitting. The model used here is VGG19 which is an updated version of VGG16 whose architecture is mentioned in Fig. 4. It contains a series of convolutional and pooling layers together termed as feature extractors, and further, a fully connected layer is present which is called as classifier. Then, it is followed by a softmax activation function for the classification of images into one of three types. By using the VGG19 model, transfer learning was applied. The features are easily located by using the autonomous feature extraction ability of VGG19 which saves the time spent on inspecting them manually. Adam is the optimization function used here. The 100 epochs were used for training the network with batch size 50 and the learning rate 0.01. The metrics for evaluation were accuracy and loss, and 600 images were used in testing the network. The model was 97.5% accurate by the end of the 100 epochs with loss function 0.1185. In addition, the stable network was observed between epochs 60 and 70.

2.4 Additional Classification Models

Kreutz [3] classified the lesion based on ABCD rule. Physical characteristics marginally differentiate between melanoma and benign skin lesions; a blend of these characteristics is necessary for accurate diagnosis. The image of a skin lesion is considered after cleaning, and the output border is determined, which is termed as segmentation. Here, the hybrid method is employed for segmentation. Once the image is segmented, the extraction of features takes place. The asymmetry percentage

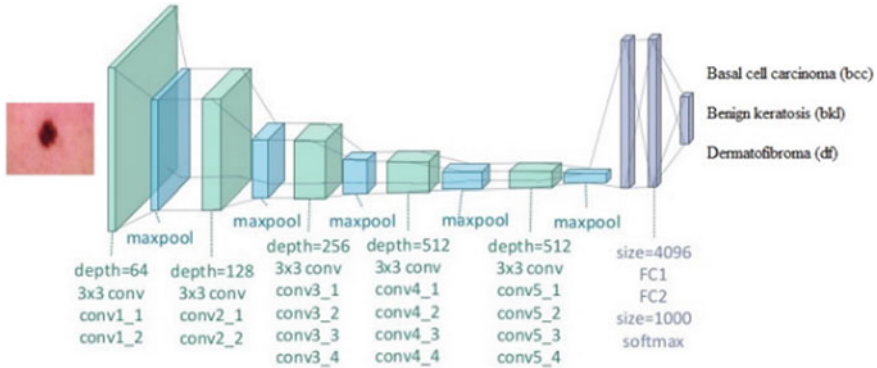
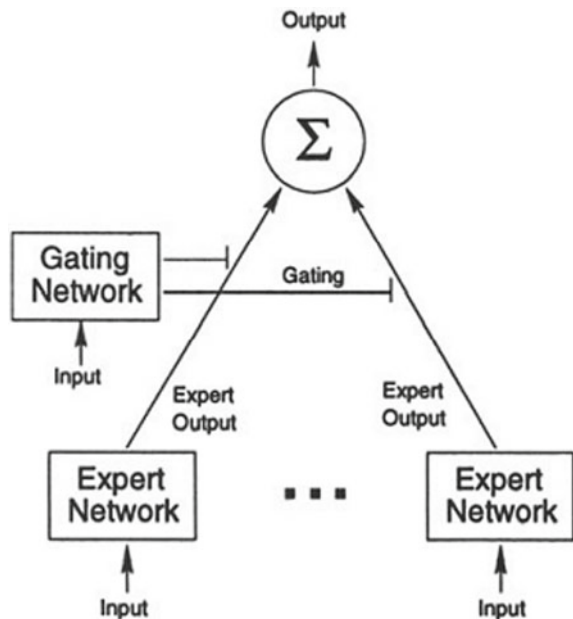


Fig. 4 VGG19 general architecture [14]

and border irregularity of a tumor is estimated. The variation in color is determined by whirling together of red, black, and brown. Here, the mixture-of-experts (MOE) architecture is used instead of a fully connected network where the problems is divided into sub-problems and follows the divide and conquer principle as shown in Fig. 5. The specificity and sensitivity were the metrics considered. According to the results, the mixture-of-experts did not perform better than the multi-layer perceptron in classifying the lesion as benign or malignant. Overall, when compared to the multi-layer perceptrons, MOE exhibits to be less overfitting.

Fig. 5 Architecture of MOE [3]



The authors of this study [15] present an approach for rapidly classifying and interpreting reflectance confocal microscopy (RCM) images. The method entails the creation of a visual vocabulary using speed up robust features (SURF), vanilla neural network classifiers, and their interpretation systems. It is divided into four steps for the classification task and two stages for the interpretation task. The four stages of classification steps include augmentation, feature extraction, modeling, and classification. Augmentation of the initial images in the dataset of 136 RCM is done using contrast limited adaptive histogram equalization and non-local means denoising in the first stage, yielding 408 images in the augmented dataset. SURF and Haralick features are the two distinct descriptors used during feature extraction. SURF algorithm uses a fast Hessian detector detects and assigns a 64-dimensional vector for each point of interest. Haralick features can help improve the accuracy of configuration built on SURF features. The RCM image is represented as an m -dimensional vector with $k + 14$ values in the modeling step. The 14 Haralick values are concatenated with the histogram of k values formed from the interest areas in each image for the subsequent phases. A vanilla neural network with $k + 14$ input neurons and two output neurons perform binary classification with softmax as output layer activation function at a learning rate of 5 for the ADAM optimizer. The visual pattern weighted localization (VPWL) conducted during the modeling stage and the visual word weight calculation (VWWC) performed during the classification stage are the two steps of the interpretation process, as shown in Fig. 6. The K-means clustering variations called the K-means plus-plus and K-means canopy are used in the experiments. The models are tested with various featured extracted and loss function. The highest accuracy of 91.17 is obtained by K-means canopy method with SURF and Haralick features and cosine proximity loss function.

This study [16] employed a histopathology image set that was gathered and processed from hospitals in China. For training and testing, the photos have been downsized to 640×480 pixels in RGB color space. The dataset contains 1167 photos of skin diseases ranging from malignant tumors to benign tumors to inflammatory dermatoses. The dataset is uneven in terms of the various categories, so data augmentation operations are used in each epoch to resolve these issues. In this deep residual attention network, first, a convolution layer is employed. The convolution layer acts as a feature extraction layer with kernel size of 7×7 . Then, a MaxPooling layer is used. Further, for each DRA module, one residual unit is connected in front of it to enhance the information it learns. In the CAM module, the last dense layer comes after a global summation pooling (GSP) layer which is introduced. To enrich the levels of features, squeeze-excitation attention modules are layered and coupled in a residual fashion in each DRA module. Three branches of the SEA modules work in parallel: the trunk branch, the spatial attention-oriented mask branch, and the channel attention-oriented mask branch. Trunk branch performs the main part of feature extraction and processing. It produces a feature map from which the spatial attention—orientation mask branch captures region of interest. Squeeze and excitation operations are used to model interdependencies between channels in the channel attention-oriented mask branch. The CAM module, along with a GSP layer and a dense layer, is added right before the final output layer, resulting in the desired

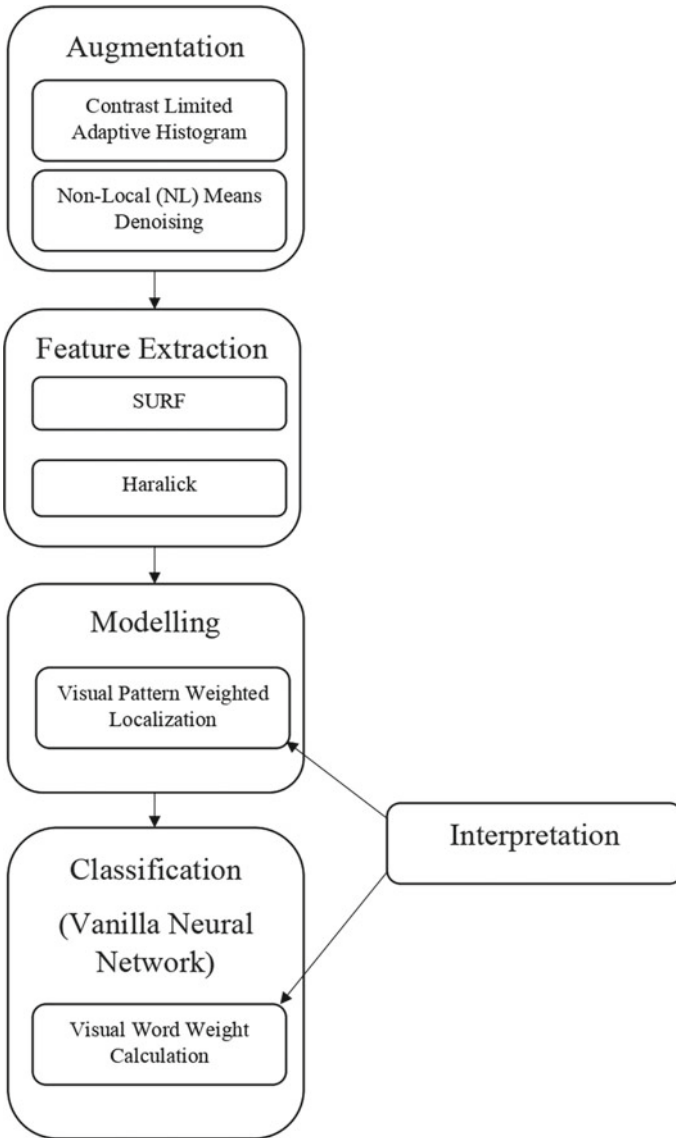


Fig. 6 Stages of classification

output. The accuracy of DRANet is calculated to be 0.868 and is compared with various other lightweight models such as MobileNet and EfficientNet. Comparison of different models for classification of skin lesions is shown in Table 1.

Table 1 Comparison of different models for classification of skin lesions

No	Model	Advantage	Disadvantage	Performance metrics
1	VGG19 [14]	Autonomous feature extraction capability	–	Accuracy: 0.975 and loss: 0.1185
2	CNN(ReLU, tanh, sigmoid) [6]	Reduced the initializer weight dependency and obtained highest accuracy	Requires more no of training datasets to improve efficiency	Accuracy: 80.2%
3	SVM (ABCD, GLCM) [1]	Very precise	Highly complex	Accuracy: 96.25% and precision: 96.32%
4	MoE [3]	Less overfitting than multi-layer perceptron	Performance is less than multi-layer perceptron	Specificity and sensitivity
5	MFCEC using GP [12]	Outperformed the commonly used classifiers such as NB, SVM, and KNN, on PH ² dataset	–	Accuracy: 98.03
6	Random Forest Classifier [10]	Gives better results as compared to SVM and Quadratic Discriminant	–	Accuracy: 93.89
7	KNN–MDCA method [13]	Reduced misclassification due to combined per-processing techniques to remove noise	–	Accuracy: 75.86%, Precision: 76.18%, Recall: 72.22%
8	K-means Canopy (SURF and Haralick) [15]	Higher accuracy for SURF + Haralick features with negative log likelihood loss function	Dataset is limited, model interprets skin folds as a feature and misclassification can incur	Accuracy: 91.17
9	MobileNet (SoftMax) [9]	High categorical accuracy, lightweight model, and prediction are done in 2–3 s	Needs more iterations for accuracy to increase	Accuracy: 97.07%, Recall: 91.34%
10	DRANet [16]	Light and relatively less computational costs	–	Accuracy: 0.868

3 Conclusion

An effective model is required for the classification of malignant melanoma from the benign ones since there is a drastic increase in the melanoma cases in recent years. Identification of lesions in the initial stages plays an important role in curing skin cancer. This paper covers the various models used for skin lesions classification. While building a model that classifies the skin lesion, a greater number of training dataset images is needed to increase the efficiency of the classification, but some types of skin lesions do not have enough images required for the machine learning model. It is significant that comparing the different classification models is difficult. It results in different outcomes, upon using the same dataset with the same classification problem.

Acknowledgements All the authors would like to acknowledge that this research work was supported in part by the Karnataka State Council for Science and Technology (KSCST), an autonomous organization under the Department of Science and Technology, Government of Karnataka, India, under the Student Project Program (SPP) scheme.

References

1. Monika MK, Vignesh NA, Kumari CU, Kumar MN, Lydia EL (2020) Skin cancer detection and classification using machine learning. *Mater Today* 33(7) 4266–4270
2. He K, Sun J, Tang X (2013) Guided image Fltering. *IEEE Trans Pattern Anal Mach Intell Jun*(2013):1397–1409
3. Kreutz M, Anschütz M, Gehlen S, Grünendick T, Hoffmann K (2001) Automated Diagnosis of Skin Cancer Using Digital Image Processing and Mixture-of-Experts. In: Handels H, Horsch A, Lehmann T, Meinzer HP (eds) *Bildverarbeitung für die Medizin 2001*. Informatik aktuell. Springer, Berlin, Heidelberg
4. Ali A-RH, Li J, Yang G (2020) Automating the ABCD rule for melanoma detection: a survey. *IEEE Access* 8:83333–83346
5. Nour A, Boufama B (2020) Convolutional neural network strategy for skin cancer lesions classifications and detections. In: 11th ACM international conference on bioinformatics, computational biology and health informatics, BCB '20, Article 91, pp 1–9
6. Subha S, Wise DCJW, Srinivasan S, Preetham M, Soundarlingam B (2020) Detection and differentiation of skin cancer from rashes. In: 2020 International conference on electronics and sustainable communication systems (ICESC), pp 389–393
7. Hasan M, Barman SD, Islam S, Reza AW (2019) Skin cancer detection using convolutional neural network. In: 2019 5th International conference on computing and artificial intelligence, ICCAI '19, pp 254–258
8. Dai X, Spasić I, Meyer B, Chapman S, Andres F (2019) Machine learning on mobile: an on-device inference app for skin cancer detection. In: Fourth international conference on fog and mobile edge computing (FMEC), pp 01–305
9. Younis H, Bhatti MH, Azeem M (2019) Classification of skin cancer dermoscopy images using transfer learning. In: 15th International conference on emerging technologies (ICET), pp 1–4
10. Javaid A, Sadiq M, Akram F (2021) Skin cancer classification using image processing and machine learning. In: 2021 International Bhurban conference on applied sciences and technologies (IBCAST), pp 439–444

11. Vidya M, Karki MV (2020) Skin cancer detection using machine learning techniques. In: 2020 IEEE international conference on electronics, computing and communication technologies (CONECCT), pp 1–5
12. Ain QU, Al-Sahaf H, Xue B, Zhang M (2020) A genetic programming approach to feature construction for ensemble learning in skin cancer detection. In: 2020 Genetic and evolutionary computation conference, GECCO '20, pp 1186–1194
13. Qian Y, Zhao S (2020) Detection and recognition of skin cancer in dermatoscopy images. In: 2020 International conference on pattern recognition and intelligent systems (PRIS 2020), Article 20, pp 1–5
14. Abuared N, Panthakkan A, Al-Saad M, Amin SA, Mansoor W (2020) Skin cancer classification model based on VGG 19 and transfer learning. In: 3rd International conference on signal processing and information security (ICSPIS, 2020)
15. Kallipolitis A, Stratigos A, Zarras A, Maglogiannis I (2020) Explainable fully connected visual words for the classification of skin cancer confocal images: interpreting the influence of visual words in classifying benign vs malignant pattern. In: 11th Hellenic conference on artificial intelligence (SETN 2020), pp 67–73
16. Jiang S, Li H, Jin Z (2021) A visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis. *IEEE J Biomed Health Inform* 25(5):1483–1494

IOT-Based Weather Monitoring and Weather Prediction Using ML



Brijesh Vadalia, Nisarg Thakkar, Shalin Tilva, Pratham Thakkar, and Ritiksha Modi

Abstract A perfect local weather forecasting system can be greatly helpful to the lives of local people like farmers, people living around coastal regions, etc. Besides this it also provides accurate prediction of weather at local level so that any worse condition can be dealt with and possible to plan ahead of time and take the required precautions to avoid it. The circuit implemented, primarily makes use of a NodeMCU, for data collection it requires sensors like temperature, pressure, humidity, and rainfall. Also, it exhibits ML models for the prediction of the weather with great accuracy. Here with the help of this model it has the function to collect data and ML for prediction (for prediction we require a dataset of basic parameters like temperature, pressure, humidity, and rainfall) so that prediction model can be more accurate at local level and be cheaper. The Internet of Things (IoT) is the use of the Internet to interact with physical objects. Weather information can be accessed from a remote location via an Internet of Things (IoT)-based weather information system on any device, including Windows, Android, iOS, and Web sites.

Keywords Humidity sensor · Temperature sensor · Rainfall sensor · Pressure sensor · NodeMCU (ESP8266)

1 Introduction

Weather conditions estimating is the act of anticipating the condition of the air for a given area of various climate boundaries. Weather conditions gauges are made by social event information about the present status of the climate [1]. Precise weather forecasting is essential in guarding individuals and property from dangers of cyclones, floods, and extreme weather conditions. Business, agriculture, the food

B. Vadalia (✉) · N. Thakkar · S. Tilva · P. Thakkar · R. Modi
Department of Information Technology, Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India
e-mail: brijeshvadaliya1@gmail.com

R. Modi
e-mail: ritikshamodi.dit@charusat.ac.in

industry, airports, and naval structures all rely heavily on precise weather conditions. In agriculture, earlier data about weather conditions assists ranchers with taking important choices to further develop their harvest yields. Air terminals [2] or maritime frameworks require constant climate information to be aware in the event that there is an unexpected change in climatic circumstances. For wind homesteads to govern the operation of wind turbines throughout the wind power age, precise breeze speed predictions are essential. Weather conditions have benefited from the rapid progress of inventions such as the Internet of Things, Wireless Sensor Networks, and Cloud Computing as we enter the Big Data era [3]. Big data technologies aid in the more precise prediction of future climatic states [4]. Weather and environmental forecasts can also be made more accurate and precise with the advancement of deep learning algorithms and correct information representation methodologies [5]. Consequently it is sensible to utilize profound ways to deal with important data from climate information. Profound learning methods use layers of brain organizations to recognize and remove significant examples from the datasets.

The regularly applied measurable models incorporate Autoregressive Moving Average mode (ARMA), Autoregressive integrated moving average (ARIMA), and its variations [3]. Man-made consciousness models are additionally ordered into AI indicators and profound learning indicators. AI and profound learning models are better for taking care of nonlinear datasets.

This task will focus on advancement of the Thing Speak, an IoT platform [6] that can show the information of the sensor. The strategy is separated into two sections which are equipment and programming improvement. The equipment advancement includes the circuit development and fosters the model [5]. In the interim, the product part includes the IoT coding, circuit schematic chart, circuit reproduction and information securing.

This assignment is made to diminish the obligation of individuals and to recognize the limits like Temperature, Humidity, Rain, and Pressure. Resulting in distinguishing those limits depending upon the circumstance the ARM controller will take an appropriate action [6]. Sensors are the most important modules in this task. This whole model can be put wherever [7, 8]. This model has three Sensors as a data contraption to distinguish the Weather condition and depending [9] upon the limit assessed the controller will take a fitting action.

2 Literature Survey

It will use suitable sensors attached to the NodeMCU to measure various climate boundaries such as temperature, pressure, and water level, and it will use remote technology to provide continuous information.

They employed sensors for the live data that was present at the time in this paper [10], as well as a sensor panel, a power unit with rechargeable and detachable batteries, and an LCD display (with help of Arduino Mega). Weather stations are built to collect quantitative data about the weather conditions in a specific location.

Weather conditions in an environment must be monitored in today's society. It is deemed crucial due to the severe weather. Every day is unique. The goal of this research is to improve weather stations. This can be found on the website of the IoT platform.

In this paper [6], The Raspberry Pi platform is used to describe one possible weather station system in this study. Information on the current temperature, humidity, and pressure in the atmosphere. The relative altitude and air pressure are both measured. After the data is received it is processed by transmitting data from the measurement point to a distant location. A User Datagram Protocol (UDP) server and sensor data gathering are incorporated. It was implemented on the Raspberry Pi platform. An android application that performs requests and displays the results.

In this paper [11], Climate change and environmental monitoring are two topics that have gotten a lot of press in the last few years. Management has recently received a lot of attention, and an integrated information system (IIS) is regarded to be quite useful. With a case study on regional climate change and its ecological effects, this article introduces a new Internet of Things (IoT) IIS (IoT), Geoinformatics [remote sensing (RS), geographic information system (GIS), and global positioning system] Cloud Computing, Geoinformatics (GPS), as well as e-Science for environmental monitoring and management.

In this paper [12], Weather has a big influence on people's lives. Weather variations can affect a large group of peoples including agriculture and transportation. The major goal of this project is to track and report meteorological conditions so that informed decisions may be made ahead of time and the appropriate steps can be taken to prevent disaster damage [13–16]. On the basis of the current facts, we create projections. Models that are now available are unreasonably pricey. It will make local area surveillance more difficult, in contrast to ours. It is possible because it will be less expensive.

The Internet of Things (IoT) is a network of interconnected gadgets that can send and receive data without involving humans. Checking the climatic boundary [5] is expected to assess the present status of the climate so that the best living decisions can be made based on the information gained from the gadget. In today's world, various environmental characteristics are used to create a variety of pollution monitoring systems [17]. The existing system model is a web-based weather monitoring and reporting system that allows you to collect, process, analyze, and display your observed data.

The wireless sensor network administration paradigm is described in this paper [11], which includes a device, router, gateway node, and management monitoring center [18]. After assessing and packing data from the wireless sensor network into Ethernet format, the gateway node delivers it to the server after receiving [2].

3 Existing Method

The current weather conditions observing frameworks for the most part utilize weather conditions stations that utilize different instruments, for example, thermometers, indicators, wind vanes, downpour measure and so forth to quantify climate and environment changes. The greater part of these instruments utilizes basic simple innovation which is later genuinely recorded and put away in an information base.

As shown in Fig. 1, First in data acquisition is used for digitizing data from the world around us so it can be displayed. The process of putting raw data into a comprehensible format is known as data preparation. Model training is the phase to train data to learn from. And last visualization in that we have to visualize data.

Weather conditions at the time of observing systems in the field are typically made up of unusual and huge equipment with numerous moving components that necessitate constant maintenance and should be checked and altered on a regular basis [19]. Because these instruments are typically located far from a power supply, one of the most important requirements is electricity. This raises the cost of using such devices. The use of thermometers to determine outside temperature is still in use, but they should be physically examined for any temperature changes on a regular basis. Instrument data should be physically transported from the equipment to a computer or computer via a link. Existing system consists of huge and weighty instruments that consume a great deal of room thus making it challenging to introduce them in far off areas and spots which have restricted space. The instruments utilized in the current system are costly and amount to the generally significant expense of establishment and support.

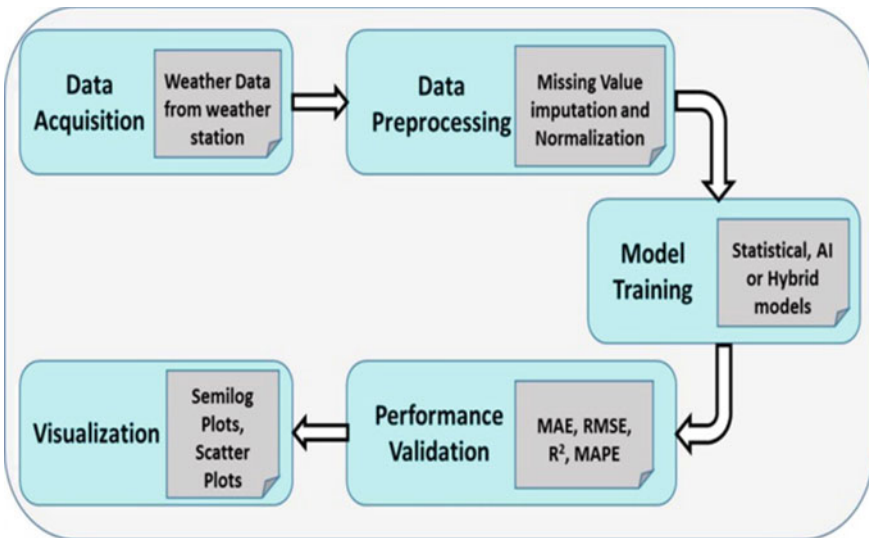


Fig. 1 Flowchart of existing method for data collection and prediction

4 Proposed Architecture/Methodology

There are high end systems available these days for weather monitoring systems. However, these systems are implemented on a large scale, for checking the continuous climate for an entire city or state.

As shown in Fig. 2, The NODEMCU microcontroller acts as the system’s main processing unit, and all sensors and devices can be attached to it. The microcontroller can control the sensors and retrieve data from them, as well as do analysis on the data. The processed data can be posted and kept in a website to serve as a database using NodeMCU.

In an implanted framework, a microcontroller is a conservative coordinated circuit designed to oversee a specific activity. A typical microcontroller consists of a single chip with a CPU, memory, and input/output (I/O) peripherals.

A microcontroller is embedded in a framework to control a specific function in a device. It accomplishes this by using its native processor to decode data from its I/O peripherals. The microcontroller saves the ephemeral data it receives in its information memory, from which the processor retrieves it and uses instructions stored in its program memory to unravel and apply the incoming data. It then uses its I/O peripherals to communicate and authorize the fitting operation. Carrying out such a framework for a little region is not achievable, since they are not intended for itself and the upward trend for keeping up with such a framework for a little area is astoundingly high. Our proposed framework utilizes 3 sensors to quantify the climate/climate factors, for example, temperature, rain, and pressure sensor. The qualities perused from the sensors are handled by the NodeMCU and stored in a text file which is used to derive analysis. All these requirements [20] are fed into the database and these values are essentials and recorded over time. These recorded boundaries are fundamental and differ from spots to places. This large number of necessities are taken care of into the information base and these values are basics and recorded after some time. Involving these qualities as info we can plot a weather

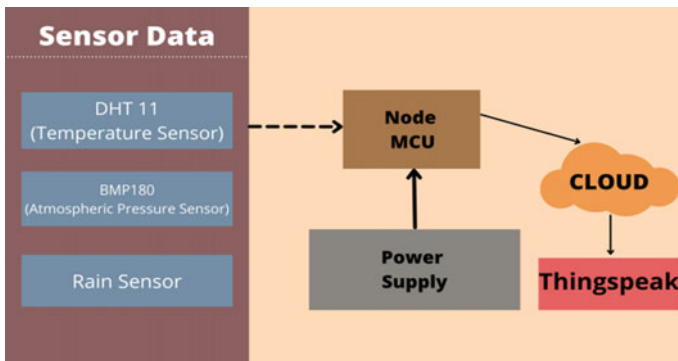


Fig. 2 Flow chart of proposed architecture

conditions diagram of a specific region over the long haul. In view of the current weather conditions, factors and present qualities the set activities are finished.

ThingSpeak comes with a Web Service (REST API) that allows you to collect and store sensor data in the cloud and build Internet of Things applications.

ThingSpeak is a cloud-based IoT analytics tool for aggregating, visualizing, and analyzing live data streams. ThingSpeak visualizes data posted to ThingSpeak in real time.

The NodeMCU ESP8266 Wi-Fi module uploads humidity and temperature values to the Cloud at regular intervals. Humidity and temperature measurements can be graphically shown on the ThingSpeak platform from anywhere in the world thanks to cloud. Pictorial Diagram of system framework is shown in Fig. 3.

Hardware specifications are shown in Table 1 and the operating voltage to the sensors are shown in Table 2.

Mode wise explanation of the system in shown in Fig. 4. Flowchart for Weather Forecasting is shown in Fig. 5. In Fig. 9 NodeMCU consists of a Wi-Fi module and when connected or when provided with power supply it first writes the data and then

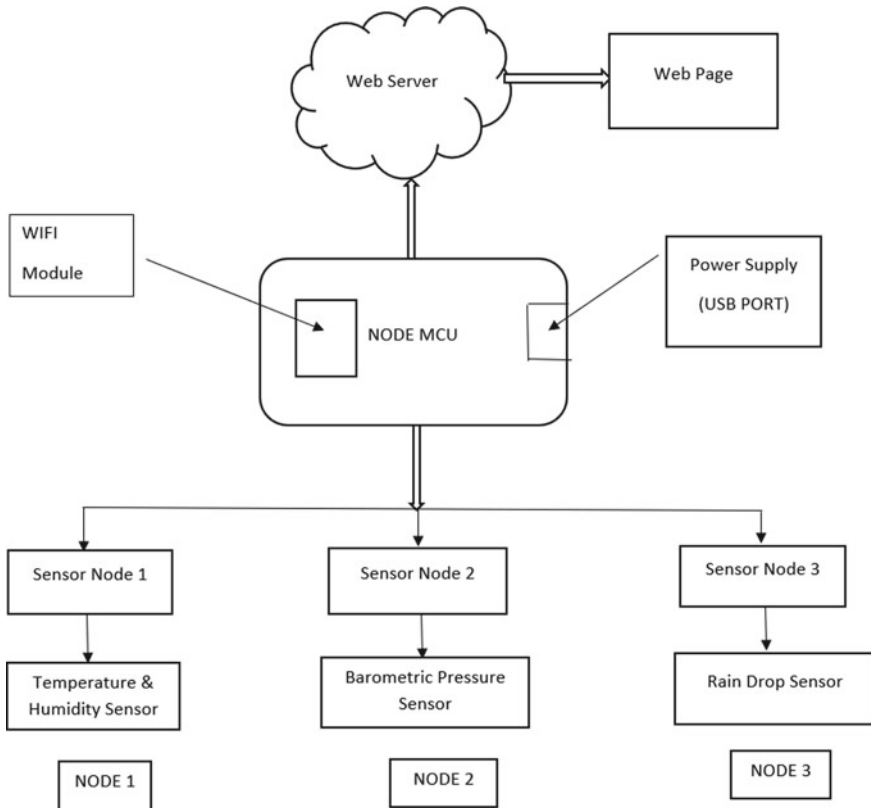


Fig. 3 Pictorial diagram of system framework

Table 1 Contains small description of hardware used in our circuit

Device overview	
Hardware	Use
Breadboard	Building temporary circuits
NodeMCU (ESP8266)	GPIO, PWM, and network provider module (able to connect to any nearby Wi-Fi network)
DHT11	To get temperature and humidity
BMP180	To get absolute pressure
Rain sensor	To detect the rainfall

Table 2 Contain the supply voltage requires by each sensor to operate

Devices and its supply voltage	
Device	Supply voltage (V)
DHT11	3.5–5.5
BMP180	1.8–3.6
Rain sensor	3.3–5
NodeMCU	5–12

it will read the data and that data is then displayed in the webpage through the web server. Mode wise explanation is there in Fig. 10.

At the initial stage we first fit all the sensors to the breadboard and then connect all the sensors to NodeMCU. Then we write the code in Arduino ide and then import the related library of Arduino. And upload the code on Arduino ide. Then it will check the necessary conditions. And then the process goes on and it will end.

To complete all of the operations in a project, the programming section is quite vital. The whole code is provided at the end, as is customary. Begin by incorporating all of the necessary libraries and specifying all of the variables.

```

// #include <SFE_BMP180.h>
#include <Wire.h>
#include <ESP8266WiFi.h>
#include "DHT.h"
    
```

Mode 1	If mode 1 presents, DHT11 i.e., temperature and pressure, if it is on, then it captures the temperature value and represents it on the webpage
Mode 2	If mode 2 presents, BMP 180 i.e., it shows the barometric pressure of the environment.
Mode 3	If mode 3 present, the Rain sensor module i.e., it shows the rain value of the environment.

Fig. 4 Mode wise explanation of the system

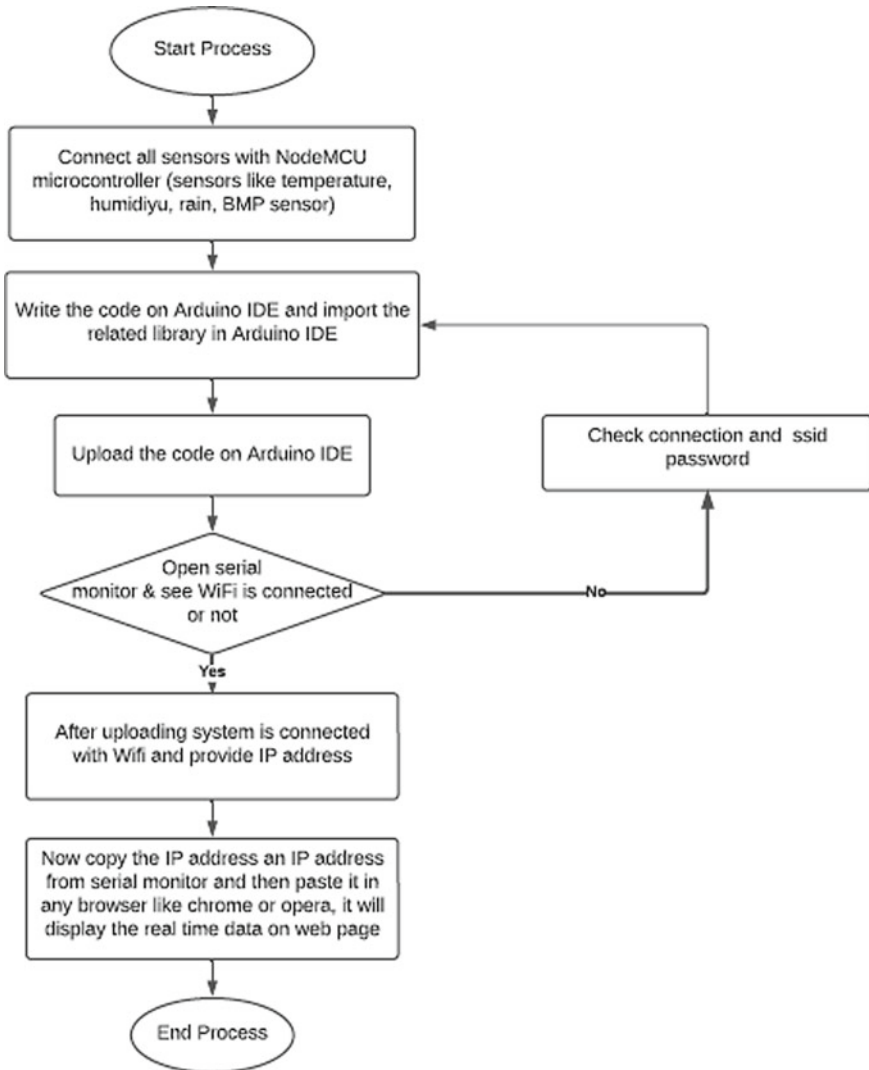


Fig. 5 Flowchart for weather forecasting

The SSID and password—that are expected to associate the NodeMCU to the web are presently characterized. Then the ThingSpeak account certifications are characterized, including the channel number and compose API, which were recently recorded.

```
String apiKey = "7SWU2QX265Q7VGG1";  
const char *ssid = "FREEWIFI";  
const char *pass = "123456780";  
const char* server = "api.thingspeak.com";
```

Fig. 6 NodeMCU (ESP8266)

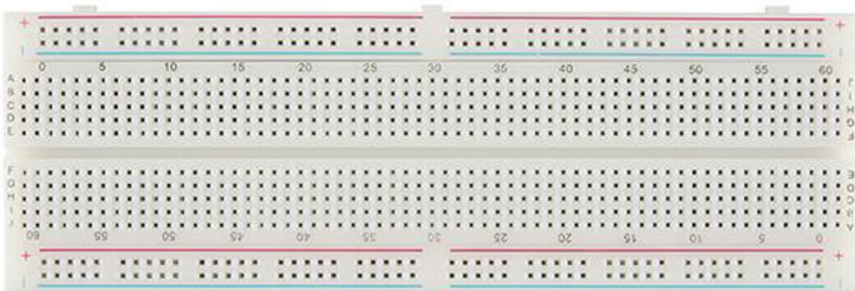


Fig. 7 Breadboard

Then call Wi-Fi to connect NodeMCU to the internet. Begin by passing the network SSID and password as parameters to it. Check for a successful Wi-Fi network connection. `status()`, and then print a message on Serial Monitor with the local IP address after a successful connection.

4.1 Hardware Requirements

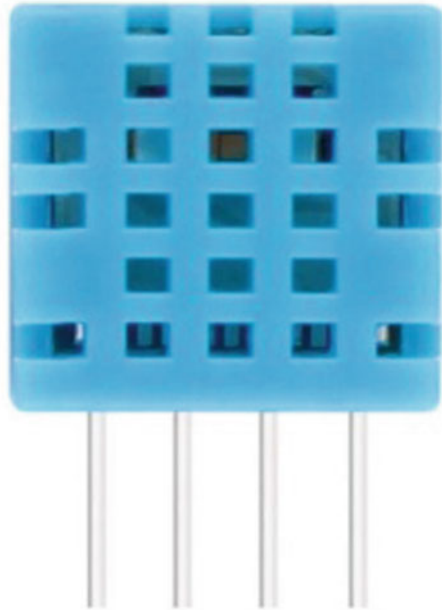
NodeMCU: The Node Micro Controller Unit (NodeMCU) is an open-source programming and development environment based on the ESP8266, a low-cost System-on-a-Chip (SoC) as in Fig. 6.

Breadboard: A breadboard is a rectangular board with many mounting holes as in Fig. 7.

4.1.1 Sensors Used

DHT11: This sensor senses temperature and mugginess as in Fig. 8.

Fig. 8 DHT11



Rain Sensor: A rain sensor is made out of a downpour recognition plate with a comparator who oversees insight as in Fig. 9.

BMP-180: It is concerned with the fundamentals of air weight. Our surrounding air has a certain weight and specified strain, which BMP180 can measure as in Fig. 10.

Fig. 9 Rain sensor



Fig. 10 Barometric pressure sensor (BMP180)



5 Implementation and Result (Simulation)

The capacity of “weather conditions checking framework utilizing IOT”, the plan of this framework is to monitor weather patterns and at the same time estimating ecological factors through the web of things.

In light of the outcomes that are gotten by the sensors send and show to Thing Speak for client seeing. This will make checking climate boundaries all the more effective with the Wi-Fi association. This framework will begin and Thing Speak will start showing sensor information by diagram and furthermore this information can be examined in Thing Speak.

With remote observing organization gadgets, individuals can check online on the page the weather pattern to make specific strides and issues even in the most pessimistic scenario for observing the climate boundaries. With every one of the information additional weather conditions announcing framework is to observe unquestionably climate to defeat the main variable deciding farming endeavor achievement or disappointment and have facilitated the client from getting incorrect estimates from Forecast Company for their place.

Thing Speak is an open information stage for the Internet of Things. It sends information to the cloud. Utilizing this we can dissect and picture our information. At last based on those we can respond or set off an activity. It gives continuous information assortment and different gadgets and innovations like Particle Photon, Raspberry Pi, Twitter, Electric Imp, Esp, and so on.

As shown in Fig. 11, It is the Physical model of our system. Primarily all the sensors are connected with breadboard and the all the sensors is connected with NodeMCU with help of jump wire and when we provide power supply the data with be fetched with the help of NodeMCU and it will be shown in the website.

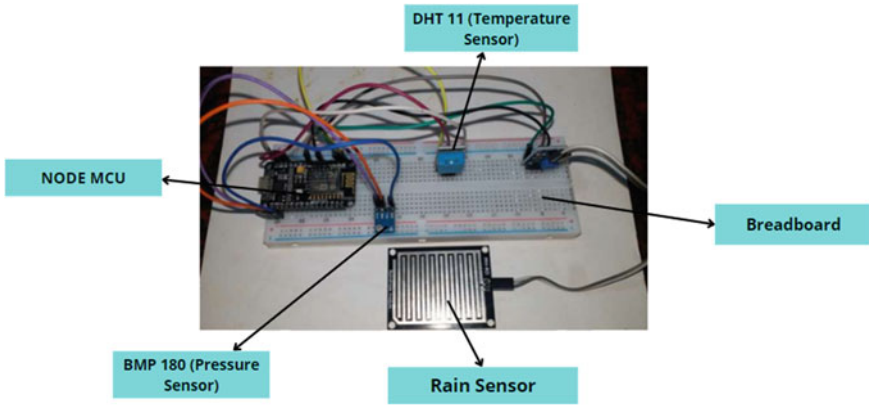


Fig. 11 Implementation and result

As shown in Fig. 12, We have connected our circuit to the ThingSpeak Server, with the help of that it is showing our data in a graph format for different attributes such as Humidity, Temperature, and Pressure. Data process flow is shown in Fig. 13.

As shown in Fig. 14, we have used a Power BI tool for better data visualization. It shows maximum/minimum temperature along with the rain prediction for tomorrow.

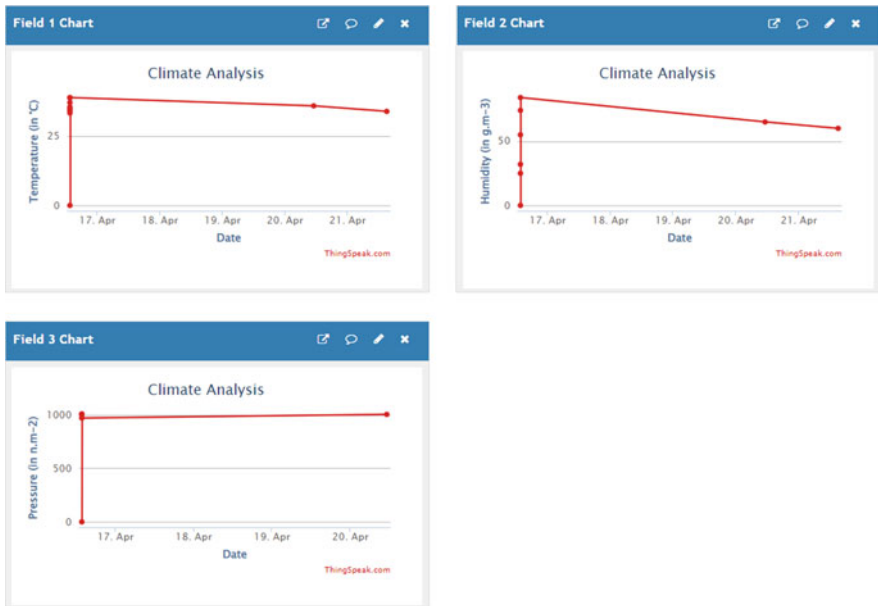


Fig. 12 ThingSpeak website

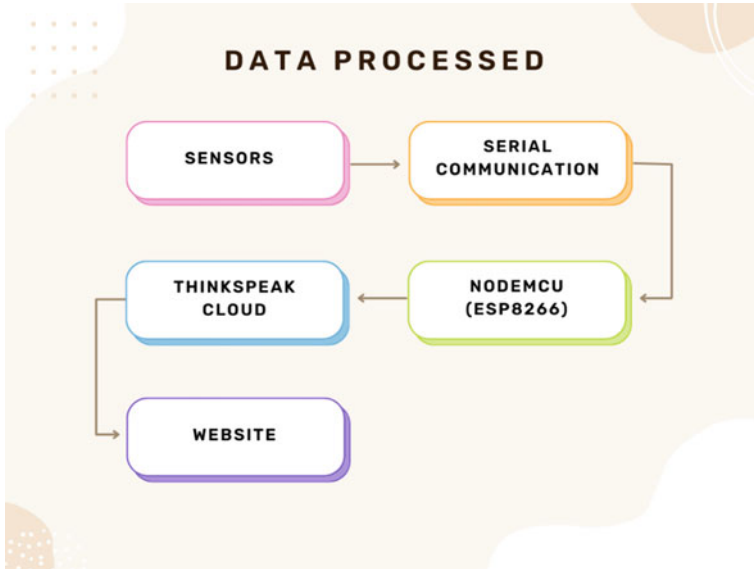


Fig. 13 Data processed

Along with this we have also added a map pointing to the rainfall location. It is fully responsive.

As shown in Fig. 15, We have created a website containing a form, in which users need to enter this required field and once the user clicks on the predict button it fetches the data from the database and displays the result.

6 Future Scope and Conclusion

Including new sensors to monitor other ecological boundaries, such as soil PH sensors, CO₂ sensors, and oxygen sensors, while allowing present sensors to be replaced if a broader range of estimations is required. Integration of additional observation devices, such as a Wi-Fi camera, to monitor the development of horticultural objects. Furthermore, the data can be consistently sent to the web server.

This project’s major purpose is to analyze live data which is collected from various Sensors like DHT11 (temperature), BMP180 (barometer pressure), Rain Sensor and sent to ThingSpeak to display the live data. After collecting a large amount of data, we perform predictions of rainfall by the help of our ML model. For high accuracy, additional sensors are employed to analyze more parameters. After collecting the data, we can display it on Power BI for future analysis.

In order to predict rainfall, we have used random forest classifier model which works on the algorithm of dividing our data into number of trees and training each

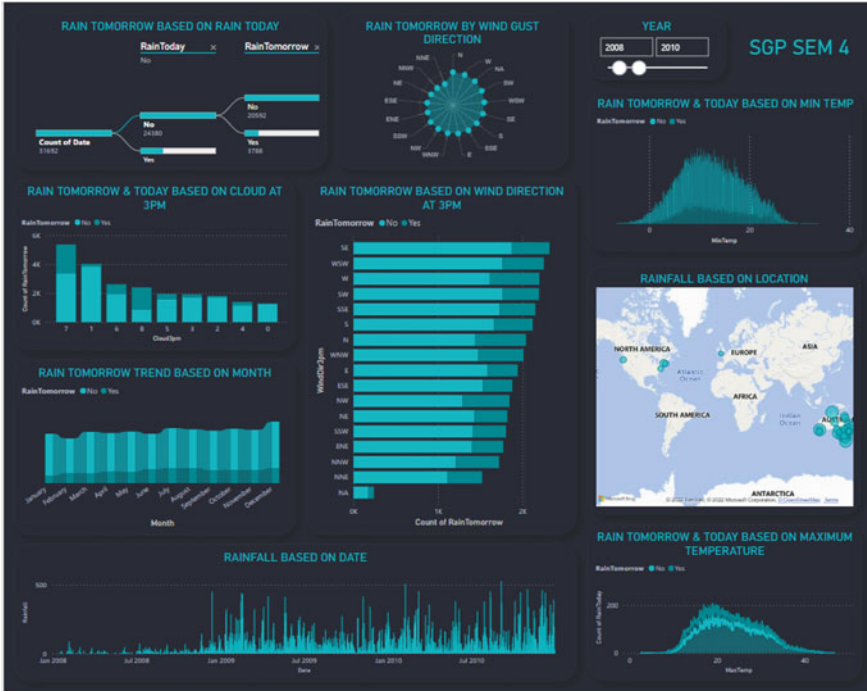


Fig. 14 Power BI

Predictor

Date <input type="text" value="dd-mm-yyyy"/>	Minimum temperature <input type="text"/>
Maximum Temperature <input type="text"/>	Rainfall <input type="text"/>
Evaporation <input type="text"/>	Sunshine <input type="text"/>
Wind Gust Speed <input type="text"/>	Wind Speed 9am <input type="text"/>
Wind Speed 3pm <input type="text"/>	Humidity 9am <input type="text"/>
Humidity 3pm <input type="text"/>	Pressure 9am <input type="text"/>
Pressure 3pm <input type="text"/>	Temperature 9am <input type="text"/>
Temperature 3pm <input type="text"/>	Cloud 9am <input type="text"/>
Cloud 3pm <input type="text"/>	Location <input type="text" value="Select Location"/>
Wind Direction at 9am <input type="text" value="Select Wind Direction at 9am"/>	Wind Direction at 3pm <input type="text" value="Select Wind Direction at 3pm"/>
Wind Gust Direction <input type="text" value="Select Wind Gust Direction"/>	Rain Today <input type="text" value="Did it Rain Today"/>

Fig. 15 Website

Table 3 Information about different classification algorithms

Accuracy (%)	83	82.6	85
Parameters used	Temperature, humidity, and light intensity	Max and min temperature, humidity, and wind speed	Max and min temperature, rainfall, wind direction, and wind stream
Algorithm used	Logistic regression	Decision tree	Random forest classifier
How to implement	<pre> From sklearn.linear_model import LogisticRegression reg = LogisticRegression() reg.fit(X_train, y_train) y_pred = reg.predict(X_test) </pre>	<pre> From sklearn.tree import DecisionTreeClassifier dtc = DecisionTreeClassifier() dtc.fit(X_train, y_train) y_pred = dtc.predict(X_test) </pre>	<pre> From sklearn.ensemble import RandomForestClassifier rfc = RandomForestClassifier() rfc.fit(X_train, y_train) y_pred = rfc.predict(X_test) </pre>

of them. After predicting each of the tree maximum voted one is selected. It has the accuracy of 85%, which is more precise compared to other models. Information about different Classification Algorithms are tabulated in Table 3.

References

1. Vanmathi C, Mangayarkarasi R (2020) Real time weather monitoring using internet of things. In: 2020 international conference on emerging trends in information technology and engineering (ic-PETITE). IEEE, pp 1–6
2. Aram S, Troiano A, Pasero E (2012) Environment sensing using smartphone. In: 2012 IEEE sensors applications symposium, SAS 2012—proceedings, pp 110–113
3. Folea SC, Mois G (2015) A low-power wireless sensor for online ambient monitoring. IEEE Sens J 15(2):742–749
4. Princy SE, Nigel KGJ (2015) Implementation of cloud server for real time data storage using Raspberry Pi. In: 2015 online international conference on green engineering and technologies (IC-GET), Nov 2015. IEEE, pp 1–4
5. Shete R, Agrawal S (2016) IoT based urban climate monitoring using Raspberry Pi. In: 2016 international conference on communication and signal processing (ICCSP), Apr 2016. IEEE, pp 2008–2012
6. Savić T, Radonjić M (2015) One approach to weather station design based on Raspberry Pi platform. In: 2015 23rd telecommunications forum (TELFOR), Belgrade, pp 623–626
7. Baste P, Dighe D (2017) Low cost weather monitoring station using Raspberry Pi. Int Res J Eng Technol (IRJET) 4(05)
8. Ugale N, Navale M (2016) Implementation of IoT for environmental condition monitoring in homes. Int J Eng Appl Technol (IJFEAT)
9. Ram KSS, Gupta ANPS (2016) IoT based data logger system for weather monitoring using wireless sensor networks. Int J Eng Trends Technol (IJETT) 32(2)
10. Kusriyanto M, Putra AA (2018) Weather station design using IoT platform based on Arduino Mega. In: 2018 international symposium on electronics and smart devices (ISESD), Bandung, pp 1–4

11. Fang S et al (2014) An integrated system for regional environmental monitoring and management based on the internet of things. *IEEE Trans Ind Inform* 10(2):1596–1605
12. Parashar A (2019) IoT based automated weather report generation and prediction using machine learning. In: 2019 2nd international conference on intelligent communication and computational techniques (ICCT). IEEE, pp 339–344
13. Varghese L, Deepak G, Santhanavijayan A (2019) An IoT analytics approach for weather forecasting using Raspberry Pi 3 Model B+. In: 2019 fifteenth international conference on information processing (ACINPRO). IEEE, pp 1–5
14. Kodali RK, Rajanarayanan SC, Boppana L (2019) IoT based weather monitoring and notification system for greenhouses. In: 2019 11th international conference on advanced computing (ICoAC). IEEE, pp 342–345
15. Vivek Babu K, Anudeep Reddy K, Vidyapathi CM, Karthikeyan B (2017) Weather forecasting using Raspberry Pi with internet of things (IoT). *ARNP J Eng Appl Sci* 12(17):5129–5134
16. Susmitha P, SowyaBala G (2014) Design and implementation of weather monitoring and controlling systems. *Int J Comput Appl* 97(3)
17. Bindhu V (2020) Design and development of automatic micro controller based weather forecasting device. *J Electron Inform* 2(1)
18. Fan S, Mao CX, Chen LN (2005) Peak load forecasting using the self-organizing map. In: *Advances in neural network—ISSN 2005*, pt III. Springer-Verlag, New York, pp 640–649
19. Saini LM, Soni MK (2002) Artificial neural network-based peak load forecasting using conjugate gradient methods. *IEEE Trans Power Syst* 12(3):907–912
20. Xiaojun C, Xianpeng L, Peng X (2015) IOT-based air pollution monitoring and forecasting system. In: 2015 international conference on computer and computational sciences (ICCCS). IEEE, pp 257–260

EverTrack: A System to Track History of Used Cars Using Blockchain



Rex Rony Jacob, Rohit Mahesh, Shannon Gomez, Syed Sahabuddin, and Vijitha Robinson

Abstract Because there are so many different parties involved in the automobile trade, it is one among the major sectors that suffers from fraudulent activities and lack of integrity. So, a proper outcome to track and log data of these vehicles are necessary to satisfy a high level of trust, integrity, and traceability. A systematic approach to change the way specific industries operate by eliminating middle layers between the seller and buyer. In this paper, the system that we introduce which we have named it as EverTrack will make it easier for auto owners, repair shops, and insurance-based companies for registering and adding new information regarding their vehicles, as well as include an anti-theft system. EverTrack will address the difficulties of centralised used automobile history reporting systems, including authority and repair agencies, as well as central storage databases.

Keywords Blockchain · Centralised · Vehicle · History

1 Introduction

Buying and selling used cars is a lucrative business all around the world. Customers in the second-hand vehicle market, on the other hand, frequently find themselves in the situation of not knowing whether the details given out by the car dealer is matching with the original condition, resulting in the purchase of vehicles with issues (e.g. odometer fraud and accident vehicle) [1]. Blockchain technology is a new approach to data management. Because of its qualities, a set of users are given the privilege to work together to store, administer, and defend the blockchain without anybody having to trust anyone else, avoiding the risks of centralised data control from an outside source. Furthermore, every node copies the entire data using P2P networks,

R. R. Jacob · R. Mahesh · S. Gomez (✉) · S. Sahabuddin · V. Robinson
Department of Computer Science and Engineering, Mar Baselios College of Engineering and Technology, Thiruvananthapuram, Kerala 695015, India
e-mail: shannongomez60@yahoo.in

V. Robinson
e-mail: vijitha.robinson@mbcet.ac.in

making the data nearly impossible to remove. As a result, this technology can be used to track the situation of second-hand cars. In this paper, a motor vehicle history tracking interface, named EverTrack, is proposed. This system creates a networking medium that regular customers, car vendors, mechanics, insurance firms, and governments can access. Distinct user categories belong to different transactions. EverTrack allows automobile owners to show their vehicle's information to various consumers and dealers without the need for any central authorities. When an automobile is transferred from one owner to the other, the information is updated to the new owner.

2 Overview of Blockchain Technology

Blockchain is a digital ledger that is immutable, decentralised, and distributed. It keeps track of the increasing set of blocks connected to one another and encrypted in an immutable chain [1]. This chain is spread, with numerous copies physically dispersed among network nodes. Because there is not any single point of authority and control is shared among numerous independent entities like as individuals, computers, or businesses demonstrates blockchain decentralisation. Integrity and sense of trust between the participants that cannot be trusted is an issue in distributed systems. The immutability of blockchain assures that blocks in the append-only ledger cannot be altered with. The act of adding new blocks, on the other hand, may result in several copies of the ledger being disseminated throughout the system's nodes [1]. To establish a majority agreement among network participants, a powerful consensus method is required (51% of the nodes or more). The source of truth is represented to the network as the accepted and chosen ledger after the next updated state of the ledger is agreed upon. As a result, every member of the network will possess similar copy of the particular ledger [2–5].

3 Related Works

Since the inception of blockchain, other similar systems have been built. Some are concerned with financial transactions, like decentralised exchange. On the cloud storage platform, some address the problem of personal data ownership. Others include application of supply chain, insurance, IoT integrated with blockchain, application based on gaming, and trade involving a radio spectrum. EverTrack focuses on data collection, and this type of application is related to supply chain management [6]. In [7], a supply chain blockchain framework management is proposed. This framework consists of various IoT sensors and functionalities of blockchain. It helps in creating a secure storage with a variety of high quality data, assets, logistical data, and transaction data. In this system, the smart contract functionality provides privacy and intelligence, while sensors collect information from the actual world. Every information in the product's production process and other information in the

transit phase will be recorded on blockchain framework. Because of the blockchain's tamperproof feature, recording all the data in the product's production can enable users rapidly comprehend the genuine information of the goods [8–12]. This method of establishing a biography can be used to record condition of a vehicle. Similar study [13] is a blockchain-based huge information exchange solution for China's automotive industry. They have formed an automobile industry partnership in China and released a paper that focuses on developing a blockchain platform for collecting vehicle data. The unveiling was made official towards the end of 2018, as platform dubbed engine chain (EGCC) [6, 14–16], with the goal of using it to store information from a car's entire life cycle. The research mentioned above is close to what the system implemented here wishes to accomplish, yet it differs slightly. Data querying would be free in our system. The group with record permission is the second distinction. All users that are registered can upload data at EGCC, however, our system only enables trustworthy third parties, preventing malevolent users from tampering with data. As a result, blockchain technology is employed where car condition news is immutable, hence assisting buyers in having legitimate vehicle information for reference when purchasing used vehicles.

4 System Architecture

As previously said, our aim is to create and build a system that not only provides consumers with car history information, but also analyses the data and provides a logical framework that is immutable. For crypto-currencies, blockchain has been proposed as a dynamic and decentralised ledger. This framework, on the other hand, has shown versatility in the implementation of any decentralised computing resource applications. This property of blockchain is extended by the EverTrack structure. Different types of users or subscribers can join and register in EverTrack, including regular users, car owners, car repair shops, and insurance companies. EverTrack leverages smart contract for block updates, registration of new owner, and new organisations and agencies. Figure 1 shows the interaction process using a use-case diagram. The user has permission to access the information in the system provided by the various service providers such as manufacturers, DMVs, insurance companies, mechanics, etc. These service providers enter the data into system via the blockchain platform which makes the data immutable. The information entered will be validated using with the help of an admin.

4.1 High Level Architectural Diagram

The architecture of EverTrack consists of (1) infrastructure layer, (2) service layer, and (3) application layer as in Fig. 2. The infrastructure layer will upload and store a large amount of vehicle data in auxiliary storage, freeing up more storage capacity.

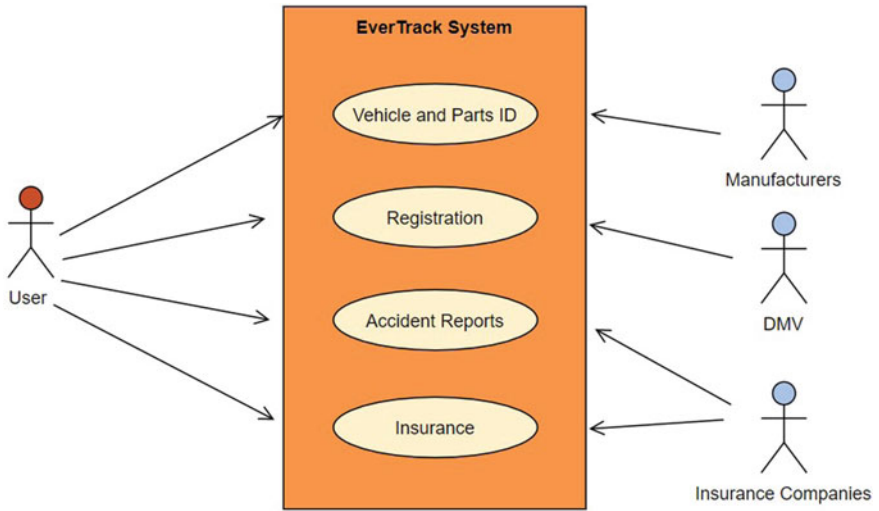


Fig. 1 Interaction diagram

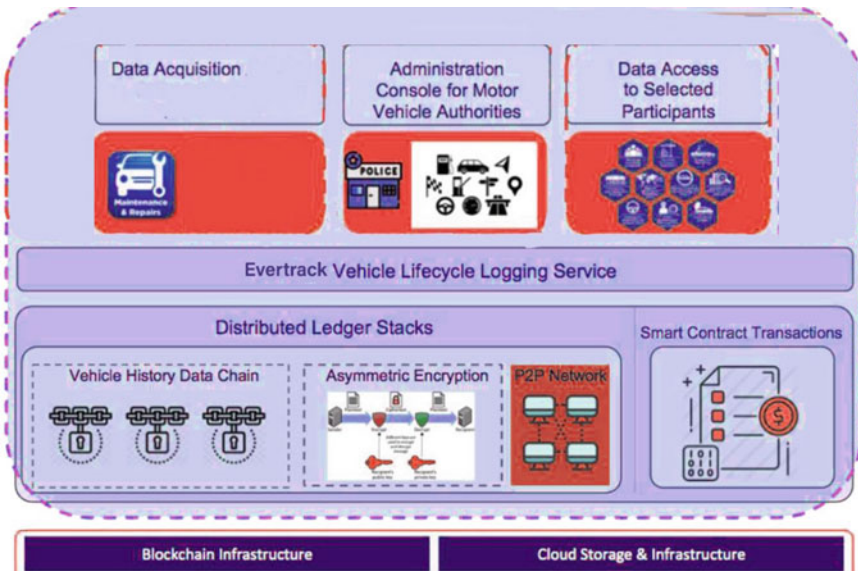


Fig. 2 Architectural diagram

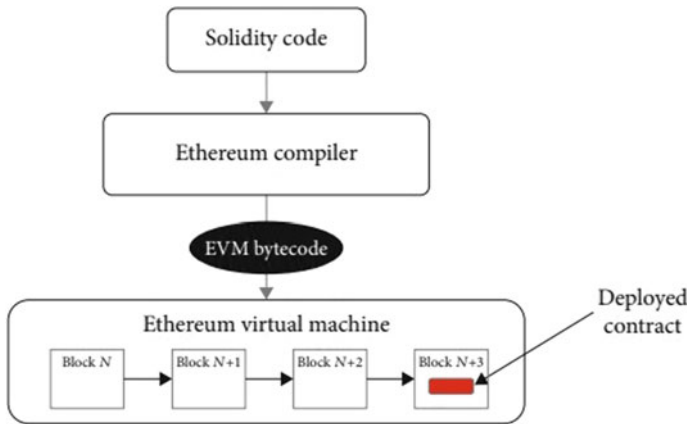


Fig. 3 Ethereum smart contract execution flow

Instead of the blockchain, these less frequently utilised data are stored in auxiliary storage. Users can interact with the service layer through three different types of interfaces (apps, web pages, and application programming interfaces). The application layer is where external partners can integrate EverTrack service to construct a variety of services such as auto loan services, car insurance, certified second-hand cars, and so on.

4.2 Smart Contract

Solidity is the programming language used to create the smart contract. Solidity programmes must first be compiled into Ethereum-based code before adding it to a transaction and later being sent to a network. After that, verification will be done by Ethereum miners and a particular block is used to record the same, thus completing the smart contract deploying function. By sending a transaction, users using Ethereum can invoke a smart contract operation. Before running the smart contract’s bytecode on Ethereum virtual machine to fulfil the call of the contract, verification will be done by Ethereum miners and a particular block is used to record the same. Ethereum smart contract execution flow in Fig. 3.

4.3 Smart Contract Functions

Several smart contract functions have been called in order to perform the particular tasks (as shown in Fig. 4). They are mentioned below:

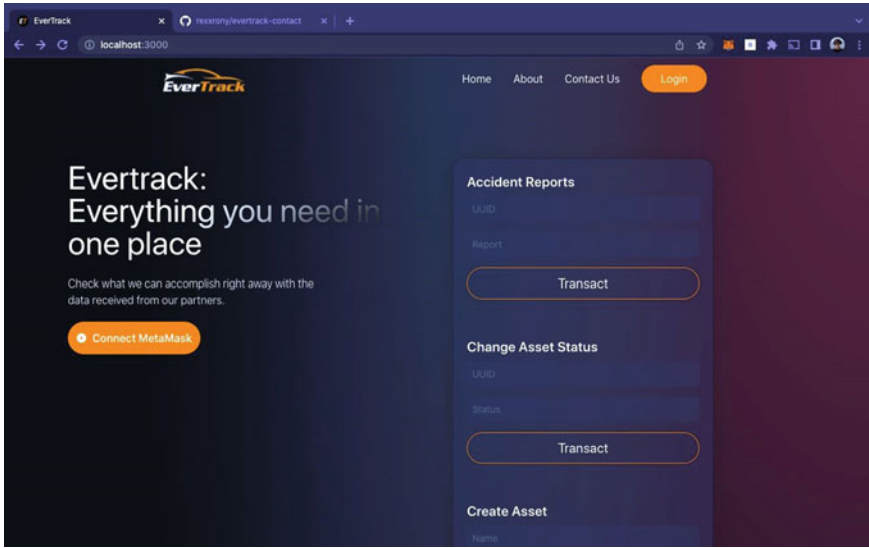


Fig. 4 User interface of EverTrack

- CreateCar: This function provides an interface to create or add a new vehicle along with its details into the blockchain.
- GrantRole: This function is used by the admin on deciding what kind of role is to be given out to each individual or organisation.
- TransferCar: It allows us to change the ownership of the car from one user to another along with proper verification.
- AccidentReports: This function gives us a platform to provide details regarding any accidents that the vehicle has gone through.
- GetCarbyUUID: It provides an overall information regarding the vehicle using a unique key. This smart contract function uses a 'struct' keyword which is basically a key-value binding mechanism.
- IsOwnerOf: This checks the ownership of the vehicle.
- PartDetails: It shows all the details of various parts of the vehicle along with its verification.
- ChangeCarStatus: This function shows the current status of the car and provides an option to change it as well. The status of the vehicle is 'NIL' by default. The main aim of this function is to implement the anti-theft mechanism where the various parts of vehicle would be having a unique number which is numerically linked with the UUID of the vehicle. The status of the car can be changed if the ID does not match.

4.4 Proof of Work

Bitcoin and various other well-known cryptocurrency networks employ the PoW consensus algorithm. To have the ability to add new transactions to the blockchain, participating nodes must successfully solve a computationally challenging challenge. Since they may earn Bitcoins by validating the blocks, the people that attempt to solve the difficulty are referred to as ‘miners’. A similar approach has been established where the verification of vehicle information is verified by the admins.

5 Results and Discussions

With a broad consumer base and a high demand market, used vehicles are one of the most essential second-hand commodities. The industry has huge economic repercussions, with a business volume of 180.4 billion euros in Europe alone [1]. According to some studies, the used vehicle market is two to three times greater than the new vehicle industry. Because it involves multiple parties from the original distribution company, vehicle data gathering and tracking is difficult. This happens to people who work with a variety of repair businesses and insurance agencies from around the world. It is critical to obtain real car history and erase any deceptive information in the used vehicle market to minimise fraud. As a result, we were able to make an accurate vehicle valuation evaluation. Information regarding the vehicle must be transferred and retrieved in a secure storage format where data authenticity and integrity may be easily verified. Figures 4 and 5 show the system functioning result and the user interface.

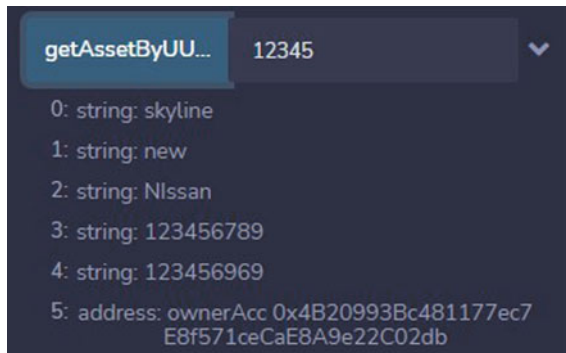


Fig. 5 Information accessed by the user

6 EverTrack Challenges and Opportunities

EverTrack can be described as a distributed ledger technology application. It is also considered to be a peer-to-peer (P2P) application. This diversity presents issues as well as numerous technical details in networking, storage, and security. Few of the technical issues will be discussed in this section. For starters, EverTrack has a problem with computational power. In these platforms, the application should work without any computational burden. Because organisation nodes are likely to be strong servers, the mining feature has been disabled in normal nodes to reduce computational concerns. A computing power assessment function with a threshold algorithm may be added in future to lessen the mining burden. The execution of this article is based on an existing platform (Ethereum). As a result, their system's capability limits their functionality. Furthermore, given the method in this study is a PoC, it may possess challenges in online utilisation, like how a good trusted third party is chosen for chain recording. Also, because PoS cannot be considered as a consensus method like PoW, it is not decentralised as Bitcoin. It could prevent second-hand traders from keeping car information on their own. As a result, it is still possible to remedy this issue.

7 Conclusion

This paper uses blockchain technology to create a trusted car data source system. Thanks to the help of reliable third parties who swiftly register information regarding the vehicle into our network, customers need not worry about getting untrustworthy information from used car sellers. Furthermore, because to the nature of this technology, we will be able to assure that information about vehicle condition is transparent across every user and is immutable. The system is used by automakers for smart contract deployment that shows a vehicle. The suggested system allows the maintenance unit and government-oriented branches to capture and update information of the vehicle condition. Through the interface of the system, users can simply search car information. All of the aforementioned operations are free of charge. The transaction throughput of the suggested system was found to be sufficient for daily transactions.

References

1. El-Switi S, Qataweh M (2021) Application of blockchain technology in used vehicle market: a review
2. Beck R, Avital M, Rossi M (2017) Blockchain technology in business and information systems research. Published online: 15 Nov 2017
3. Vishnu Gopal PG, Mathew G (2021) Blockchain based verification of vehicle history for pre-owned vehicle industry. In: 2021 international conference on communication, control and information sciences (ICCISc), pp 1–6. <https://doi.org/10.1109/ICCISc52257.2021.9484896>

4. Meyliana et al (2021) Blockchain technology for vehicle maintenance registration. In: 2021 international conference on information management and technology (ICIMTech), pp 608–613. <https://doi.org/10.1109/ICIMTech53080.2021.9534974>
5. Hossain MP, Khaled M, Saju SA, Roy S, Biswas M, Rahaman MA (2020) Vehicle registration and information management using blockchain based distributed ledger from Bangladesh perspective. In: 2020 IEEE region 10 symposium (TENSYPMP), pp 900–903. <https://doi.org/10.1109/TENSYPMP50017.2020.9230781>
6. Jiang Y-T, Sun H-M (2021) A blockchain-based vehicle condition recording system for second-hand vehicle market
7. Chen S, Shi R, Ren Z, Yan J, Shi Y, Zhang J (2017) A blockchain based supply chain quality management framework. In: 2017 IEEE 14th international conference on eBusiness engineering (ICEBE), Shanghai, China, pp 172–176
8. Notheisen B, Cholewa JB, Shanmugam AP (2017) Trading real world assets on blockchain. *Bus Inf Syst Eng* 59:425–440. <https://doi.org/10.1007/s12599-017-0499-8>
9. Dorri A, Steger M, Kanhere SS, Jurdak R (2017) BlockChain: a distributed solution to automotive security and privacy. *IEEE Commun Mag* 55(12):119–125. <https://doi.org/10.1109/MCOM.2017.1700879>
10. Khanna H, Mehta H (2020) Decentralized vehicle registry on blockchain
11. Yoo SG, Ahn B (2021) A study for efficiency improvement of used car trading based on a public blockchain. *J Supercomput* 77:10621–10635. <https://doi.org/10.1007/s11227-021-03681-z>
12. Joby PP (2020) Expedient information retrieval system for web pages using the natural language modeling. *J Artif Intell* 2(02):100110
13. Chen J, Ruan Y, Guo L, Lu H (2020) BCVehis: a blockchain-based service prototype of vehicle history tracking for used-car trades in China. *IEEE Access* 8:214842–214851. <https://doi.org/10.1109/ACCESS.2020.3040229>
14. Masoud MZ, Jaradat Y, Jannoud I (2019) CarChain: a novel public blockchain-based used motor vehicle history reporting system
15. Farr Z, Azab M, Samir E (2020) Blockchain-based cooperative autonomous detection of suspicious vehicles
16. Subramanian G, Thampy AS (2021) Implementation of hybrid blockchain in a pre-owned electric vehicle supply chain. *IEEE Access* 9:82435–82454. <https://doi.org/10.1109/ACCESS.2021.3084942>

Identification of Respiratory Diseases and Diabetes by Non-invasive Method Using IoT



S. Suthagar , G. Mageshkumar , K. Hemalatha ,
Saggurthi Prabhakara Rao, R. Mahesh, and S. M. Kural Eniyavan

Abstract The objective of this paper is to identify respiratory diseases such as Asthma, Covid-19, pulmonary disease, and diabetes from the human breath odor using a non-invasive method. For detecting diseases using a non-invasive method, temperature sensor (to identify body temperature), pulse oximeter sensor (to identify blood oxygen level and heartbeat rate), and acetone sensor (to identify respiratory diseases from human breath odor) with Arduino ATmega328 microcontroller unit (MCU) were used. If the temperature is greater than 37.2°C, the heartbeat rate is greater than 100 bpm, and the oxygen level is less than 92% SpO₂, Covid-19 will be detected. If the oxygen level is less than 95% SpO₂, the heartbeat rate is at (100–125) bpm, and the temperature is at 36.1–37°C, asthma will be detected. If the heart rate is greater than 86 bpm, the temperature at 36.1–37°C, the oxygen level at 92–97% SpO₂, and the acetone level at (354–496) ppm, diabetes will be detected. If the oxygen level is less than 92% SpO₂, the temperature at 36.1–37°C, and the heartbeat rate is greater than 110 bpm, the pulmonary disease will be identified. After disease detection, suggestions will be provided to the patients based on their health reports. Finally, suggested medicines will be sent to the patient's registered mobile phones by connecting node MCU with blynk using IoT technology. The results will be stored and the patients can compare their health conditions for future analysis. The traditional method of laboratory tests is considered to consume more time. In our method, the duration of the detection process is less and the results help to identify health problems at early stages and predict diseases quickly compared to the traditional method.

Keywords Electronic nose · Diabetes · Respiratory diseases · Non-invasive method · Internet of things

S. Suthagar (✉) · G. Mageshkumar · K. Hemalatha · R. Mahesh · S. M. Kural Eniyavan
Kongu Engineering College, Perundurai, Erode 638060, India
e-mail: susjou@gmail.com

S. Prabhakara Rao
Joginpally B R Engineering College, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes
in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_30

1 Introduction

There are more than 422 million people who have diabetes according to the World Health Organization. Diabetes is the seventh main cause of death and it was high risk due to complications such as cardiovascular disease, blindness, the risk of amputation, and renal failure [11]. Blood sugar is an important parameter for diabetics. It needs to keep it in good condition. Levels can delay these complications. Diagnosis prioritizes non-invasive procedures that exclude not only vascular damage but also skin surface damage. Such procedures have no pain and avoid the possibility of infection or trauma to the patient [17]. However, despite frequent efforts, the problem due to blood glucose levels by non-invasive measurement remains unsolved. Visual methods for determining blood sugar levels has limitations, especially near-infrared spectroscopy, which has failed to provide the needed sensitivity [13]. Therefore, analysis of excreted fluid should be considered for non-invasive diagnosis. Several attempts have been made to determine blood glucose levels by applying negative pressure to the skin and measuring the fluid generated on the opposite surface of the skin, including glucose iontophoresis [22]. The latter was spotted on sale machines but disappeared after its release.

Traditional methods of laboratory testing are very slow. The non-invasive method does not hurt any blood tissue and does not hurt the skin of the human body. This is the safest method for analyzing human diseases.

Asthma: Asthma is a disease that causes lung inflammation. It is identified by frequent and recurrent symptoms, obstruction of the airway, and bronchospasms that can be easily caused. Symptoms include GASP, cough, breast, and respiratory episodes [16]. This can happen several times a day or several times a week. Depending on the individual, symptoms of asthma may subside at night with exercise.

Covid-19: Coronavirus 2019 (COVID-19) is a highly infectious disease that is caused by a virus, SARSCoV2. COVID-19 is transmitted when people breathe the infected virus by air droplets and tiny particles of air that contain the virus. People stay infected for up to 14d and can spread the virus even without symptoms. The risk of breathing problems in this situation is very high when people are near but it can be sniffed for long distances, especially indoors [19].

COPD: Chronic obstructive pulmonary disease (COPD) is defined as obstruction of the respiratory tract which does not change much over several months. It is a disease that includes chronic obstructive pulmonary disease, bronchiolitis, and emphysema, the effects of which vary from person to person. COPD is a major cause of illness and death worldwide [2]. Although this is a serious health problem with limited treatment options, most Drugs have been developed to treat asthma. The successful development of new COPD drugs requires a better understanding of natural history, epidemiology, genetic and environmental risk factors, and the pathophysiology of the disease.

2 Literature Review

Yatabe et al., have developed an odor sensor system that uses a chemically sensitive resistor that outputs multi-channel data. A mixture of fixing material (GC material) for gas chromatography and carbon black was used as a chemical-sensitive resistor. The interaction between the chemically sensitive resistor and the gas molecules has shifted the electric resistance of the resistor. In addition, a small measuring instrument was created. The instrument's gas type readings were used to obtain 16 channel data. Machine learning techniques in the Weka software were used to examine the collected data. As a consequence, the sensor system was able to detect alcoholism successfully [25].

Karpova et al., presented on Diabetes Monitoring by frequent Analysis of Diabetic Winding Immediately after creating Flood Ingle Course Biosensor Used Biosensor is based on Perfluoro sulfonated ionomer and glucose oxidase immobilized in alkoxysilane. There is the sensitivity and calibration range obtained in batch mode. Based on glucose tolerance testing a clinically relevant positive link between rising glucose concentrations and a strategy for imitating hyperglycemia and non-invasive collecting sweat speed was observed. The observed interaction between blood and sweat assumed by low-grade metabolites was better than those previously observed between capillary blood and venous blood. The kinetics of weld glucose concentrations collected by the proposed biosensor is well matched to fluctuations in blood sugar content without delay, which gives the impression of non-diabetic precautions [10].

João et al., presented that exercise and fitness tend to increase health benefits. The human body loses a lot of electrolytes during sweating and needs to be replaced with special fluids, thus introducing a smart wearing sensor that allows you to monitor sweat using biosensors that provide real-time feedback with a smartphone app that notifies the user when necessary they are needed to change electrolytes. The glucose solution changed into organized with numerous concentrations tested using chronoamperometry. The primary effects have been received through the usage of commercial glucose sensors and compared with a glucose meter. The system appears to be effective in measuring high and low glucose levels in sweat [14].

Roosbeh et al., have presented a Wearable skin-to-skin device with integrated colorimetry, microfluidic channels, and electrochemical sensors for non-irritating, real-time sweating. The hypothesis explains current developments in the creation and application of novel wearable sensors for evaluating sweat variations and biomarkers utilizing precise samples and real-time analysis. Sensitivity, system stability, and high usage in remote places are significant areas of opportunity that enable wider dissemination in field investigations, clinical trials, and current sales. increase. Body measurements show good consistency in these conditions when compared with conventional laboratory-based sweat analysis methods [7].

Rydosz et al., have presented the blood volume of volatile organic compounds (VOCs) that occurs in a person's respiratory tract due to metabolic changes as a promising tool for non-invasive medical diagnosis. The amount of acetone released

is associated with diabetes monitoring. Traditional methods of analyzing exhaled air are based on spectroscopic techniques, but the gas sensors development makes them more attractive in the medical aspect [18].

Asrosa et al., presented MQ138 semiconductor sensor was incorporated into a three-block system. It has advantages such as portability, speed, and ease of use. Experiments with breath acetone samples from diabetic and non-diabetic patients confirmed the system's accuracy. According to the results of the measurements, the greater the sensor output voltage, the higher the blood glucose values obtained, implying that sensor readings of the MQ138 sensor have a linear relationship with breath acetone. As a result, the MQ138 semiconductor sensor has the potential to be employed for breath acetone detection in diabetic patients [15]. Other works related to Internet of things to connect the wearable device to collect medical data and store the information of patients in the cloud database as described by Suma in their article [20] and more advanced applications can be developed using blockchain technology to monitor the clinical sensor information through internet of things as proposed by Wang et al. [23].

3 Proposed Work

3.1 Hardware Design

The non-invasive method [25] does not damage the blood tissue or the skin of the human body. This is the safest way to analyze human illness. Sensors such as pulse oximeters, temperature sensors, heart rate sensors, and acetone sensors MQ138 are used to detect illness in a non-invasive manner [10]. Drugs are proposed based on the identified illness and sent to the patient's registered mobile phone. The block diagram of components. Connect the sensor to the MCU node and connect to the Blynk app via the Wi-Fi protocol. To analyze the degree of diabetes in a patient, a person's acetone concentration is needed. Disease analysis is performed using the Arduino ATMEGA 328 microcontroller as shown in Fig. 1. Pulse oximetry sensors are used to determine a patient's blood oxygen level and heart rate. Body temperature sensors are used to measure a patient's temperature level. Acetone sensors are used to measure human acetone levels. Heart rate sensors are used to determine a patient's heart rate. After disease detection, suggestions will be provided to the patients based on their health reports. Finally, suggested medicines will be sent to the patient's registered mobile phones based on Table 1. The results will be stored and the patients can compare their health conditions for future analysis.

Arduino is a single-board microcontroller that makes it easy to use electronic devices for interdisciplinary projects [21]. The hardware consists of an open hardware board built around an 8-bit or 32-bit Atmel ARM microcontroller as shown in Fig. 2. This software contains a combination of standard programming language and a launcher that works on a small controller. Arduino boards can be purchased

Fig. 1 Functional diagram of the proposed system

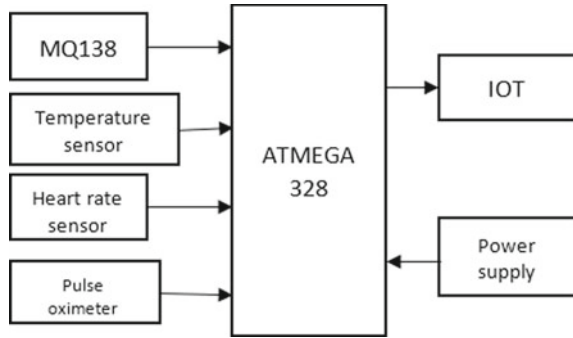


Table 1 General medicine prescribed for various diseases at initial stage

Diseases	General medicines prescribed
Diabetes	Metformin, glibenclamide
Asthma	Salbutamol, terbutaline
COPD	Roflumilast
Covid-19	Remdesivir, paracetamol

in combination or as a custom kit. For those who want to integrate Arduino themselves, hardware design information is available. Arduino ATmega328 microcontroller chip (black, bottom right); 14 digital I/O pins are present at the top and 6 analog input panels at the bottom right. The Arduino board includes a small Atmel 8-bit AVR controller with additional configuration and integration components in some regions. A main feature of Arduino is the procedural way of connecting the connectors, which allows the CPU board to be connected to various flexible modules including so-called shields. In comparison to other devices that require an external programmer, Arduino’s microcontrollers come pre-programmed with a bootloader that makes downloading programs into on-chip flash memory much easier. This microcontroller has been chosen as it has a maximum clock frequency of 20Mbpsm, that runs at 20 Million Instructions per Minute. It has a good memory resource (i.e., an EEPROM off1KBN) suitable for real-time implementation. MIPS When utilizing the Arduino software stack, all boards are programmed over an RS232 serial connection on a basic level, although how this is done varies by hardware version. The Arduino serial board contains a level switch to convert between RS232 and TTL level signals. The system operates with a supply voltage of 5 V.

ESP8266 Node MCU is the type of small controller which is designed by Espressif Systems. Node MCU ESP8266 is an independent Wi-Fi network solution that provides a bridge between the existing microcontroller and Wi-Fi and is capable of using standalone applications [3]. Node MCU module includes a built-in USB connector and various pins shown in Fig. 3 with a small USB cable, it can connect the Node MCU upgrade kit to the laptop and turn it on without any problems, like Arduino. It also quickly integrates with breadboards. This device can be used to connect to

Fig. 2 Arduino Uno with ATmega328 MCU

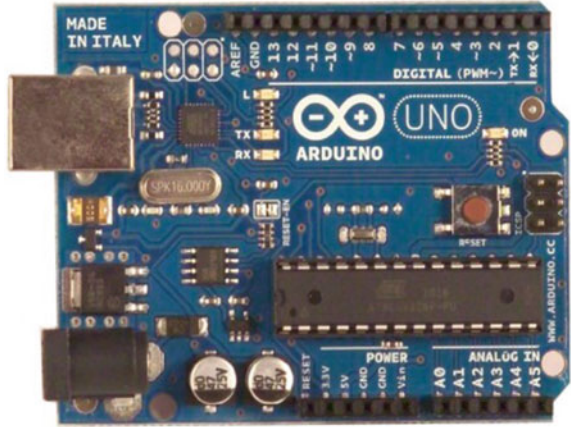
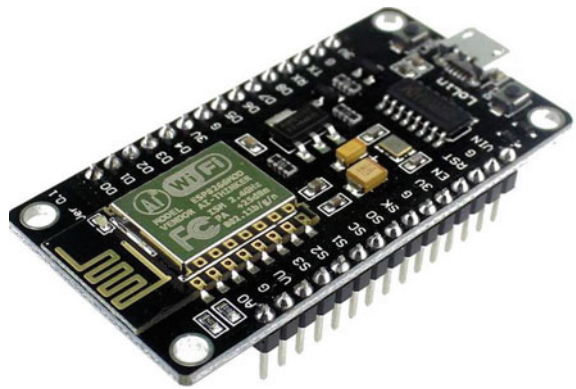


Fig. 3 NodeMCU-ESP8266 Wi-Fi module



any Wi-Fi network and transfer the sensor information to the cloud from a remote place to the server for collection of data and to make decisions over the actions to perform [12]. ESP8266 NODE MCU is an independent Wi-Fi network solution that provides a bridge between the existing microcontroller and is mainly used to display the output in the Blynk app by a data transfer using the Wi-Fi protocol.

The thermistor for measuring the human body temperature is shown in Fig. 4. The thermistor is based on the simple concept of resistance changing with temperature [14]. When the ambient temperature changes, the thermistor's elements begin to self-heat. As the temperature varies, the device's resistance value changes. This variation is determined by the type of thermistor used. Varied varieties of thermistors have different resistance temperature properties. Thermistors, also known as Thermally Sensitive Resistors, are temperature sensors that work by changing resistance as temperature changes. They are made from semiconducting materials. Thermistors are available in a variety of shapes and sizes on the market. Smaller thermistors are made up of beads with diameters ranging from 0.15 to 1.5 mm [7]. The time

Fig. 4 Thermistor

constant of thermistors operating in sheaths is minimal due to this size advantage, albeit the size reduction also reduces their heat dissipation capability, increasing the self-heating impact. This effect has the potential to permanently harm the thermistor. In comparison to resistance thermometers, thermistors must be operated at low levels of electric current to avoid this, resulting in decreased measurement sensitivity.

The sensitive material in the MQ138 gas sensor is SnO_2 [18], which has lower conductivity in clean air is shown in Fig. 5. The conductivity of the sensor grows in lockstep with the concentration of VOC gas. A simple circuit can be used to translate the change in conductivity to the matching output signal of gas concentration. Toluene, acetone, alcohol, and methanol, as well as hydrogen and other organic vapors particularly sensitive to the MQ138 gas sensor. It is a low-cost sensor that can detect a range of organic gases and can be utilized in a variety of applications.

- Creating a new account in the Blynk app
- Syncing email account with the Blynk app
- Creating a new project
- Creating a new virtual terminal
- Entering the authentication code received via email and Wi-Fi details in the programming code
- Output will be displayed in the Blynk app using Node MCU.

The oximeter operates on the principle of spectrophotometry: the associated red absorption (absorbed by oxygen-rich blood) and the infrared light (absorbed by oxygen-rich blood) partially. The systolic absorption waveform is associated with arterial blood oxygen saturation. Relative light absorption rates are taken a few times per second, and these rates are processed by a machine to provide a new reading every 0.51 s [4], which measures the reading 3 s ago. Two light-emitting diodes, red and

Fig. 5 Acetone sensor—MQ138



Fig. 6 MAX30100 pulse oximeter module



infrared, are placed across their receivers in a sequence of 510 mm [6] of tissue. The transducer is usually placed on the finger, although sometimes the ears and forehead are used as alternatives. A pulse oximeter is an important test used to measure oxygen levels, i.e., oxygen saturation in the blood. This is a simple, painless measure of the quality of oxygen delivered to parts of your body far away from your heart, such as your arms and legs. The MAX30100 is a solution for pulse oximetry [24] and heartbeat rate. It incorporates dual LEDs, a photodetector advanced optics and low analog signal processing for pulse oximeter and heartbeat rate signals.

The Pulse Oximeter measures oxygen saturation when Oxygen enters the lungs and is transmitted to the bloodstream that is shown in Fig. 6. Blood takes oxygen to the various parts of our bodies [9]. The main source of oxygen in the blood is hemoglobin. We will call oxygen-deprived hemoglobin or deoxidated hemoglobin (deoxy Hb). The SpO_2 output should always take as identification of oxygen saturation. For example, if an FDA-approved pulse oximeter gives a reading of 90%, the true blood oxygen saturation is usually in the range of 86–94% [1]. The accuracy of a pulse oximeter is very high at saturation of 90–100%, the average is 80–90% [8], and the

Table 2 Parameter threshold for diagnosis of respiratory diseases and diabetes

Diseases	Sensors			
	Acetone sensor (ζ)	Heart rate sensor (α) (bpm)	Temperature sensor (β) ($^{\circ}$ C)	Pulse oximeter sensor (γ)
Asthma	-	100–125 (α_1)	31.6–37 (β_1)	92–95% SpO ₂ (γ_1)
Diabetes	354–496 ppm (ζ_2)	80–86 (α_2)	36.1–37 (β_2)	92–97% SpO ₂ (γ_2)
COPD	-	55–67 (α_3)	36.1–37 (β_3)	<92% SpO ₂ (γ_3)
Covid-19	-	> 100 (α_4)	>37.2 (β_4)	<92% SpO ₂ (γ_4)

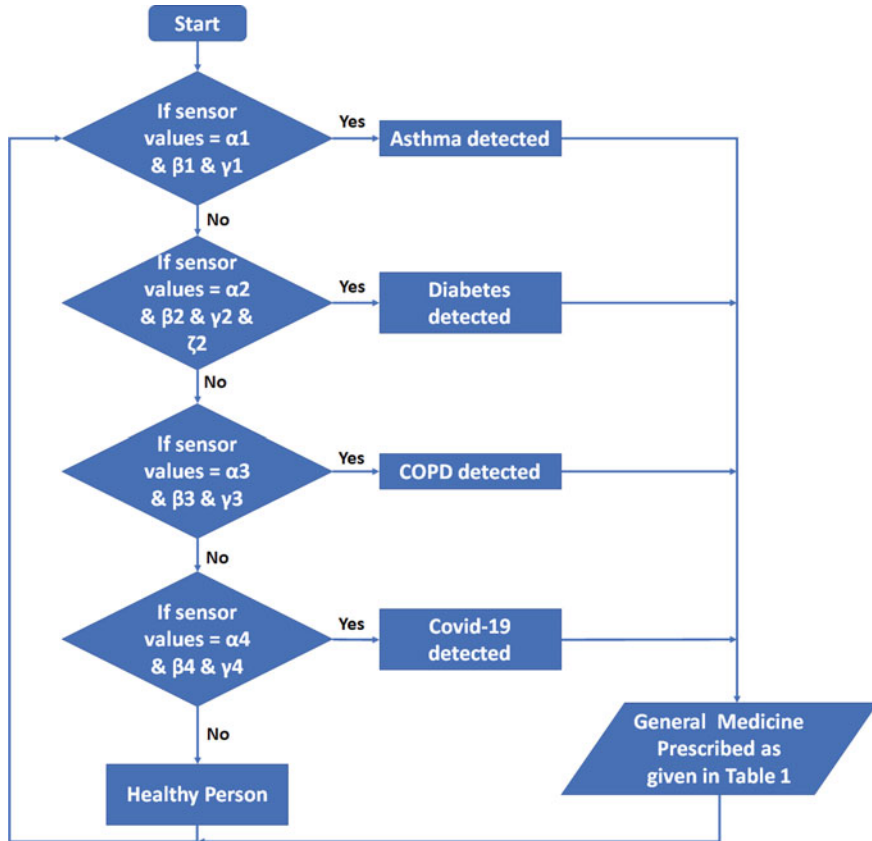


Fig. 7 Flowchart of the proposed work

lowest is less than 80%. Due to the limitations of individual-level accuracy, SpO₂ provides more accuracy for trends over time rather than absolute thresholds. Based upon various parameter thresholds shown in Table 2, the disease can be diagnosed by comparing it with the sensor values observed from the hardware implementation. The comparison of sensor values are performed as shown in Fig. 7

4 Result and Discussion

The input is given by patients to sensors namely, temperature sensor (thermistor), Pulse oximeter sensor (MAX30100), and Acetone sensor (MQ138). These sensors were connected with Arduino ATMEGA 328. The obtained result will be displayed using the node MCU module and it will be stored in a cloud database and used for future purposes by connecting an Arduino terminal with a tera term software and the data is saved in the excel sheet to compare the existing results.

There are no diseases identified as parameters does not match to given input by a person as shown in Fig. 8a. Diabetes is identified as parameters matched to given input by a person. So, the related medicine is suggested for the person as shown in Fig. 8b. Asthma is identified as a parameter match that has to be given input by a person. So, the related medicine is suggested for the person as shown in Fig. 9.

The hardware implementation of the proposed work to detect the respiratory disease and diabetes is shown in Fig. 10. It contains the sensors which include a temperature sensor (thermistor), Pulse oximeter sensor (MAX30100), and Acetone sensor (MQ138). These sensors have connected with Arduino ATMEGA 328 and node MCU. The hardware is non-invasive as the sensors used does not require, blood samples or body liquids.

The input taken from the pulse oximeter sensor by placing a finger is shown in Fig. 11a. The pulse oximeter sensor is used to identify the heartbeat rate and oxygen level of the patients. The input taken from the thermistor sensor by placing a finger



Fig. 8 The analysis report shows that the candidate is not having fever, but has diabetes



Fig. 9 The analysis report diagnoses that the candidate is having asthma

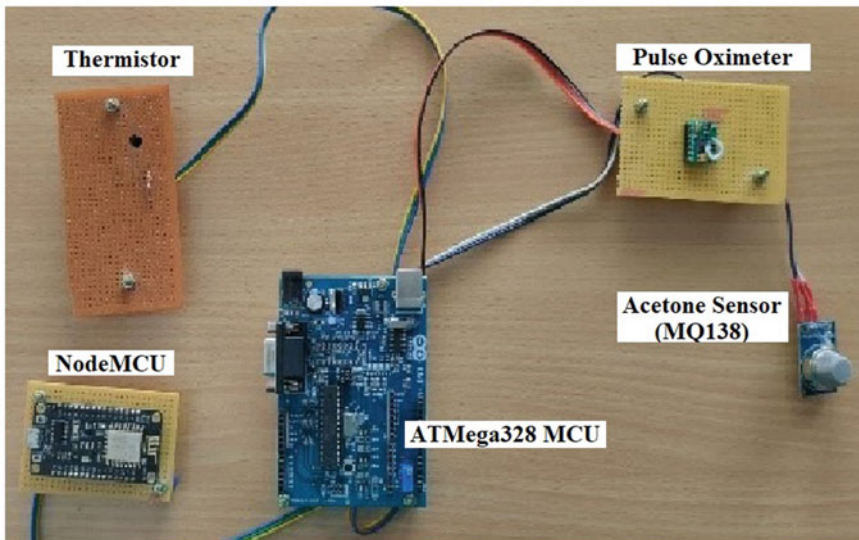


Fig. 10 Hardware implementation of the proposed work

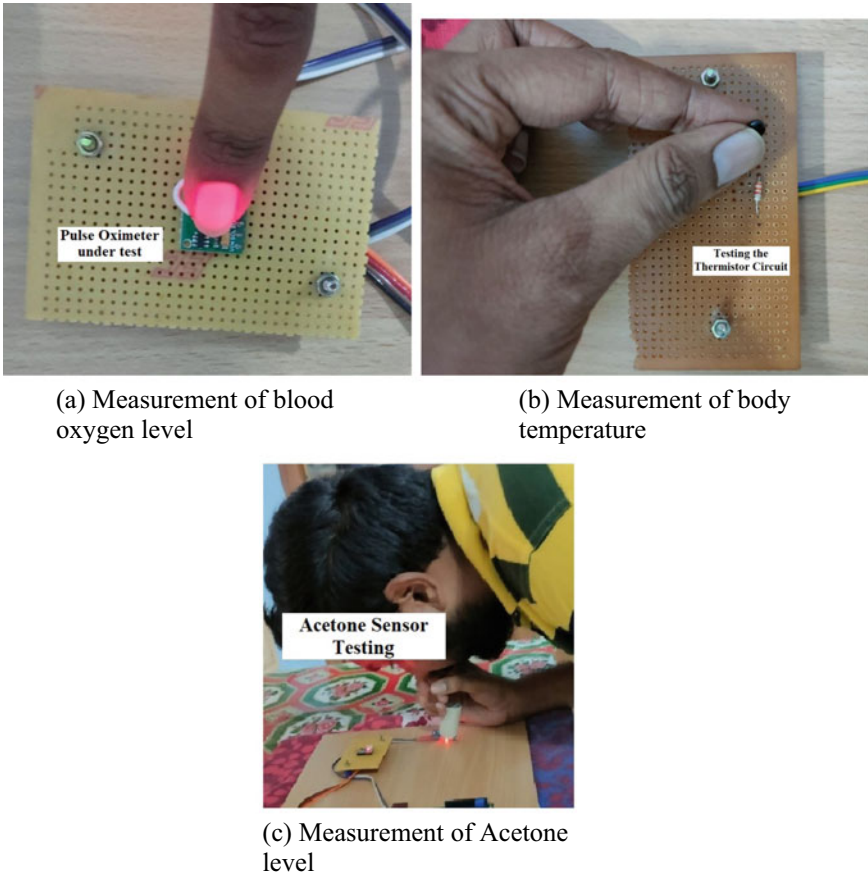


Fig. 11 Testing of sensor modules to measure the parameters for diagnosis of disease

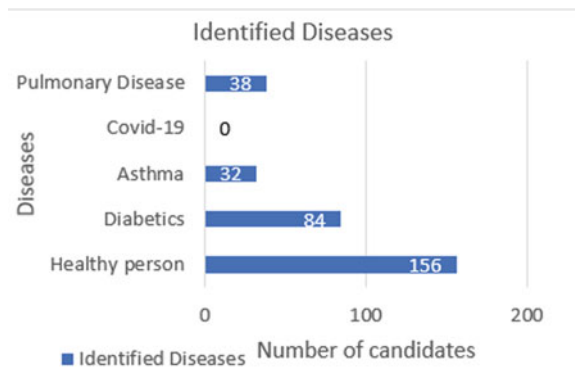
is shown in Fig. 11b. It is used to identify the body temperature of the patients. The input taken from the acetone sensor is shown in Fig. 11c. Acetone sensor detects the acetone level by using mass spectrometry which calculates the exact molecular weight of the sample volatile organic components and gas chromatography which is important in separating Volatile Organic components like acetone in a mixture from human mouth odor. Finally, the detected output is displayed in the Blynk app and it is stored in a database for future reference as shown in Fig. 12.

The collected data from the candidates are represented as shown in Fig. 13. The X-axis represents the diseases and Y-axis represents the number of candidates. Out of 310 candidates 38 candidates suffered from pulmonary disease, 32 candidates suffered from asthma, 84 candidates suffered from diabetics, and 156 candidates are healthy persons. Here, the collected data from the candidates were analyzed and compared as shown in Fig. 14. The X-axis represents the values obtained from

Fig. 12 Data stored in remote system

	A	B	C
1	Heart rate:54bpm / SpO2:94%		
2	Temperature:34		
3	Beat!		
4	maresh	02-04-2022	21
5	Heart rate:54bpm / SpO2:94%		
6	Temperature:34		
7	The person is healthy		
8	Beat!		
9	Heart rate:54bpm / SpO2:94%		

Fig. 13 Diagnosis report analysis of 310 candidates



the input and Y-axis represents the number of candidates. The acetone level of the candidates varies drastically with every candidate whereas the oxygen level and temperature of the candidates were almost similar.

Here, the data collected from the candidates (samples) are analyzed age-wise as shown in Fig. 15. The Y-axis represents the values obtained from the output of the proposed system and X-axis represents the ages of the candidates. The acetone level is lying above 350ppm for the candidates of 25–35 and greater than 60 years old when compared to other ages. The heartbeat rate is less than 100 bpm for candidates below 25 years old compared to other age candidates. People between the age of 25–35 and above the age of 60 having acetone levels above 350 ppm and a heartbeat

Fig. 14 Data visualization 310 candidate's body temperature, heart beat rate in bpm, oxygen level in percentage and acetone level in ppm

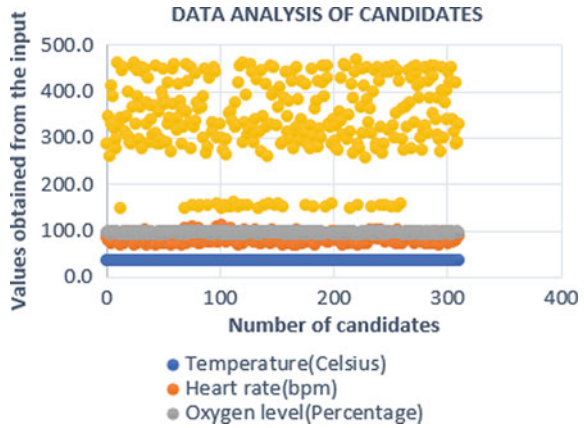
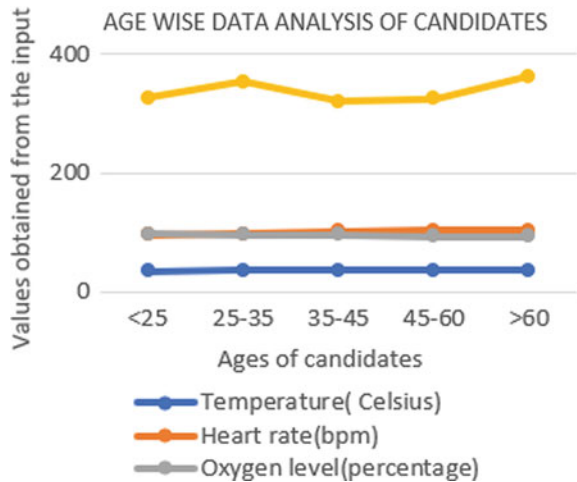


Fig. 15 Age-wise data visualization of 310 candidates



rate more than 100 bpm. So, they have symptoms of diabetes. Most of the people below the age of 25 are healthy and people above the age of 45 have SpO_2 levels of less than 95%. So, they have symptoms of asthma and COPD disease.

5 Conclusion

The acetone doses are used as a main diagnostic tool for diabetes in diagnostic tests or diabetes based on the relationship between acetone concentration in exposure and glucose levels in the blood. The interaction between medicine and technology in obtaining VOC has the potential to change the way of monitoring people's health [5]. Breath tests have several advantages over traditional laboratory testing, including

non-invasiveness, painlessness, ease of use, and real-time data. However, they have yet to be implemented in clinical practice, probably due to the lack of commercially available devices. The development of acetone respirator sensors over the next few years is expected to focus on the manufacturing process to bring the end-to-end device to market to accurately capture and reproduce inspired respiration. Long-term testing should be done so that all targeted acquisition strategies are 100 percent sure that the results obtained are the same.

References

1. Arai T, Kayashima S, Kikuchi M, Kaneyoshi A, Itoh N (1995) A non-invasive portable blood-glucose monitoring system: sampling of suction effusion fluid. *Nihon rinsho. Jpn J Clin Med* 53(4):1017–1022
2. Barnes PJ, Shapiro SD, Pauwels R (2003) Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J* 22(4):672–688
3. Behera B, Joshi R, Vishnu GA, Bhalerao S, Pandya HJ (2019) Electronic nose: a non-invasive technology for breath analysis of diabetes and lung cancer patients. *J Breath Res* 13(2):024001
4. Broza YY, Haick H (2013) Nanomaterial-based sensors for detection of disease by volatile organic compounds. *Nanomedicine* 8(5):785–806
5. Curran AM, Rabin SI, Prada PA, Furton KG (2005) Comparison of the volatile organic compounds present in human odor using spme-gc/ms. *J Chem Ecol* 31(7):1607–1619
6. Esfahani S, Wicaksono A, Mozdiak E, Arasaradnam RP, Covington JA (2018) Non-invasive diagnosis of diabetes by volatile organic compounds in urine using faims and fox4000 electronic nose. *Biosensors* 8(4):121
7. Ghaffari R, Yang DS, Kim J, Mansour A, Wright JA Jr, Model JB, Wright DE, Rogers JA, Ray TR (2021) State of sweat: emerging wearable systems for real-time, noninvasive sweat sensing and analytics. *ACS Sens* 6(8):2787–2801
8. Ito N, Saito A, Kuriyama T, Kimura J, Arai T, Kikuchi M, Kayashima S, Nagata N (1992) Development of a non-invasive transcutaneous blood glucose monitoring method using an isfet biosensor. *Front Med Biol Eng Int J Jpn Soc Med Electron Biol Eng* 4(1):35–45
9. Kao KW, Hsu MC, Chang YH, Gwo S, Yeh JA (2012) A sub-ppm acetone gas sensor for diabetes detection using 10 nm thick ultrathin inn fets. *Sensors* 12(6):7157–7168
10. Karpova EV, Shcherbacheva EV, Galushin AA, Vokhmyanina DV, Karyakina EE, Karyakin AA (2019) Noninvasive diabetes monitoring through continuous analysis of sweat using flow-through glucose biosensor. *Anal Chem* 91(6):3778–3783
11. Li S, Wang J, Zhang B, Li X, Liu Y (2019) Diabetes mellitus and cause-specific mortality: a population-based study. *Diab Metabol J* 43(3):319–341
12. Mageshkumar G, Kasthuri N, Tamilselvan K, Suthagar S, Sharmila A (2020) Design of industrial data monitoring device using iot through modbus protocol. *Int J Sci Technol Res* 9:1392–1396
13. Mhishi SB, Van Zyl DG (2020) Evaluation of the accuracy of visual glucose estimates by healthcare providers and patients at Kalafong Hospital, City of Tshwane, South Africa. *J Endocrinol Metabol Diabet South Afr* 25(1):18–23
14. Monge J, Postolache O, Plopa O, Trandabat A, Schreiner O, Schreiner T (2019) Glucose detection in sweat using biosensors. In: 2019 E-health and bioengineering conference (EHB), pp 1–5. IEEE
15. Nasution T, Asrosa R, Nainggolan I (2018) Application of mq-138 semiconductor sensor for breath acetone detection. *J Phys Conf Ser* 1116:032023
16. Pattemore PK, Asher MI, Harrison AC, Mitchell EA, Rea HH, Stewart AW (1990) The interrelationship among bronchial hyperresponsiveness, the diagnosis of asthma, and asthma symptoms. *Am Rev Respir Dis* 142(3):549–554

17. Rizzoni D, Mengozzi A, Masi S, Agabiti Rosei C, De Ciuceis C, Virdis A (2022) New noninvasive methods to evaluate microvascular structure and function. *Hypertension* 79(5):874–886
18. Rydosz A (2018) Sensors for enhanced detection of acetone as a potential tool for noninvasive diabetes monitoring. *Sensors* 18(7):2298
19. Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, Spijker R, Hooft L, Emperador D, Domen J et al (2021) Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has covid-19. *Cochrane Database Syst Rev* 2
20. Suma V (2021) Wearable iot based distributed framework for ubiquitous computing. *J Ubiquit Comput Commun Technol (UCCT)* 3(01):23–32
21. Suthagar S, Tamilselvan K, Priyadarshini M, Nihila B (2021) Determination of apple, lemon, and banana ripening stages using electronic nose and image processing. *Innovations in cyber physical systems*. Springer, Berlin, pp 755–769
22. Tang L, Chang SJ, Chen CJ, Liu JT (2020) Non-invasive blood glucose monitoring technology: a review. *Sensors* 20(23):6925
23. Wang H et al (2020) Iot based clinical sensor data management and transfer using blockchain technology. *J ISMAC* 2(03):154–159
24. Wyszynski B, Yatabe R, Nakao A, Nakatani M, Oki A, Oka H, Toko K (2017) Array of chemosensitive resistors with composites of gas chromatography (gc) materials and carbon black for detection and recognition of vocs: a basic study. *Sensors* 17(7):1606
25. Yatabe R, Shunori A, Wyszynski B, Hanai Y, Nakao A, Nakatani M, Oki A, Oka H, Washio T, Toko K (2020) Odor sensor system using chemosensitive resistor array and machine learning. *IEEE Sens J* 21(2):2077–2083

Use of E-Wallet as a Substitute for Physical Money in Transactions at Malls



Achmad Rizky Kurniawan, Rafif Imam, Rifqi Farhan Zhalifunnas, Armando Rama Putra, Ford Lumban Gaol, and Tokuro Matsuo

Abstract Several challenges have emerged as a result of the rapid spread of Covid-19 in Indonesia. To combat the pandemic, the government has also issued general guidelines for avoiding/limiting direct contact with common people who do not live with them. However, in some unavoidable cases, such as cash transactions, the social distancing has not yet been followed properly. These type of challenges can be avoided by the introduction of e-wallets and e-money. This research study examines the transactions made using e-wallets and e-money and how these methods are increasing sales in public places. The proposed research study has attempted to utilize both qualitative and a quantitative method to observe the customer behavior while using e-wallet as a substitute for physical money when making a transaction in a shopping mall. The result of the proposed research work will be compared with the existing literature to provide a conclusion based on existing research data and literature.

Keywords E-wallet · Behavior · E-money · Transactions · Pandemic

A. R. Kurniawan · R. Imam · R. F. Zhalifunnas · A. R. Putra
School of Information System, Bina Nusantara University, Alam Sutera, Jakarta, Indonesia
e-mail: achmad.kurniawan@binus.ac.id

R. Imam
e-mail: rafif.imam@binus.ac.id

R. F. Zhalifunnas
e-mail: rifqi.zhalifunnas@binus.ac.id

A. R. Putra
e-mail: handikalimanto@binus.ac.id

F. L. Gaol (✉)
Binus Graduate Program—Doctor of Computer Science, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aait.ac.jp

1 Introduction

People are now becoming increasingly familiar with the term “e-wallet” in this digital era. E-wallet is an electronic service that enables the storage of data and payment methods such as credit cards and electronic money, as well as the storage and payment of funds. Additionally, e-wallet can be considered as a piece of software, application, or a service that is designed to store digital currency and facilitate online transactions among users [1].

E-wallets are gaining popularity in Indonesia as the number of fintech companies increases. The government recently issued new rules for governing the legality of electronic money transactions by ensuring that technology can easily transform to the digital era [2]. Furthermore, in the current Covid-19 period, the government encourages people to limit physical interactions, which makes e-wallet purchases more accessible.

Due to the ease and speed with which e-wallet transactions can be completed, they are gaining popularity and are now starting to compete with transactions involving actual money [1].

Money usage is beginning to trend downward in 2021, as many people are now using e-wallets as a daily payment method [2]. This new payment mechanism simplifies, accelerates, and eliminates the need for physical money at a time when the majority of people are cautious of it. The goal of this research is to examine whether a country can eliminate physical money and replace it with e-money as the primary method of payment in order to reduce spending on money printing, increase transaction volume, and increase competition in the field of e-money and e-money wallets.

The significance of this research stems from the fact that the country is entering a new normal period, where people are restricted to live with minimal physical interaction. This research is especially significant since there is a possibility of innovation in which many nations are developing new methods for performing daily transactions [3].

Electronic money or e-money is a type of digital payment instrument that makes use of electronic media, specifically computer and internet networks [4]. E-money has been available in Indonesia since 2009, when it was ratified by Bank Indonesia Regulation 11/12/PBI/2009 about Electronic Money (Electronic Money), which was subsequently modified by PBI Regulation 18/17/PBI/2016. Electronic Cash, Digital Money, Digital Cash, Electronic Currency, or Digital Currency are all terms that are frequently used to refer to e-money. The Flazz BCA, e-money Mandiri, Brizzi BRI, Tap Cash BNI, Blink BTN, Mega Cash, and JakCard Bank DKI are all examples of chip-based electronic money (e-money) usage.

Meanwhile, an e-wallet or electronic wallet refers to a software program or feature that enables users to make payments more easily. GoPay, OVO, LinkAja, Kasku, and DANA are all examples of e-wallets. Apart from avoiding the present propagation

of the COVID-19 virus, there are other reasons why non-cash or cashless transactions have recently increased in Indonesia. Several other benefits of utilizing e-money and e-wallet include simplifying cash handling and reducing money management, obtaining discounts and rewards, tracking transactions more easily, being safer against theft, and being used to purchase a variety of essential things [5].

The primary significance of this research is to determine whether electronic money can truly enhance the function of physical money in our daily lives. The term “e-wallet” refers to an electronic money server that operates via a link on the publisher’s server [6]. E-Wallets are not interchangeable with credit or debit cards. Moreover, e-wallet transactions are not routed via third parties or intermediaries. According to halomoney.com, the following is the distinction between e-wallet and e-money:

- E-money in Indonesia is usually referred to a chip based on cards or other media, while e-wallet is an e-money-based server.
- E-money is usually used for everyday purchases such as toll roads, public transportation, and transactions at retail stores. While e-wallet is now commonly used for online shopping, shopping at offline retail stores, buying phone credit, etc.
- The maximum balance that e-money can accommodate is only Rp. 5.000.000, while e-wallet can accommodate Rp. 10.000.000.

According to Olsen et al. [7], the growth of e-wallets is not attributable to the features offered by e-wallets but to the growing number of mobile devices [8]. Nowadays, almost everyone owns a mobile device, and they are inextricably linked to daily life. Businesses are finally driving innovation by utilizing mobile devices as intermediaries for e-wallet as a result of the growing popularity of mobile devices. As a result, it is possible to conclude that e-wallets, in addition to existing currency, exist. Economic growth is a quantifiable measure of an economy’s progress in a given year when compared to the previous year [2]. GDP is a proxy used for economic growth as it represents the total monetary value of all goods and services produced in an economy over a given time period.

According to Untoro et al. [9], changes in the volume of transactions can serve as an early warning system for emerging economic trends. Through consumption and investment indices, more transactions can drive the economic growth [10].

- E-wallet: e-wallet is a form of prepaid card in which users deposit funds on their card in order to conduct future online purchases [2, p. 2, 11].
- E-money: e-money is typically used to refer to money that is stored in computers and banks and is used to enable electronic transactions. Its value is typically supported by fiat currency/government money.

This fundamental hypothesis is confirmed by various studies on the growing usage of electronic money and by government rules that urge to reduce contact whenever it is not necessary. Thus, by using e-money, it is possible to decrease the incidence of touches, as all purchasing and selling transactions are conducted online [12].

Since the transaction process with e-money is faster and more cost effective, it will be practical and popular with many people because they will no longer have to

worry about securing the physical money in their wallets or calculating change from transactions. With e-wallets, the transactions will become smoother and faster [13].

By taking a closer view on the e-wallet and e-money debate in terms of its benefits and drawbacks. Positively, by using e-wallets and e-money, criminal/theft activities can be prevented by tracking the path of money sent as well as the money transferred and received within a bank account. The time and space required to retain and store actual money will be reduced while maintaining data about the money in e-wallet [14]. An e-wallet and e-money will make it simple to transfer money and conduct fiat money transactions with few clicks.

On the other side, if someone hacks into a user's e-wallet or e-money account, they may access all of the user's information and steal the identity [15]. Individuals who are unfamiliar with e-money and e-wallets are prone to feel overwhelmed and make errors due to the absence of actual money. Certain individuals will struggle to maintain a control over their behavior when performing transactions since they do not perceive money being physically lost or given away.

Despite the fact that e-money and e-wallets are now widely used in Indonesia, they are currently limited to toll and parking payments [16]. Lack of understanding is also a result of Indonesia's eradication of physical money, as transactions involving e-wallets and e-money are typically limited to a few provinces or large cities, while some other provinces continue to use physical money.

Research Questions:

RQ.1: Is e-money safe enough to replace cash?

RQ.2: Is e-wallet advantageous for users and profitable for merchants?

RQ.3: Is the number of e-wallet transactions more than physical money?

2 Research Technique

A qualitative research has been carried out in this study through field observations, analysis of official sources, and quantitative research using questionnaires. The main objective of this study is to learn more about electronic money's effectiveness when compared to real money. Here, a scenario of shopping with e-money in a shopping mall is considered.

3 Research Method

A questionnaire has been used to perform a quantitative research. Since it is considered as exploratory due to the fact that it examines the preferences and factors that impact people's e-wallet and e-money decisions. Exploratory research identifies areas that have not been examined previously. The primary objective of this form of study

is to define new facts, insights, and understanding about a subject. Additionally, the researchers have employed a descriptive research. Descriptive research’s main objective is to characterize a phenomena and its properties [17, 18]. It is intended to elicit the respondents’ perceptions and opinions about the subject under consideration.

Q.1. Do you have an e-wallet account?

Q.2. How often do you make e-wallet transactions at the Mall?

Q.3. How easy do you think transactions are using e-wallet?

Q.4. Do you feel more consumptive when making transactions with e-wallets at malls?

Q.5. Do you prefer to use an e-wallet when transacting at the Mall?

Where:

Q.1 = frequency of people who have e-wallet accounts.

Q.2 = probability of people using e-wallet on transactions at malls.

Q.3 = frequency of people who like transactions with e-wallet.

Q.4 = probability of people increases in transactions at the mall when using e-wallet.

Q.5 = frequency of people’s preferred method of transaction using e-wallet at the mall.

4 Research Results

According to Fig. 1, 100% of participants responded affirmatively. 0% of respondents indicated Nil.

According to Fig. 2, the X axis ranges from 1 to 5, where 1 indicates people who have not made an e-wallet purchase while in a mall and 5 indicates people who have

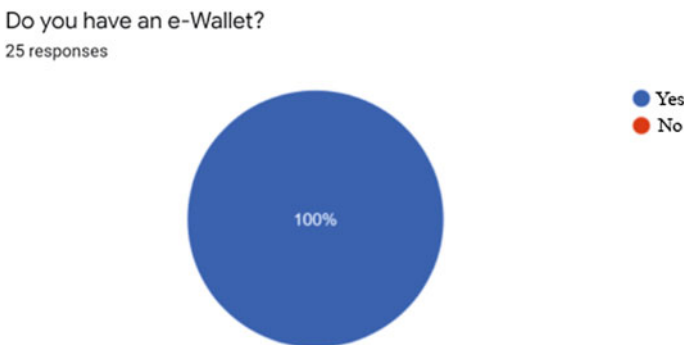


Fig. 1 Owner of an e-wallet

How often do you make a transaction using e-Wallet in a Mall?
25 responses

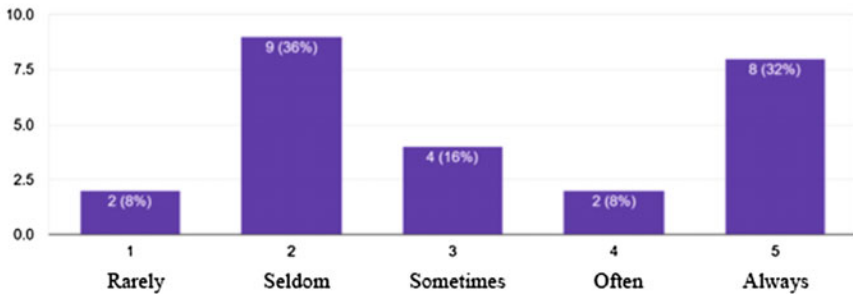


Fig. 2 The frequency of using e-wallet in a shopping mall

frequently made an e-wallet transaction while in a mall. The Y axis represents the response rate.

The vertical line indicates the number of respondents to this questionnaire and the percentage of respondents who share the same opinion. The horizontal line 1–5 shows that they have not made an e-wallet transaction while in a mall, whereas number 5 indicates that they frequently made an e-wallet transaction while in a mall. We can observe from this graph that the majority of those who have completed a transaction have done so using an e-wallet. Number 2 receives the most responses out of the five; this indicates that they rarely utilize an e-wallet to conduct a purchase at a mall. Number 5 has the second highest number of responses in this graph. This implies that some consumers frequently use an e-wallet to make purchases in a mall. The discrepancy between numbers 2 and 3 in this graph indicates that consumers either frequently use e-wallets as their primary way of payment at the Mall or they use them infrequently and just on occasion.

According to Fig. 3, the X axis is 1–5, where 1 indicates that the transaction using an e-wallet is extremely difficult and 5 indicates that the transaction using an e-wallet is extremely simple. The Y axis represents the response rate.

The vertical line indicates the number of respondents to this questionnaire and the percentage of respondents who share the same opinion. The horizontal line 1–5 indicates that on a scale of 1–5, transactions using an e-wallet are extremely difficult to perform. On a scale of 1–5, transactions involving an e-wallet are extremely simple to complete. As observed from the graph, majority of responses choose ease of use as a method of transaction. Number 5 has the most responses, indicating that the majority of people agree that using an e-wallet as the method of transaction is quite simple. However, as indicated by this graph, some people encountered some moderate difficulty while using e-wallet as a method of transaction. Further, it can be attributed to the user’s first time by using e-wallet as a method of transaction.

According to Fig. 4, the X axis is 1–5. A value of 1 indicates that the respondent does not experience any consumptive behavior from themselves, while a value of

How easy do you think making transactions is using an e-wallet?
25 responses

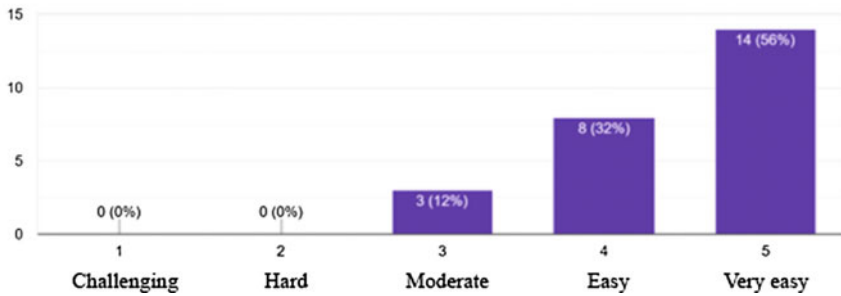


Fig. 3 Ease to use using an e-wallet

5 shows that the respondent experiences much more consumptive behavior when conducting an e-wallet transaction. The Y axis represents the response rate.

The vertical line reflects both the total number of respondents and the proportion of respondents, who share the same opinion. The horizontal line 1–5 shows that the responder feels no consumptive behavior from themselves and significantly more consumptive behavior while completing an e-wallet transaction on a scale of 1–5. The fifth question is the most often answered. This demonstrates that many customers who conduct a purchase using an e-wallet regard themselves to be more consumptive. The remaining numbers on the horizontal scale suggest that some customers continue to have conflicting feelings about feeling consumptive while doing transactions using an e-wallet. When customers utilize an e-wallet to complete a transaction, we may deduce that the majority of consumers have a sense of consumption. This is explained by the fact that people feel they do not possess money but instead see the numbers (Fig. 5).

According to the results of a questionnaire survey sent to young people who frequently visit shopping malls, we can conclude that while most young people own

Do you feel more consumptive when making a transaction using an e-Wallet in a mall
25 responses

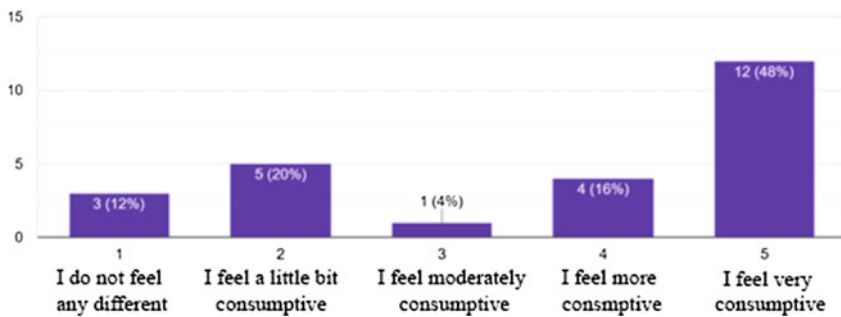


Fig. 4 The consumptive behavior

Do you prefer to use an e-Wallet when transacting at the Mall?
25 responses

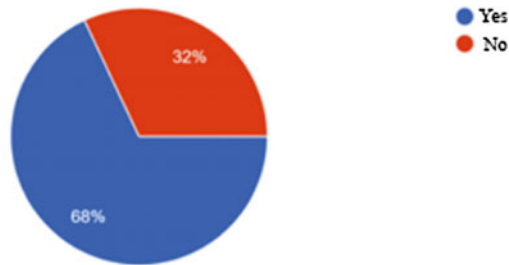


Fig. 5 The preference of the use e-wallet when transaction at the mall

an e-wallet, whether they use it to transact at the mall or not, the comparison is not too far off. According to the poll, almost 40% of people who almost never use an e-wallet in their transactions and almost 60% of people enjoy using an e-wallet. Moreover, all respondents believe that using an e-wallet is simple and convenient; however, 48% report feeling more consumptive when using an e-wallet. As a result, e-wallets are preferred by 68% of respondents, while 32% of people dislike them.

In conclusion, using e-wallet as a method of payment in a shopping mall is quite popular because it is convenient and efficient, it may also backfire, resulting in poor customer behavior.

5 Discussion

Based on various articles and the findings of the proposed research study, we believe that e-wallets will always evolve in the future. Other researchers' findings indicate that the mall's use of e-wallets can be increased again depending on the implementation [5]. People are typically homogeneous or identical, and their homogeneity can be used to develop the acceptance of e-wallet as a primary method of purchasing transactions at the mall, thereby eliminating the need for physical money entirely.

6 Conclusion

Utilizing e-wallet as a means of payment at the mall is a way that is favored by many people since it is practical, quick and fast to use. This assertion may be reinforced by 100% of respondents who have an e-wallet account and more than 56% who accept its usefulness and ease. Therefore, e-wallet is very beneficial and it will continue to expand in the future as a means of payment for a larger sector, additionally it comes

under the government's policy of social distancing. With the presence of e-money, there is no need for actual exchange. Although behind its usefulness, it is inescapable that many individuals become lazy and do not manage their spending adequately as they are complacent with its practicality. The future research study will focus on developing new ways to avoid undermining their individual economies.

Acknowledgements The research is funded by the Research and Technology Transfer Office of Bina Nusantara University as a portion of Bina Nusantara University's International Research Grant contract number: No: 017/VR.RTT/III/2022.

References

1. Wijyanthi I (2019) Behavioral intention of young consumers towards e-wallet adoption: an empirical study among Indonesian users. *Russ J Agric Soc-Econ Sci* 85(1):79–93. <https://doi.org/10.18551/rjoas.2019-01.09>
2. Angelini K, Koesrindartoto DP (2019) E-money or e-wallet? A study of university students' preference in choosing cashless payment systems, 5
3. Vlasov AV (2017) The evolution of e-money. *Eur Res Stud J* XX(1):10. https://www.um.edu.my/library/oar/bitstream/123456789/28785/1/The_Evolution_of_E-Money_2017.pdf
4. Cashless society pros and cons (2021) *The Balance*, 30 Oct 2021. <https://www.thebalance.com/pros-and-cons-of-moving-to-a-cashless-society-4160702>
5. Khatimah H, Halim F (2014) Consumers' intention to use e-money in Indonesia based on unified theory of acceptance and use of technology (UTAUT). *Am Eurasian J Sustain Agric* 7. <http://www.aensiweb.net/AENSIWEB/aejsa/aejsa/July%202014/34-40.pdf>
6. Miliiani L, Purwanegara MS, Indriani MTD (2013) Adoption behavior of e-money usage. *Inf Manag Bus Rev* 5(7):10. <https://ojs.amhinternational.com/index.php/imbr/article/view/1064/1064>
7. Olsen W (2011) A study of satisfaction, loyalty, and market share in Kuwait banks. *Proceedings of the academy for studies in international business*, vol 10, pp 2–7. Retrieved from <https://sloap.org/journals/index.php/irjmis/article/view/664>. Accessed on 4 August 2020
8. Ismanda F (2019) Analisis pengaruh APMK dan E-money sebagai instrumen pembayaran non tunai terhadap tingkat suku bunga dan pertumbuhan ekonomi di Indonesia. *J Dinamika Ekon Pembangunan (JDEP)* 2(2):11. <http://jdep.upnjatim.ac.id/index.php/jdep/article/view/94/49>
9. Untoro R, Dewi K (2013) Pemetaan Produk dan Risiko Pembayaran Bergerak (Mobile Payment) dalam Sistem Pembayaran di Indonesia
10. Foster B (2020) Pengaruh servqual terhadap kepuasan yang berdampak pada loyalitas pengguna E-wallet (OVO). *J Bisnis Manajemen* 14(1):48–56. <https://jurnal.unej.ac.id/index.php/BISMA/article/view/10990/7995>
11. Perkasa HR, Setiawati CI (2020) Analysis of consumer preferences in choosing electronic wallet (e-wallet) in Bandung city. *e-Proc Manag* 7(2):11. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/management/article/view/13454/12997>
12. Wulandari D, Soseco T, Narmaditya BS (2016) Analysis of the use of electronic money in efforts to support the less cash society. *Int Finance Bank* 3(1):10. https://www.researchgate.net/profile/Dwi-Wulandari-5/publication/303478948_Analysis_of_the_Use_of_Electronic_Money_in_Efforts_to_Support_the_Less_Cash_Society/links/57f60a8a08ae886b8980613f/Analysis-of-the-Use-of-Electronic-Money-in-Efforts-to-Support-the-Less-Cash-Society.pdf
13. Wulantika L, Zein SR (2020) E-wallet effects on community behavior. *IOP Conf Ser Mater Sci Eng* 8. <https://doi.org/10.1088/1757-899X/879/1/012121/pdf>

14. Bakar NA, Rosbi S, Uzaki K (2020) E-wallet transactional framework for digital economy: a perspective from Islamic financial engineering. *Int J Manag Sci Bus Adm* 6(3):8. https://www.researchgate.net/profile/Nashirah-Abu-Bakar/publication/340452710_E-Wallet_Transactional_Framework_for_Digital_Economy_A_Perspective_from_Islamic_Financial_Engineering/links/5e873b72a6fdcca789e3b/E-Wallet-Transactional-Framework-for-Digital-Economy-A-Perspective-from-Islamic-Financial-Engineering.pdf
15. Silaen E, Prabawani B (2019) Pengaruh persepsi kemudahan menggunakan E-wallet dan persepsi manfaat serta promosi terhadap minat beli ulang saldo E-wallet OVO. *J Ilm Adm Bisnis* 8(4):9. <https://ejournal3.undip.ac.id/index.php/jiab/article/view/24834>
16. Bich N, Thi HN (2020) An investigation of generation Z's intention to use electronic wallets in Vietnam. *J Distrib Sci* 18(10):89–99. <https://www.koreascience.or.kr/article/JAKO202031064817089.page>
17. Alif MS, Pratama AR (2021) Analisis kesadaran keamanan di kalangan pengguna E-wallet di Indonesia. *Automata* 2(1):7. <https://journal.uui.ac.id/AUTOMATA/article/view/17279/10856>
18. Nassaji H (2015) Qualitative and descriptive research: data type versus data analysis. *Lang Teach Res* 19(2): 129–132

Predictive Study and Classification of Diabetes Using Machine Learning Techniques



Krishan Kumar and Sanjay Patidar

Abstract Diabetes mellitus is a common but deadly disease in humans. It is caused by having excessive sugar levels existing for a long time. It causes around 30–40 lakh deaths worldwide each year. Technology plays a consequential role in the medical industry to assess diabetes prediction. In this research, we trained four machine learning techniques so as to make predictions on whether a person is diabetic or not. The Pima Indian Diabetes dataset is used, which consists of 768 samples, and each sample contains 8 attributes and one target class attribute. Data preprocessing techniques are used to update the raw dataset into a dataset which is suitable for the machine learning models. KNN, Logistic Regression, Support Vector Machine, and Artificial Neural Networks are the techniques used for prediction of diabetes in this research. As a result, K-Nearest Neighbor performed the best, with an accuracy of 76.17%.

Keywords Diabetes prediction · Machine learning · Neural network · Classification · Data preprocessing

1 Introduction

In today's world, technology surrounds us in every possible way, usage of modern day technologies to get some ease in our work, some of the examples we can see are driverless cars, maps GPS navigation, voice controlled devices and many others. All these modern technologies use the data of the real world to train their models and finally test them with the help of data again and again to generate better accuracy. The degree of ease and accuracy of the technology is directly proportional to its usage. Machine learning is also a similar technique in which various algorithms are used to train the model using some part of data and then test the model with some other part of data to generate the outcome. This modern day technique can be used in the prediction of some disease with the help of associated symptoms. Algorithms can

K. Kumar · S. Patidar (✉)

Department of Software Engineering, Delhi Technological University, New Delhi, India
e-mail: sanjaypatidar@gmail.com

train the model in which symptoms values will be used as inputs and generation of outcome will be tested to check the best accuracy and some other different measures, which will help the patient and the medical field to deal with the patient's condition in the earlier stage [1].

Diabetes is a major cause of death, metabolic disorders in humans as well as leads to commercial and productivity loss throughout the world due to lower levels of efficiency of man power. It is a metabolic disorder, characterized by high blood sugar levels which is caused by low insulin production in the pancreas. It increases the risk of long-term complications. It increases the chances of heart disease and about 75% people having this disease die due to coronary artery disease.

In this research, different machine learning algorithms are compared in order to predict risk of someone having diabetes. Classification algorithms are used to classify the target outcomes (1 for diabetic or 0 for non-diabetic) independently. Our study is structured in the following order—Sect. 2 contains literature review. The next, Sect. 3 explains the procedural approach, the machine learning techniques used and the model evaluation. Section 4 discusses about the final result obtained from the research. The last, Sect. 5 contains the conclusion and future work.

2 Literature Review

Diabetes diagnosis and treatment has been a crucial topic in medical research from a very long period of time. With the help of Machine learning, a really good progress has been made in the process of predicting diabetes in people. This prediction is made by the help of machine learning models, which are trained on the dataset consisting of medical information of patients, along with the information whether patients are diabetic or not. After the training phase the model is evaluated by passing the testing data to the model, to check how efficiently the model is working. Kahramanli and Allahverdi [2] used amalgamation of Artificial Neural Networks and fuzzy logics to make a model with good accuracy to predict diabetes. Kumar Dwivedi [3] compared five machine learning algorithms to predict diabetes. The algorithms used were artificial neural networks, classification tree, KNN, SVM, and logistic regression. The author in [4] used two classification algorithms, deep neural networks and artificial neural networks. And also used principal component analysis. Using deep neural networks they achieved better accuracy of 82.67%. Khan and Mohamudally [5] used k-means clustering, neural networks and C4.5 decision tree algorithm to predict diabetes in patients. Bayesian network, Artificial neural network, SVM, Decision tree, and KNN were used to predict diabetes, by Heydari et al. [6]. Temurtas et al. [7] made a model which was trained by Levenberg–Marquardt (LM) algorithm, and the model was combined with multilayer neural network structure.

Rajesh and Sangeetha used classification technique. They used C4.5 decision tree algorithm to find hidden patterns from the dataset for classifying efficiently [8]. Butwall and Kumar proposed a model using Random Forest Classifier to forecast diabetes behavior [9]. Ashiquzzaman et al. [10] proposed a prediction framework for

the diabetes mellitus using deep learning approach where the overfitting is diminished by using the dropout method. Patil proposed Hybrid Prediction Model which includes Simple K-means clustering algorithm, followed by application of classification algorithm to the result obtained from clustering algorithm. In order to build classifiers C4.5 decision tree algorithm is used [11].

Patients can have several symptoms and some of the symptoms and factors are included in the data set like Age, Insulin level, Glucose level, Diabetes Pedigree Function, Blood Pressure level, Skin Thickness, and BMI. Prediction of the outcome from data has been done using various traditional machine learning techniques and artificial neural networks. In order to apply these algorithms, we need to preprocess the data which includes cleaning of the data [12]. Then proposed algorithms are applied and their performances are validated. These prerequisite actions are necessary so that optimal levels of accuracy, precision, and recall can be obtained.

In this paper, classification algorithms are used on the diabetic patient’s data set to predict the outcome of diabetes presence in patients and we achieved a success rate on the test set of 76%. Moreover, we were able to obtain this much accuracy with traditional machine learning approaches, by adding some data preprocessing techniques.

3 Procedural Approach and Methodology

The procedural approach is as follows, as shown in Fig. 1.

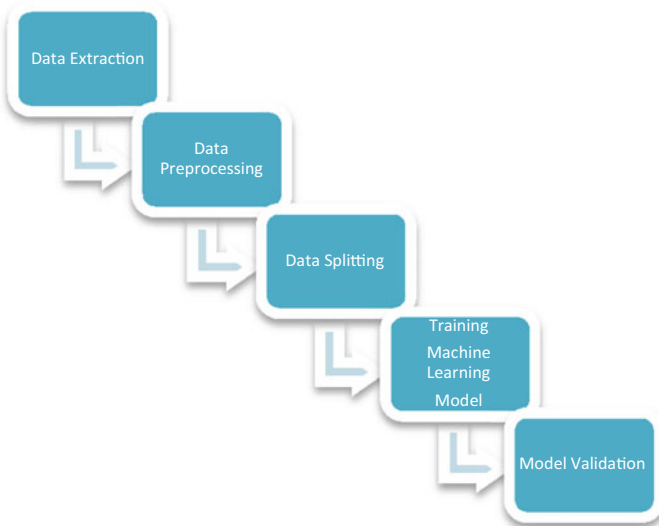


Fig. 1 Proposed architecture

Table 1 List of attributes present in the dataset and their data type

Attribute name	Data type
Pregnancies	Integer
Glucose	Integer
BloodPressure	Integer
SkinThickness	Integer
Insulin	Integer
DiabetesPedigreeFunction	Float
BMI	Float
Age	Integer
Outcome	Integer

3.1 Data Extraction

The data used is the PIMA Indian diabetes dataset. Aim of which is to predict whether or not a patient is diabetic, on the basis of several attributes included in the dataset. Different criteria were used on the selection of these values from the database.

The dataset contains the medical details of 768 different patients, and these medical details were used for classifications. These medical details were stored in 768 rows and 9 columns. Nine columns consisting of 8 attributes and one class column 'Outcome' (diabetic or non-diabetic), as shown in Table 1.

3.2 Data Preprocessing

After the data is collected, it cannot be directly used for the study, therefore it needs to be processed and cleaned to gather suitable information from the raw data useful for the study. The raw data is expected to have many inconsistencies, anomalies, out of bound values, missing values or a format not suitable for our model. Hence, the data needs to be processed in order to use it for our study. Moreover, vast data in present day business, science, industry, and academia scenarios needs complex mechanisms to analyze it. It includes data cleaning, transformation of data; and irregular data reduction tasks, used to reduce the convolution of data, determine and eliminate irrelevant and boisterous elements from the data through feature selection or discretization processes.

Elimination of Null Values

The data was checked for any null values across all features and secondly in individual feature columns, Elimination of Not a number (NaN) values: We replaced the null values of 'glucose' and 'blood pressure' by the mean of the respective attributes, and replaced the null values of the 'skin thickness', 'insulin,' and 'BMI' with the median of the respected attributes.

Table 2 Null values count

Attribute name	Null values	
	Before elimination	After elimination
Pregnancies	0	0
Glucose	5	0
BloodPressure	35	0
SkinThickness	227	0
Insulin	374	0
DiabetesPedigreeFunction	0	0
BMI	11	0
Age	0	0
Outcome	0	0

Table 2 shows the count of number of null values present in the data set before elimination of null values, and also after the elimination of null values.

Evaluation of Class Distribution

The data was checked to be distributed evenly between the target variable outcomes.

3.3 Data Splitting

Entire data was divided into training and testing data. Two-thirds of main dataset was the training data and the rest one-third was used for testing. The training data is the dataset which is given to the model in the beginning for model’s training purpose, which is, to learn from the dataset about the input attributes and the output attribute. The testing data is the dataset which given to the model to model after the training of the model is complete, to check of efficiently the model is working.

So, here the testing data is 1/3rd of the whole data set, and the training dataset is 2/3rd of the whole dataset, as shown in Table 3.

Table 3 Verification of data splitting

Dataset	Percentage of data w.r.t original dataset
Dataset before splitting	768 (100%)
Training dataset	512 (66.6%)
Testing dataset	33 (33.3%)

3.4 Machine Learning Methodologies

K-Nearest Neighbor

Aim of the algorithm is to find the class for the given input. It is a supervised machine learning algorithm. K is the number of neighbors with which we will compare the given input. The input will be assigned to the class whose maximum number of data will be near to the input itself. And the calculation is done with the help of Euclidean distance KNN formula, where x and y are the values of the independent attributes of the neighbor and the new point, and m is the number of independent attributes:

$$\text{dist}(A, B) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (1)$$

Support Vector Machine

It is a labeled training data algorithm that creates a hyperplane that separates the points according to their classes. This hyperplane can be seen in 2D space as a plane splitting line into two pieces, one for each segment. Linear SVM is a technique for generating a classifier that can distinguish between labeled datasets. Given two sorts of points, it tries to maximize the margin geometrically. The letter ‘ Z ’ is utilized to solve the problem of maximum margin and the reparability limitation.

Logistic Regression

It is an algorithm for calculating binary outcomes like zero and one (in our case diabetic or non-diabetic). A linear regression is ineffective for categorizing a binary variable because it predicts continuous values that are beyond the range.

Artificial Neural Networks

The output layer, hidden layer, and input layer are the three layers of an ANN, which are made up of interconnected neurons. The hidden layer has multi-layered structure. The nodes in successive layers are all linked together. Every neuron has an activation function, which is a transformation function that is applied to the node before it is sent to the next layer as input. The result of a node is computed as in Fig. 2.

3.5 Model Validation

The methodologies were performed on the jupyter notebook. The data was analyzed using data visualization techniques and conformed via performance evaluation metrics such as accuracy, precision, recall, $F1$ -score. Cross-validation method was used for evaluation. In k -fold cross-validation, we broke the data into k distinct sets

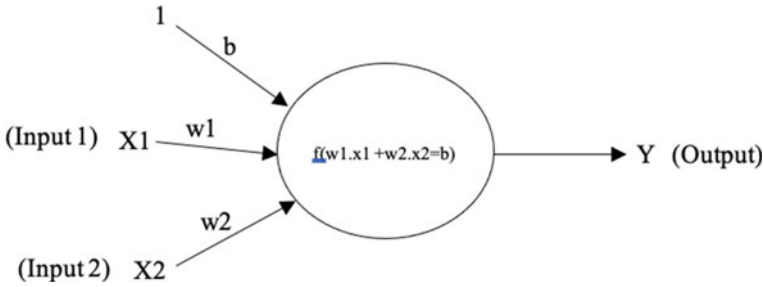


Fig. 2 Neuron in ANN

which are exclusive in nature and have equal size, with one set used for training purpose and other for testing.

Evaluation Metrics

The study is evaluated/validated via confusion matrix using metrics such as accuracy, precision, recall, and *F1*-score.

- True Positives (TP): Total predicted diabetic cases, validated as diabetic.
- True Negative (TN): Total predicted non-diabetic cases, validated as non-diabetic.
- False Positives (FP): Total predicted diabetic cases, validated as non-diabetic.
- False Positives (FN): Total predicted non-diabetic cases, validated as diabetic.

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP}, & \text{Recall} &= \frac{TP}{TP + FN}, \\
 \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN}, & \text{F1-score} &= \frac{2TP}{2TP + FP + FN}
 \end{aligned}$$

In this study for model validation confusion matrix have been used. In Fig. 3, confusion matrix of KNN model is shown in Table 4. Figure 4 shows the learning curve of KNN model which represents the training score and cross-validation of the KNN model. Table 5 shows the accuracy, precision, recall, and *F1*-score of the KNN model. Similarly, all the other algorithms were validated.

4 Results

In this research, we have performed diabetes prediction on PIMA Indian dataset, to predict diabetes a person is diabetic or not. First data is preprocessed by eliminating all the Not a number (NaN) values, by replacing them by the mean or the median of the respective attribute. Then the prediction was made by using four different machine learning algorithms KNN, SVM, logistic regression, and artificial neural networks. And among all the four algorithms, KNN showed the best accuracy of

Fig. 3 Learning curve of KNN

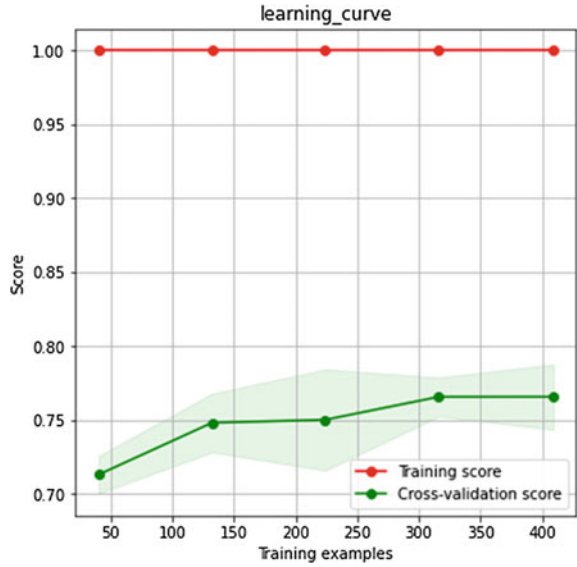


Table 4 Confusion matrix

Output		Predicted values	
		Diabetic	Non-diabetic
Actual values	Diabetic	TP	FN
	Non-diabetic	FP	TN

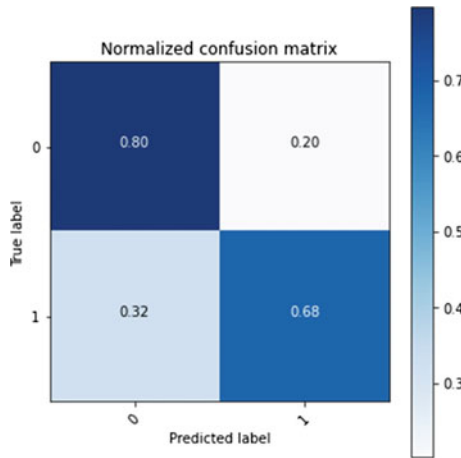


Fig. 4 Confusion matrix of KNN

Table 5 Accuracy, precision, recall, and *F1*-score of the KNN model

Accuracy	0.76
Precision	0.67
Recall	0.59
<i>F1</i> -score	0.63

Table 6 Performance measures of all the four algorithms

	Accuracy	Precision	Recall	<i>F1</i> -score
KNN	0.76	0.67	0.59	0.63
SVM	0.75	0.69	0.53	0.60
Logistic regression	0.73	0.63	0.52	0.57
Neural network	0.73	0.62	0.61	0.62

76%. Table 5 shows all the values of accuracy, precision, recall, and *F1*-score of all the four machine learning algorithms are shown in Table 6.

5 Conclusion and Future Work

In this study we ought to resolve the complications occurred during diagnosis of diabetes disease. The study put forwards an light on different machine learning algorithm such as the SVM, KNN, logistic regression, and ANN for predicting whether a patient is diabetic or not. It was concluded that out of all KNN performed best with an accuracy of 76%, hence it is a better option for classifying complex data.

In future we will try to come up with much better mechanisms and a much larger data set in order to increase the accuracy to help medical practitioners to treat patients and overcome this deadly disease.

References

1. Kalyankar GD, Poojara SR, Dharwadkar NV (2017) Predictive analysis of diabetic patient data using machine learning and Hadoop. In: International conference on I-SMAC. 978-1-5090-3243-3
2. Kahramanli H, Allahverdi N (2008) Design of a hybrid system for the diabetes and heart disease. *Expert Syst Appl Int J* 35(1–2)
3. Kumar Dwivedi A (2017) Analysis of computational intelligence techniques for diabetes mellitus prediction. *Neural Comput Appl* 13(3):1–9
4. Vijayashree J, Jayashree J (2017) An expert system for the diagnosis of diabetic patients using deep neural networks and recursive feature elimination. *Int J Civ Eng Technol* 8:633–641
5. Khan DM, Mohamudally N (2011) An integration of K-means and decision tree (ID3) towards a more efficient data mining algorithm. *J Comput* 3(12)

6. Heydari M, Teimouri M, Heshmati Z, Alavinia SM (2015) Comparison of various classification algorithms in the diagnosis of type diabetes in Iran. *Int J Diabetes Dev Ctries* 1–7
7. Temurtas H, Yumusak N, Temurtas F (2009) A comparative study on diabetes disease diagnosis using neural networks. *Expert Syst Appl* 36(4):8610–8615. <https://doi.org/10.1016/j.eswa.2008.10.032>
8. Rajesh K, Sangeetha V (2012) Application of data mining methods and techniques for diabetes diagnosis. *Int J Eng Innov Technol (IJEIT)* 2(3)
9. Butwall M, Kumar S (2015) A data mining approach for the diagnosis of diabetes mellitus using random forest classifier. *Int J Comput Appl* 120(8)
10. Ashiquzzaman A, Tushar AK, Islam M, Kim J-M et al (2017) Reduction of overfitting in diabetes prediction using deep learning neural network. arXiv preprint [arXiv:1707.08386](https://arxiv.org/abs/1707.08386)
11. Patil BM, Joshi RC, Toshniwal D (2010) Association rule for classification of type-2 diabetic patients. In: *ICMLC'10 proceedings of the 2010 second international conference on machine learning and computing*, 09–11 Feb 2010
12. Fatima M, Pasha M (2017) Survey of machine learning algorithms for disease diagnostic. *J Intell Learn Syst Appl* 09(01):1–16. <https://doi.org/10.4236/jilsa.2017.91001>

Review on Modern Techniques Behind Virtual Cloth Try-On



Prajakta Joglekar and Vinaya Gohokar

Abstract The proposed study reviews different techniques such as image segmentation and image augmentation to implement virtual try-on. Generally, parser-free image segmentation leverages more accuracy in the training phase since it works with teacher–student model by considering the parser-based model as teacher. In terms of image augmentation techniques, Augmented Reality (AR)-based approach is easier to some extent and handles system level complexity efficiently compared to the image wrapping approach but this requires large computational power. The main objective of the proposed study is to assist the researchers in utilizing any one of the discussed methods such as image segmentation, image warping, and augmented reality to develop an innovative and reliable technology for enabling virtual try-on in both apparel as well as non-apparel market as per the consumer requirement.

Keywords Image segmentation · Parser · Image warping · Features · Augmented reality · Marker

1 Introduction

Technology has been growing rapidly in many industries in recent years, remarkably in the clothing sector in order to fulfill the customer needs and expectations. One of these requirements is that the trial of garments before being purchased. Retail E-commerce sales hit a worldwide profit of 4.28 trillion dollars in 2020, wherein the e-commerce revenues expected to reach 5.4 trillion dollars in 2022. When it comes to fashion, one of the most significant offline experiences that online buyers overlook is the changing room, where a garment item can be tried on by the customer. However, the recent covid-19 pandemic has changed the way of living ranging from attending

P. Joglekar (✉) · V. Gohokar
School of Electronics and Communication Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune 411038, India
e-mail: prajaktajoglekar0106@gmail.com

V. Gohokar
e-mail: vinaya.gohokar@mitwpu.ac.in

virtual work meetings to buying favorite apparel online everyone is getting more inclined toward achieving utmost advantages of technologies. Even though online shopping has become more common in today's culture, there are some customers who still opt for visiting a store for trying out an outfit before buying it. Currently, most of the shopping malls and apparel shops have trial rooms for customers to try out the apparels before buying it. In the pandemic time, these trial rooms are restricted for the customers in order to maintain no contact policy and hygiene. Moreover, trial rooms could sometime affect the users' privacy and lastly, providing trial rooms is difficult task for small shop keepers as they are more concerned with the space. The proposed idea can tackle these issues. This system allows the user to try clothes virtually without physically wearing it. In this case, real-time human body detection and tracking techniques are used to obtain the user's body shape, and augmented reality techniques are used to superimpose the cloth image on the previously obtained user's image. Clothes can be tried on virtually using Image Segmentation techniques such as Appearance Flow, Feature Mapping, Texture Preservation, Deep Learning, Neural Network, Image Warping, and Augmented Reality. Some of these techniques will be discussed further below.

The Image Processing domain has never failed to captivate researchers with its diverse range of applications. Many innovations have happened in this field over the last few decades. Since Virtual Try-on relies heavily on images as input data, image processing techniques act as a basic technology. Image segmentation is one of these techniques in which the developer can take a user image as input and divide it into the required body shapes to which a cloth image can be added. A parser, a preprocessor component that divides into small components is used for performing the segmentation task. Additionally, a significant research work has been conducted in parser-free algorithms, in which the developer takes pre-trained decomposed data as an input image and then adds a garment image to the desired body part. Next technique discussed in this paper is Image wrapping. Developer can experience a perspective transformation and wrap transformation to add the cloth image on the user's required body part. Despite the fact that all of these techniques initiate with image processing, they can be further enhanced by using deep learning and neural networks. Finally, Augmented Reality (AR) is a prominent technology that tends to fall under the broad scope of Industry 4.0. These reality techniques improve the user experience in a variety of fields by creating virtual worlds based on the requirement of the application. Text, visuals, and audio are all integrated in augmented reality, and additionally the users' interaction with the real-world is also enhanced. AR works based on the idea of superimposing digital data onto the physical world, typically in the form of graphical augmentation [1, 2]. In Augmented reality, by superimposing visual phenomena such as image, video or text on the real-time object, a virtual world can be created for user in order to perceive the experience without physically building a environment. Furthermore, AR can be classified based on the image type, which is being used for AR, they are 2D and 3D images. In 2D image type, again marker-based and marker-less are considered as two different techniques, which have been further discussed below. In next chapter, the methodology behind all these techniques are discussed in detail.

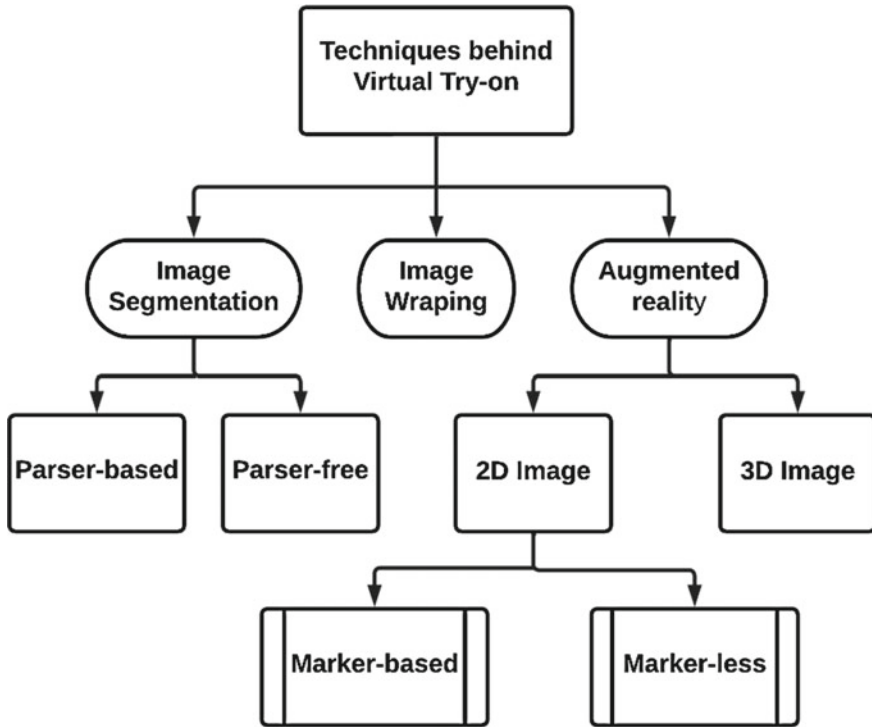


Fig. 1 Classification of virtual try-on

As shown in Fig. 1, the different techniques behind virtual try-on have been largely divided into three categories, they are classification based on image segmentation, image warping technique and augmented reality technique. Furthermore, they are sub-categorized and explained.

2 Image Segmentation Techniques

As shown in Fig. 1, this category can be divided into parser-based and parser-free with respect to the parser information, which is used to suppress the individuals' garments and generate try-on images. Human parsing is a technique used for segmenting a person image into several portions with fine-grained semantics like body organs and garments to gain a complete knowledge from the image data. Initially, the basic local formation catches the cross-layer context by hierarchically combining the global semantic data and local fine features over the distinct convolutional stages. Furthermore, in the interim tier of the Co-CNN, the global image-level label estimation is employed as an auxiliary objective, and the corresponding outputs are then used to guide feature learning in the following convolutional stages to select the global

image-level environment. Thirdly, Co-CNN incorporates semantic edge context, wherein high-level semantic constraints are applied to drive a pixel-wise labeling. Finally, the neighborhood voting are developed as instinctual sub-components of the said Co-CNN to accomplish the local label uniformity in both training and testing phases [3, 4]. A large-scale dataset composed of representative instances with varied garment manifestations, strong explication, limited self-occlusions, truncation at image borders, diverse viewpoints, and background clutters are beneficial to prepare a human parsing with a practical value in real life applications [5]. On the other hand, the inference stage of parser-free approaches only require the human image and the outfit image as input. They were exclusively created to counteract the detrimental consequences of poor parsing results. These approaches tend to start with a parser-based teacher model and work their way down to a parser-free student representation [6]. Except for the inputs, the parser-based warping and parser-free warping use the same framework. Both, Parser-based and Parser-free segmentation is discussed in the below section.

2.1 Parser-Based Image Segmentation

The esthetics of apparel is strongly influenced by user's body forms, and therefore the relocation of desired fashion product is determined by the position of a specific human anatomy as well as the body shape. A human parser is utilized to produce a human segmentation map with distinct regions for representing different areas of the users' body.

The deformation of the desired clothing image to fit the position of a person is one of the key technical challenges faced by the virtual try-on domain. To that intent, as a prior requirement, apparel human representation is provided by encompassing a combination of features such as stance, body anatomy, face, and hair. To create a realistic synthetic dressed figure in the similar posture wearing the desired garment, as well as a correlating clothing region mask, a multi-task encoder-decoder chain is used. The mask is then utilized to warp the targeted clothing item to accommodate the deformations by using it as a guide [7]. Since diverse deformations of clothes result from different human positions, pose information is modeled effectively by using a state-of-the-art pose estimator. A person's calculated pose is defined by 18 key point positions.

Figure 2 shows the heat map and pose map, which can be used in the segmentation process. To make use of their spatial arrangement, each key point is converted in a from of heatmap, with an 11×11 surrounding filled in with ones and zeros. All the key point heatmaps are combined through an 18-channel pose heatmap. A cutting-edge human parser is used to represent the human segmentation map in which multiple sections represent separate parts of the human body. The segmentation map is then converted to a 1-channel binary mask, with 1 signifying the human body (other than the face and hair) and 0 signifying everything else. To eliminate artifacts

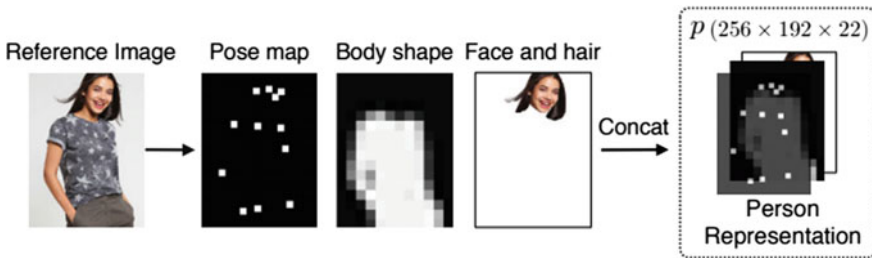


Fig. 2 Heat map and pose map for person's representation [7]

when the body image and target apparel clash, this binary mask obtained straight from I is down sampled to a lesser resolution 16×12 as illustrated in Fig. 2.

When producing additional images, the human parser captures the RGB components of the user's face regions in order to incorporate distinguishing data as shown in Fig. 3. Finally, these three feature maps are convolved to generate an apparel-agnostic user portrayal by rescaling them to the same resolution. A multi-task encoder-decoder architecture is implemented to construct a dressed person's image by including a clothing mask. The anticipated clothes mask will be applied to refine the findings in addition to directing the network to concentrate on the garment region. The refinement network is a fully convolutional four-layer model, where each of the initial three layers have $3 \times 3 \times 64$ filters followed by Leaky ReLUs, while the last layer outputs the composition mask with 1×1 spatial filters and a sigmoid activation function to scale the output to $(0, 1)$ [7].

2.2 Parser-Free Segmentation

A recent pioneering effort portrayed the parser-free network without employing human segmentation for virtual try-on in order to eliminate the necessity of utilizing precise masks to lead the try-on models. A pre-train phase and a parser-free learning phase are included in the most widely used parser-free framework. The initial phase pre-trains a PB-Warp and PB-Generator module, and the second phase distills a PF-warp module by using the PB-warp and improves a PF-generator [8].

The warp module compresses the clothes to match the human posture and body structure by retaining the garment's details. To produce the final warped garment, multiple convolution layers are tailored to resize the input images to various sizes by keeping the garment wrapped. However, while the worn garments are synced with the actual data, the human pose is not considered as it may collapse the results, especially if the pose is intricate. A posture awareness loss LW is recommended to ensure that the PF-Warp module alters the garments by focusing more on poses. It considers a collection of false images with different sampled garments but in the same pose for each intended garment. Then, for performing LW optimization, PF-Warp is used to

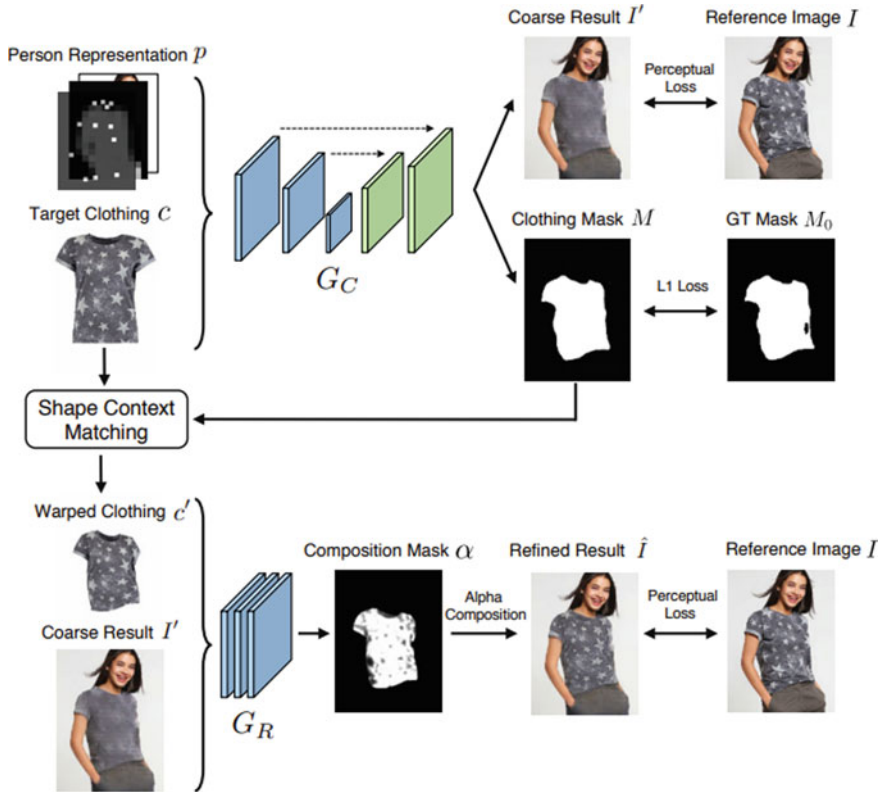


Fig. 3 Parser-based image segmenting [7]

create a set of garments. The generator module uses warped images and reference person images as input to create final try-on images, as shown in Fig. 4.

Since the existing PF-generator is unable to reduce the distraction of reference persons’ clothing, only low-resolution try-on images could be generated. Two-way feature extractor can be used to gradually fuse the retrieved data into final try-on images by using a one-way generation approach to synthesize high resolution images by including higher details of target garments and reference users.

PB approaches have recently had a lot of success but their reliance on correct parser information has restricted their usefulness in real-world situations, notably when human poses are complicated. The over-reliance on human parsing, as demonstrated in the first row makes PB systems vulnerable to erroneous parsing results and susceptible to producing unrealistic images with apparent artifacts.

As illustrated in Fig. 5, parser-free approaches do not require parser information. Recently, many researchers have proposed a student–teacher structure to guide the student model in replicating the capabilities of parser-based approaches without requiring the use of a human parser. PF-AFN redesigns the warp module to extract

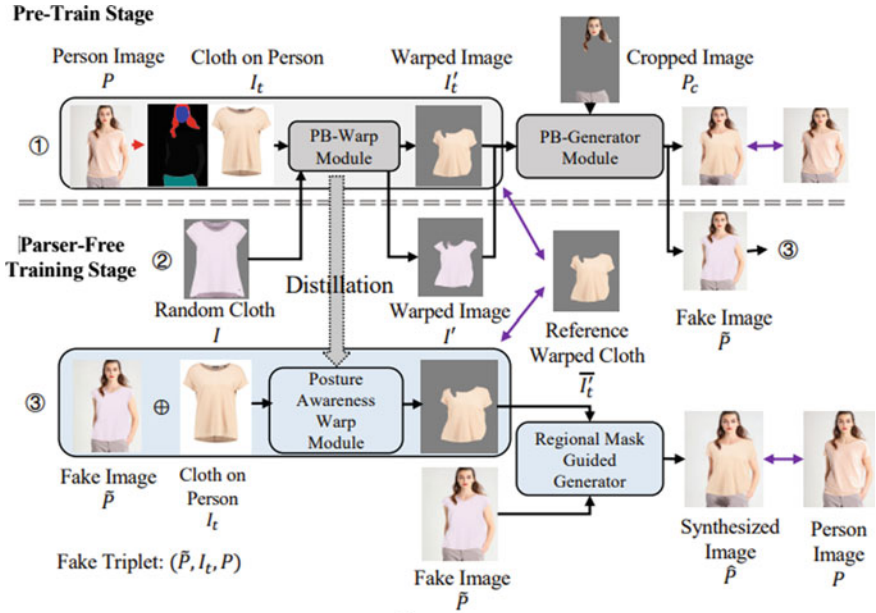


Fig. 4 Parser-free image segmentation [9]

the esthetic patterns between the reference individuals and captured cloth images in order to produce a more reliable try-on.

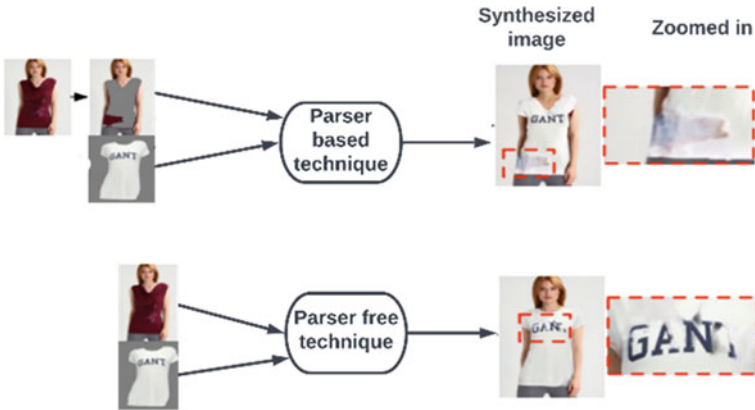


Fig. 5 Parser-based versus parser-free [9]

3 Image Warping Technique

The main concept behind this technique is to consider a try-on image of apparel that the model has tried on and warp it as per user’s physical appearance. This is defined by specific feature points before mapping the warped image into the user’s image.

In brief, 4 key techniques are used (as shown in Fig. 6):

- (1) Generating a dataset of 2D try-on images of garments with the model:

Garment images should be coupled to body models, and in a virtual try-on system, real-time physical simulation should be implemented. To cut the price of generating

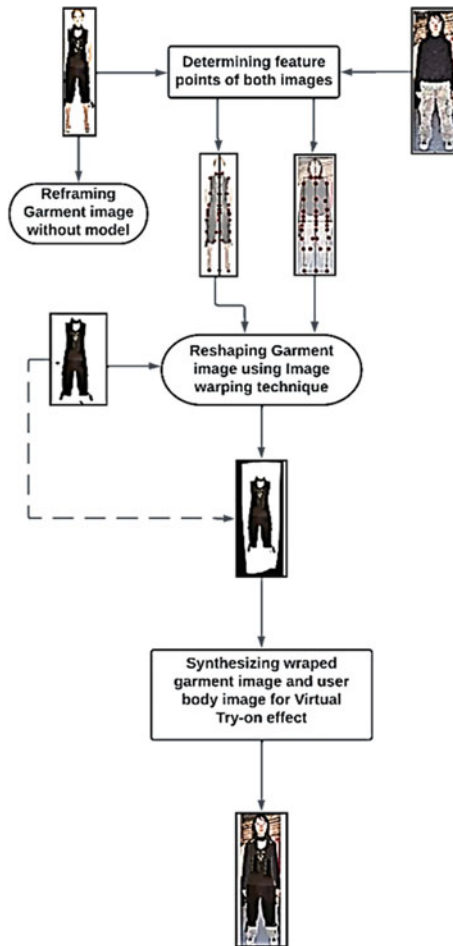


Fig. 6 Image wrapping technique for virtual try-on [10]

and computing the simulation on mobile apps, auto-skinning, local physical simulation, and transparent textures are used. The skin weights of clothing models are first initiated, followed by the correction of skin weights with immersion depending on each frame of the body model animation, and lastly the updated garment designs are generated. By lowering the upper vertex weight value and adding other weight values, the skin weights are adjusted. After that, the gap between the garment and the human body model is estimated. When the distance is greater than a limit, actual simulation based on spring-mass models will be employed; otherwise, skeletal animation will be used. For interference cancellation across clothing models and body models, many sphere colliders are employed to approach body models.

- (2) Procuring the user's local features in the 2D image as per his/her physical appearance:

Human body detection and tracking by using camera to get user body as an input can be the first step toward building a virtual try-on system. Feature points detection of human body can demonstrate estimated body position of user. At least 20 feature points are required to fetch position of the body for tracking. There is different method by which human body position can tracked, among those three predominantly use techniques are, Kalman Filter (KL), Mean-Shift Filter (MSF). Partial Least square (PLS).

- (3) Warping the garment images per the correlation of the features extracted from the model and user by utilizing an image processing algorithm:

Two-dimensional images might be considered two-dimensional entities. To put it another way, an image is a 2D plane with a bounded domain and RGB data attached with each pixel. In this perspective, image warping is technically a plane transformation, and RGB data is transformed as per the conversion of their associated coordinates. In general, the image warping approach has two stages. The first is the computation of mapping parameters that are dictated by the predetermined mapping of a fixed number of input variables. The second method is to generate warped images using a re-sampling process based on mapping variables, which contributes to interpolation theory.

Assume (p, q) and (u, v) denote the source and warped image points, respectively. The mapping link among (p, q) and (u, v) for image warping can be stated in the form:

$$\begin{cases} u = h_1(p, q) \\ v = h_2(p, q) \end{cases}$$

where h_1, h_2 are considered as the mapping functions.

- (4) Combining the warped image to the person's existing 2D image to demonstrate the apparel's try-on effect:

Attaching the body and head models is the final step in creating the user's model. The model attachment must be done in such a way that the final model appears to be

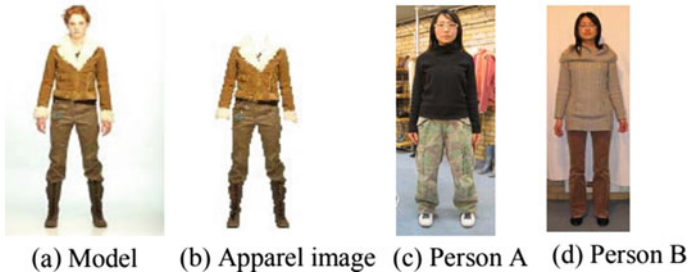


Fig. 7 Image warping for virtual try-on results [10]

a complete body as shown in Fig. 7. Two considerations must be considered in order to reach this goal. Firstly, the neck is adjusted along the x -axes and y -axes using the normalized mean of the body's neck size and the head's neck size. As a result, the link plane between the body and head models in the x - z plane is matched. This stage produces a three-dimensional model of the individual wearing the clothing.

4 Augmented Reality

Augmented reality can be further categorized based on Image which can be used for AR that is, 2-dimensional and 3-dimensional Image. According to the application and implementation model developer can choose dimensions of the image.

4.1 2D Image

Virtual 2D try-on helps buyers by providing an intuitive and realistic Try-on garment image, and it has a lot of commercial potential. Existing three-dimensional try-on solutions, on the other hand, rely often on labeled three-dimensional human shapes and apparel templates, which restricts their use in real-world circumstances. Virtual try-ons in two dimensions are a speedier way to manipulate clothed individuals. In 2D AR based on Marker-based approach and Marker-less approach. Basically, Marker as data input that might be in terms of an image or a RFID tag or a QR/Barcode which helps developer to focus on which particular area in plane, the image has to be augmented. Marker-based approach uses similar markers to track the data whereas Marker-less approach uses techniques such as Simultaneous Location and Mapping (SLAM), Multiplayer rendering, co-ordinate tracking. Both approaches are discussed briefly.

4.1.1 Marker-Based AR

A marker-based AR approach can also be used to design virtual try-ons, in which vision-based tracking methods fall back on the morphological points of interest or visual patterns of the environment to implement detection and tracking. Tracking is attained using fiducial marker detection, feature detection, or edge detection, feature extraction and mapping in this technique. Virtual fitting augmented reality is the most recent technology that allows a user to view a model of herself or himself wearing previously recorded garments. Garment database is produced to store three-dimensional designs of persons wearing clothes and user preferences. A revolutionary body modification technique is introduced to scale individual areas of the model independently while maintaining the harmony of the complete body to design a body model depending on anthropometric criteria of a user.

To solve the image classification challenge, various strategies are used. Support Vector Machines (SVM) have grown in prominence because of their strong generalization capabilities. Various feature descriptors have been used to empower this approach as shown in Fig. 8.

The following is a formal definition of the image classification challenge. A collection of training data is provided as $L = \{x_i, y_i\}_{i=1}^n \in X_n \times Y_n$ of n images along with its labels $y \in Y = \{1, \dots, c\}$, the SVM classifier is defined as:

$$y' = \arg \max_{k=1, \dots, c} gk(x) = \sum_{i=1}^n \alpha_i^k(x, x_i) + b^k$$

where c is the number of classes, and y' is the class label suggested by the SVM classifier. In this formula, $g(x)$ is considered as the discriminant function and k is considered as the index of this function, and the set $[\alpha_i^k]_{i=1}^n$, n are considered as the signed dual coefficients [10]. However, a model of a person that fits the person's dimensions is selected during the model selection process, where these models must be updated at runtime to fit the user's body as shown in Fig. 9.

Anthropometric parameters are also required to transform different body models with a high level of accuracy. Linking the user's body and head models is the final stage involved in creating the augmented try-on model. The model integration must be easily managed so that the end model will resemble as an entire body. Two

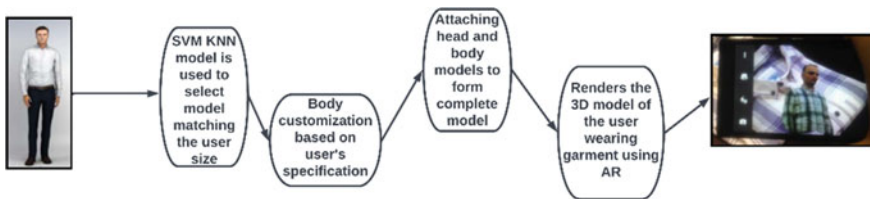


Fig. 8 Marker-based AR workflow [11]



Fig. 9 Marker-based virtual try-on using augmented reality results [11]

considerations must be addressed in order to meet this objective. Initially, the neck is resized along the x -axis and y -axis by using the harmonized means of the body's neck size and the head's neck size. As a result, the x - z coupling plane between the body and head model is aligned. This method generates a three-dimensional model of a customer wearing the outfit.

4.1.2 Marker-Less AR

In augmented reality technique, the experiments demonstrated that this technique has outperformed direct silhouette configuration and feature-based model evaluation techniques in terms of quality and reliability. The observation to assess the efficacy of body personalization and face generation components has revealed the significance of using body tailoring to provide feasible user models. Here, the example of Virtual Try-on for wristwatch is discussed.

The Kinect sensor depicts RGB and depth information from the plot. Kinect SDK is used to trace among 3 different types of data in order to interpret depth, RGB, and skeleton information. The Kinect SDK provides three-dimensional perspectives of recorded skeleton joints from which we can achieve elbow and wrist positions for one arm. Using these two joints, the primary alignment of the arm in three-dimension is identified, and a chosen digital model (e.g., virtual watch) is arranged at the wrist location with the measured orientation perspective. Two approaches are utilized to accomplish a true representation of an entire RGB plot and a precise object model at the same time. First, RGB data is retained and demonstrated as a plane background appearance, and afterward an OpenGL fragment shade is produced as shown in Fig. 10.

Both marker-based and marker-less approaches used for AR-based virtual try-on can be used effectively if sufficient resources are available as Fig. 11. Based on application and on system available computation speed, one can choose the approach. Broad Comparison between discussed two approaches is given in Table 1.

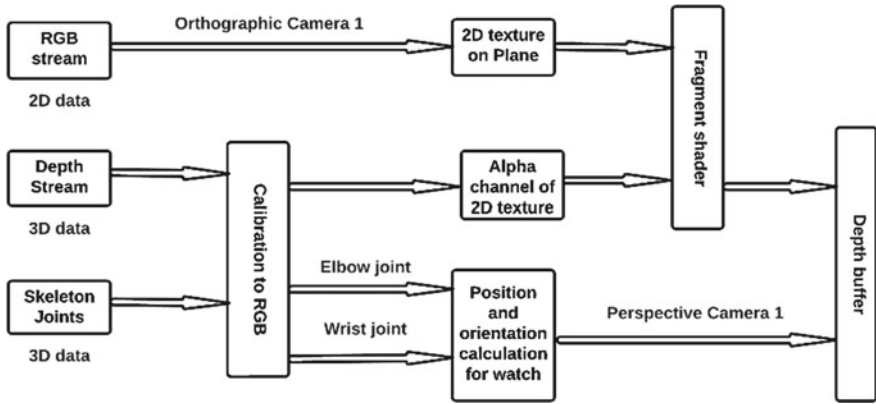


Fig. 10 Marker-less approach for virtual try-on using AR [12]



Fig. 11 Marker-less virtual try-on using augmented reality results [12]

Table 1 Comparison between marker-based and marker-less AR approach [13]

Marker-based approach	Marker-less approach
Reference image with good feature points needed	No need of reference image
Script is needed to detect feature points and to render virtual objects on target image	Vuforia plug-in inbuilt uses SIFT algorithm to detect feature points and horizontal surface detection
Use of this methodology fails in augmenting complex virtual objects	Any digital content like text, image, and video can be augmented in this method easily by importing required assets available in Unity3D

3D Image

Three-dimensional try-on is an attempt to fit a specific clothing item onto a three-dimensional person body has received a lot of research attention due to its potential scientific and commercial significance. Researchers’ attention has recently shifted away from physics or scan-based systems toward learning-based three-dimensional try-on approaches that outfit a three-dimensional person directly from two-dimensional images by eliminating the need for costly modeling or 3D detectors. Nonetheless, almost all of these learning frameworks rely on a pre-programmed digital inventory and are based on the parametric SMPL paradigm, which limits their real-world relevance. Furthermore, due to the optimization cost incurred by the parametric three-dimensional representation, the inference rate of these modern techniques remains insufficient.

By forecasting depth maps, non-parametric three-dimensional rebuilding has already been developed to accurately preserve feature descriptors. Molding individuals create a prerendered three-dimensional figure by predicting the frontal and rear depth maps by using a unique RGB image. The network comprises three components that conduct the following activities as shown in Fig. 12.

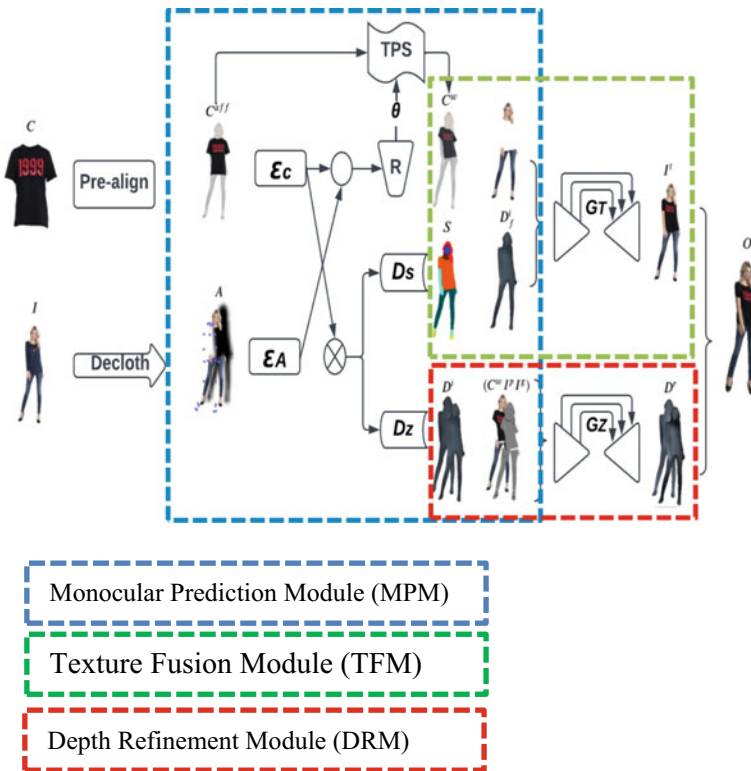


Fig. 12 3D AR workflow [14]

- (a) Monocular Prediction Module (MPM): Acquiring the cloth-agnostic individual depiction via the declcloth method reshape the in-shop clothing C to the warped clothing C^w via a self-adaptive pre-alignment accompanied by a TPS transition, forecasting a person segmentation S , and assessing a preliminary double-depth map D^i .
- (b) Depth Refinement Module (DRM): Given a double-depth map D^i , warped garments C^w , preserved user part I^p , and shadow details I^s as input data, this component modifies the interim depth map and generates more local features (including garment folds and face shape) by integrating a new depth gradient constraint.
- (c) Texture Fusion Module (TFM): Fusing the distorted garments and the preserved texture information to render the results under the guidance of the semantic arrangement from MPM. We can directly obtain colored point clouds and triangulate them to generate the three-dimensional clothed human O wearing the intended garments and with its identity retained after I^i and the corrected depth map D^r are spatially aligned, yielding an RGB-D representation as shown in Fig. 13 [13].

In this research, the authors suggest a computationally feasible three-dimensional Virtual Try-On System that combines the benefits including both 2D and 3D methodologies to generate a 3D try-on topology from 2D data. The three-dimensional try-on challenge is decomposed into a 2D try-on and a user’s body depth estimate problem using this approach. In future research, the author will look into whether the two can enhance one another in cyclic pattern. This method entails a two-stage warping



Fig. 13 3D AR virtual results [13]

mechanism, as well as segmentation and depth guiding, to achieve more accurate texture fusion outputs. A new depth gradient restriction is added to produce precise depth maps.

5 Applications and Future Scope

Technology advancements, especially in the field of computer vision, have made it possible to automate manual processes. Fashion Apparel is one of the most heavily influenced industries by technological innovation. A great experience may be created for both retailers and consumers using computer vision-based solutions. Because the apparel industry is primarily visual, establishing platforms that can comprehend fashion in images can deliver a next-level customer experience equivalent to online fashion shopping. Since the apparel industry is primarily visual, it may be dealing with computer vision to recognize images in the same way that we do by attempting to make computers comprehend images. These systems can be used for outfit recommendation, AR fashion designer, shopping assistant. Techniques like deep learning with blend of AR can be used for non-apparel try-on like makeup, jewelry, handbag, etc.

6 Conclusion

In this paper, various techniques including Image segmentation and Image Augmentation are discussed. Parser-free image segmentation has achieved high accuracy in training since it works with teacher–student model by considering Parser-based model as teacher. In terms of image augmentation techniques, Augmented Reality (AR)-based approach for 2D and 3D image is comparatively easier to some extent with less system level complexity than image wrapping approach but it requires large computational power. This paper will help researchers to proceed with one of the discussed methods.

References

1. Grubert J, Langlotz T, Zollmann S, Regenbrecht H (2017) Towards pervasive augmented reality: context-awareness in augmented reality. *IEEE Trans Visual Comput Graph* 23(6):1706–1724
2. Liang X, Gong K, Shen X, Lin L (2019) Look into person: joint body parsing & pose estimation network and a new benchmark. *IEEE Trans Pattern Anal Mach Intell* 41(4):871–885
3. Liang X, Xu C, Shen X, Yang J, Tang J, Lin L, Yan S (2017) Human parsing with contextualized convolutional neural network. *IEEE Trans Pattern Anal Mach Intell* 39(1):115–127
4. Sen H, Yi-Zhe S, Tao X. Style-based global appearance flow for virtual try-on

5. Han X, Wu Z, Wu Z, Yu R, Davis LS (2018) VITON: an image-based virtual try-on network. In: 2018 IEEE/CVF conference on computer vision and pattern recognition. IEEE, Salt Lake City, UT, pp 7543–7552
6. Ge Y, Song Y, Zhang R, Ge C, Liu W, Luo P (2021) Parser-free virtual try-on via distilling appearance flows. In: 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR). IEEE, Nashville, TN, pp 8481–8489
7. Chao L (2022) RMGN: a regional mask guided network for parser-free virtual try-on
8. Zeng X, Ding Y, Shao S. Applying image warping technique to implement real-time virtual try-on based on person's 2D image. IEEE
9. Sekhavat YA (2017) Privacy preserving cloth try-on using mobile augmented reality. IEEE Trans Multimed 19(5):1041–1049
10. Gunes S, Sanli O, Ergun OO (2015) Augmented reality tool for markerless virtual try-on around human arm. In: 2015 IEEE international symposium on mixed and augmented reality—media, art, social science, humanities and design. IEEE, Fukuoka, Japan, pp 59–60
11. Pooja J, Vinay M, Pai VG, Anuradha M (2020) Comparative analysis of marker and markerless augmented reality in education. In: 2020 IEEE international conference for innovation in technology (INOCON). IEEE, Bengaluru, India, pp 1–4
12. Zhao F, Xie Z, Kampffmeyer M, Dong H, Han S, Zheng T, Zhang T, Liang X (2021) M3D-VTON: a monocular-to-3D virtual try-on network. In: 2021 IEEE/CVF international conference on computer vision (ICCV). IEEE, Montreal, QC, pp 13219–13229
13. Ghodhbani H, Neji M, Razzak I, Alimi AM (2022) You can try without visiting: a comprehensive survey on virtually try-on outfits. Multimed Tools Appl 81(14):19967–19998
14. Mohammadi SO, Kalhor A (2021) Smart fashion: a review of AI applications in virtual try-on & fashion synthesis. JAICN 3(4):284–304

Identification and Representation of Spectral Anomalies in an Abandoned Quarry by Remote Sensing



C. Gambardella and R. Parente

Abstract Soil pollution represents a problem that must be addressed by exploiting all the resources we have available. The transformation of the built environment requires a large amount of raw materials that are extracted from the quarries, exploiting to the full abandoned places often used as a deposit of materials harmful to humans. In this work, images detected in a specific hyperspectral aerial remote sensing campaign with Itres CASI 1500 sensor were analyzed. The measurements were stored in a georeferenced image with 36 levels, one for each detected wavelength. The hyperspectral images were post-processed using vegetation indices, PCA and RXD algorithms. The survey methodology made it possible to detect spectral anomalies that require greater investigation with specific methods.

Keywords Remote sensing · Airborne sensors · Vegetation indices · Principal component analysis

1 Introduction

Although the last decades have been characterized by accelerated technological innovation and the improvement of industrial plants, production activities continue to have a profound impact on the quality of the various environmental matrices due to atmospheric and water emissions and to produce toxic waste. The processes of pollution and other forms of environmental stress have not spared the lands used for agriculture, especially those affected by high industrial pressure. Consequently, the problem of the study, protection and recovery of these territories has arisen [1].

By soil pollution we mean the chemical–physical alteration of the soil. This transformation occurs every time man exploits the soil both for agriculture and for other commercial purposes. By soil pollution it is meant the accumulation of dangerous substances, such as heavy metals or poorly degradable chemicals such as dioxins, pesticides, and other synthetic compounds. The properties of the soil are altered by

C. Gambardella · R. Parente (✉)
Benecon University Consortium, Naples, Italy
e-mail: info@benecon.it

these substances, making it unsuitable for hosting plants and animals and causing dangerous situations for humans. Human activities are directly or indirectly responsible for the accumulation of pollutants in the soil [2]. In the search for the causes of environmental contamination, alongside the emissions which have always existed, and which are defined as natural, those deriving from anthropogenic causes must also be taken into consideration. The natural sources of pollutants being randomly distributed over the entire Earth's surface, usually do not produce acute and localized pollution phenomena. Pollutants of anthropogenic origin, even if often they are chemically the same as natural ones, are more dangerous for man. This is a consequence of the fact that even if released into the air in modest quantities, they derive from confined point sources and therefore heavily alter the air quality of areas with high population density, without substantially changing the normal composition of the major constituents [3].

The evolution of the urban fabric is changing the natural balance that allows us to live with the available resources. When building an artifact, material must be found by subtracting it from its natural location [4]. This leads to the proliferation of extractive activities that produce chasms in the territory: the quarries. The built environment grows hand in hand with the creation of quarries, which are located at a distance from urban settlements, generally in the vicinity of land intended for agriculture [5]. The impacts caused by extractive activities are important and the need to restore these areas is an essential requirement, with the aim of returning them to the territory from which they were stolen [6]. The quarries appear as refused places, realities extraneous to everyday life which are marginal and distant from the interests of society, and which are therefore abandoned, pouring into conditions of decay. These areas, in most cases, are located on the edges of urban realities, in natural areas, in agricultural territories, on mountain slopes or near lakes or natural waterways [7]. The quarries do not have the related green areas, which existed prior to the beginning of the quarry's cultivation, having been the subject of quarrying activities. This continuous subtraction of natural surfaces has caused serious damage to the ecosystem balances that characterize the natural territory.

Being in many cases of abandoned activities, the quarries are the object of illegal or even criminal activities. In some areas, abandoned quarries have often been used as a repository for industrial or even toxic waste [8]. In this way, criminal organizations can obtain profits from the non-disposal of waste, which, for some types, is extremely onerous. The identification of such illegal activities by the police has highlighted the need for monitoring of disused quarries. The term remote sensing defines the set of observation methods and techniques which, by extending the perceptive capacity of the human eye, provide qualitative and quantitative information on objects placed at a distance, through the measurement of electromagnetic radiation [9]. In nature, electromagnetic radiation is continuously distributed over a wide range of wavelengths, ranging from gamma rays to radio waves. Optical remote sensing sensors measure in the regions of the electromagnetic spectrum where reflected solar radiation prevails, such as from visible to near infrared. Just as the eyes passively perceive the radiation

emitted by the light source (natural or artificial) and refracted by objects, so a hyperspectral sensor is sensitive to solar radiation reflected from the underlying territory in segments of the light spectrum, even beyond the visible range [10, 11].

Remote sensing activities with the use of airborne sensors allow remote environmental monitoring without the need for direct access to the areas to be controlled which are often inaccessible for various reasons. In addition, in cases where toxic waste is present, these areas are dangerous if frequented by operators who must carry out measurements of the environmental parameters. In the case of remote sensing, operators keep a safe distance and run no risk.

For the activity covered by this work, a specific hyperspectral aerial remote sensing campaign was carried out with the Itres CASI 1500 sensor, suitable for recording the electromagnetic reflection of the territory in the segment between 350 and 1080 nm, such as from ultraviolet (UV), visible (VIS), and near infrared (NIR). The measurements were stored in a georeferenced image with 36 levels, one for each detected wavelength. The real geometric dimensions of the pixel are a function of the planning parameters and the flight performed, equal to $1 \times 1 \text{ m}^2$. The flight planning, the hyperspectral and photogrammetric acquisition, the processing of raw data, the analysis and critical interpretation of the processed data, were performed in compliance with the technical-operational protocol defined on scientific and experimental bases by the researchers and validated on remote sensing air missions conducted in national and international territory in collaboration with the police forces.

2 Materials and Methods

2.1 Investigation Area

The area covered by this study is a rural area close to densely populated urban settlements (Fig. 1).

It is a geographical area rich in landscape and urban history, the subject of intense anthropization and a very strong exploitation of one of its natural resources: the limestone rocks of the hills, a raw material used in private and public construction. The overall result, however, is that in an area of very few square kilometers, the landscape is characterized by huge white wounds in a context of increasingly bare hills, with chaotically man-made slopes. The impact of the intensive exploitation of this resource is, therefore, of a landscape, geological and public health type. In fact, the territory is geologically destabilized, the aquifers are more vulnerable, the material transport system and intense anthropization have led to a significant increase in related pollutions. Finally, the industrial activity related to cement factories has had a significant impact on the quantity and quality of fine dust emitted into the environment. The area under study is bounded by the Tifatini mountains and the Partenio chain. There are hilly ridges with variable slopes, densely wooded but also steep, subject to the risk of landslides especially after numerous forest fires. The



Fig. 1 Representation on orthophotos of the two sites of interest and of the area to be covered through remote sensing

north side of the Tifatini mountains recognizes hilly elements with gentler profiles, but which are, on average, spare that is, devoid of woodland except for the innermost part. The climate of the area is mild on average, typically Mediterranean, however rather heterogeneous between the different points of the valley. The further west areas have a more temperate climate than the more inland areas to the east. The area has important tuff deposits in the center, intensely exploited as quarries, rich in karst layers and presents the two hill chains mainly characterized by limestone deposits.

2.2 Aircraft and Sensor

The remote sensing activity was conducted using an aircraft capable of carrying up to 14 people, including two pilots and 12 passengers. In addition, due to the presence of a sufficiently powerful propulsion system, the aircraft can accommodate on board a load with a total weight of up to 1612 kg. The aircraft is also equipped with hatches for housing sensors. The Itres CASI 1500 sensor [12] acquires images with a linear resolution of 1500 pixels, which can be discretized into 288 channels ranging from the ultraviolet to the near infrared field. Thanks to high-quality optics and the use of latest generation charged-coupled device (CCD) sensors, the system can produce images that guarantee ground resolution up to 25 cm. The sensor platform is completed by

the inertial system (INS) integrated with the GPS antenna, useful for determining the position and attitude of the aircraft during remote sensing activities.

The CASI integrated system requires a geometric calibration activity to exactly calculate the orientation of the cameras with respect to the reference system of the inertial unit used to acquire the aircraft attitude information during image acquisition. The calibration was carried out by shooting on an area where there are support points of known coordinates. The Itres CASI-1500 sensor is of the passive type, that is, it measures the reflected solar radiation of surfaces. In this sense, it is not properly able to carry out sub-superficial measurements, and therefore does not allow a direct identification of the buried bodies. However, hyperspectral measurements lend themselves well to indirect evaluations where the bodies present in the subsoil cause spectrally relevant surface alterations, such as on vegetation [13–17].

However high the resolution of the acquired data is, the sensors record an aerial measurement in each pixel of the returned image. That is, they measure an average of the radiance coming from the elements contained in the single pixel (soil and vegetation above, gardens and street furniture, different materials within a quarry). In this sense, in the absence of homogeneous areas, each acquired element is to be understood as a mixed pixel, where extraneous information can disturb that of interest (Fig. 2).

The remote sensing data processing techniques applied to the carried out survey make it possible to extract information on the components of each pixel. An exemplary case is that of vegetation, where a spectral index allows to estimate the vegetative vigor, removing the disturbance of the underlying soil. However, the problem of identifying specific surface materials generally requires many ancillary data, including the spectral signatures of the elements that are expected to be observed in the image.

2.3 Survey Methodology

The architecture of the survey methodology used for this study is represented in the diagram of Fig. 3. The planning of the hyperspectral and photogrammetric remote sensing flight identified the overflight area and circumscribed on the military map base, considers the orography, the average altitude, and the exposure of the land under investigation. These parameters are functional to the correct definition of the hyperspectral and optical digital images to be acquired. The Itres CASI-1500 sensor supports different configurations, adaptable to the specific needs of the mission. The 36-channel configuration was adopted for this flight, which allows to maximize the number of spectral channels that can be acquired, compatibly with the conditions imposed by the geometric resolution of the pixel on the ground (one meter) and the integration times required for energy detection radiant. The selected bands are uniformly distributed over the entire available spectral range.

The definition of the flight parameters that is the configuration of the acquisition of the sensors, the flight altitude, and the speed of the aircraft during shooting, is followed by the cartography drawing of the scan lines, or the ordered series of run

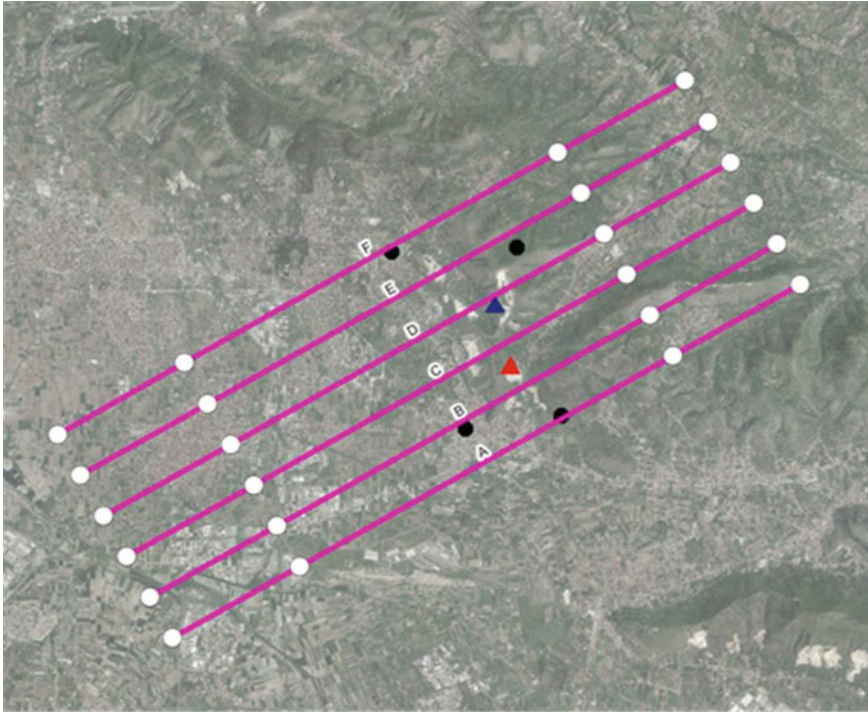


Fig. 2 Flight planning with indication of the flight lines, which, net of positioning and exit tracks, cover the entire area of interest. In the background, the orthophoto of the area under investigation

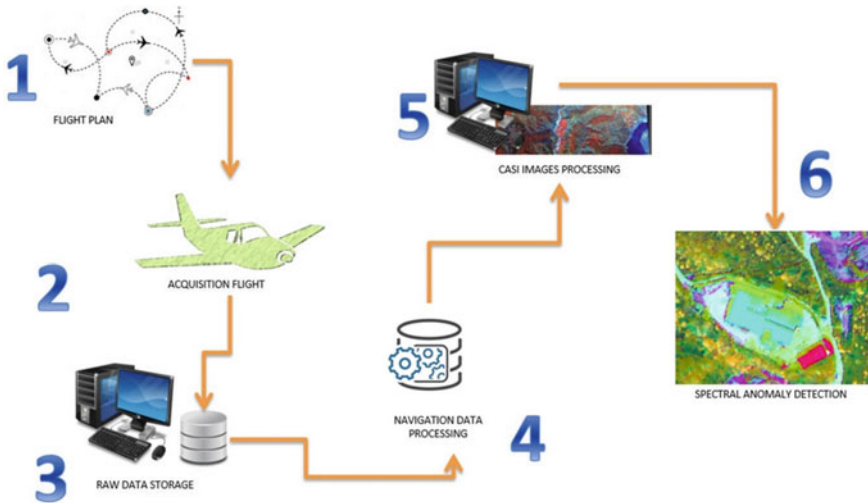


Fig. 3 Architecture of the survey methodology used for this study

lines (linear scan trajectory) to be flown with a nominal lateral overlap in plane of 20%; the overlapping area may vary with respect to the orography of the study area (Fig. 2).

The flight planning ends with the drafting of the time schedule of the acquisition campaign. This considers the technical characteristics of the sensors installed on board. In the case of flights with hyperspectral sensor, it is necessary to ensure the presence of the light source of maximum radiation. This is necessary to maximize the signal-to-noise ratio and reduce noise due to shadows.

The hyperspectral scans were suitably processed to produce:

- Images sampled in shades of red in relation to the presence of vegetation, for the representation of vegetated areas (shades from red); this false color representation associates to each red, green, and blue channel of the monitor the channels at, respectively, 789, 675, and 542 nm. This representation enhances the greater or lesser density of the vegetation with shades of red [18].
- Modified Soil Adjusted Vegetation Index (MSAVI) images, sampled with increasing gradations from white to black in relation to the greater or lesser chlorophyll activity of the vegetation, for the representation of plant areas with greater concentration of chlorophyll activity. This index is less sensitive than NDVI to soil variability. The MSAVI is related to the density of the vegetation cover which, compared to other spectral indices, has the advantage of not being influenced by the variability of the underlying soil [19].
- Principal Components Analysis images (PCA), realized through a statistical processing of the electromagnetic bands which maximizes the variance of the spectral data and filters the electromagnetic noise [20]. Principal component analysis is a technique used in the context of multivariate statistics for the simplification of source data. The primary purpose of this technique is the reduction of a high number of variables in some latent variables. This occurs through a linear transformation of the variables that projects the original ones into a new Cartesian system in which the variables are sorted in decreasing order of variance. Therefore, the variable with the greatest variance is projected to the first axis, the second to the second axis, and so on. The reduction in complexity occurs by simply analyzing the main (by variance) of the new variables [21].
- Reed–Xiaoli Detector (RXD) images, created through statistical processing for the identification of anomalies based on the statistical difference between a pixel and the surrounding area [22]. Considering the characteristics and dimensions of the area to be observed, a movable window of $21 \times 21 \text{ m}^2$ was used for the calculation. Reed et al. [reed] have developed an anomaly detection algorithm widely used in hyperspectral imaging. The RXD algorithm calculates the Mahalanobis distance between the currently detected data sample vector and the background data sample vectors, to exploit nonlinear hyperspectral data information.

The next section shows the results of the monitoring activity with the presentation of a series of maps developed with the application of the methodology just described.

3 Results and Discussion

The remote sensing data are digital images that represent the territory detected by means of a pixel matrix. Each pixel is associated with two variables x and y which identify its position within the image and the measurement of the average radiance of the ground area covered by the single pixel in a certain number of wavelengths [23]. A hyperspectral sensor measures the radiance in many contiguous channels, and therefore records, in each pixel, the spectral signature of the observed surface. A remotely sensed image is therefore a matrix that translates into numerical values the amount of energy emitted or reflected by the object under examination in each channel.

The measurements acquired by the hyperspectral sensor, as well as the aircraft attitude and in flight position data, are suitably stored according to organizational methods functional to the subsequent image preprocessing phase. This is aimed at correcting radiometric and geometric errors and georeferencing images through the following activities: differential correction of GPS data, processing of inertial data, radiometric calibration (correction) of images and ortho rectification [24].

The calculation of the position and attitude of the aircraft is performed through the integration of the data of the permanent GNSS stations (static acquisition) with the raw data of the inertial platform and on-board GNSS (dynamic acquisition). The former record satellite coverage over the large remote sensing area between the airport and the study area. The seconds are a combination of data (measured every second) between the aircraft's position and attitude (roll, pitch, drift), recorded from take-off to landing. The integrated processing of these data returns the georeferenced path made by the plane, so that all hyperspectral and thermographic images can be automatically referred to the scanned area.

Image preprocessing consists of two consecutive activities: radiometric correction and geometric correction [25]. In radiometric correction the raw data (digitized radiance measurements) are transformed into spectral radiance units. During geometric correction, the image (radiometrically corrected) is aligned with the aircraft's attitude and position data. In this phase, each across-line of the image (linear sequence of pixel-images orthogonal to the run line), acquired in a known instant of the flight (stored in the raw file through the connection to the GNSS antenna of the aircraft) is redesigned and projected in orthographic mode on the XY cartographic plane (with respect to a chosen datum) thanks to the integration with the Digital Elevation Model (DEM) of the territory under study. The geometric correction enables to make the acquired image geometrically congruent with the chosen reference, thus establishing a perfect correspondence between the position of the pixel in the image and its position on the territory.

Photogrammetry is an indirect survey technique that allows to obtain three-dimensional metric information on the shape and position of objects, through the acquisition and processing of photographic images. The basic theory of photogrammetry, represented by projective geometry, was developed long before the invention of the photographic technique itself. The photographic shots of the terrain must take

place in such a way that the entire area to be surveyed remains broken down into stereoscopic models so that each point of the terrain appears on at least two frames. This need is guaranteed by the longitudinal covering (overlap) and by the embracing of the frame (side of the plot of land contained in a frame). This produces two strips of overlap of at least 10% even in stereoscopic models compared to adjacent models which allows them to be chained. Photogrammetric flights for image acquisition must be carefully planned considering the various factors mentioned above.

Photogrammetry is therefore part of remote sensing techniques as the object is explored and analyzed at a given distance using electromagnetic radiation as an information vector. This technique, originally created to be applied in the architectural field, has found wide use in the topographic survey of the territory, developing in this sense as aerial photogrammetry. With the advent of digital technology, both photographic acquisitions and image processing have been entirely developed in a computer environment thanks to the development of dedicated software. Given a block of frames—acquired according to a specific flight plan for aerial photogrammetry as mentioned above—in the software environment we proceed to the preliminary phase of image mosaicing which consists in the relative calculation of the various frames during aerial acquisition or definition of the visual pyramid implemented by the photographic camera at every instant of shooting. The stereo-photogrammetric model is therefore scaled and oriented to the reality of the places through the identification of a suitable number of points on the ground (Ground Control Point, GCP) of which the absolute coordinates in the geodetic reference system of the whole project are known. From this stereo-photogrammetric model, the software proceeds to the three-dimensional calculation of the photographed surfaces to produce a Digital Surface Model (DSM), which represents the morphological configuration of the area subject to aerial photogrammetry within the limits of the resolution imposed on the calculation and derived in a semi-automatic by software algorithms. The DSM of the site of interest reached a resolution of 0.61 m. Figure 4 shows a 3D representation.

The curve that describes as a function of the wavelength the ratio between the incident solar radiation and the reflected one measured by the sensor, is called the spectral signature of a surface. It is linked to the biological, chemical, and physical characteristics of the surface, as well as the geometry of illumination and observation, and allows its identification and characterization. The spectral signature of the object to identify is not always available, and when it is necessary to highlight the characteristics of the vegetation, specific indices are used. Vegetation indices are generally dimensionless measures derived from radiometric data to highlight the presence of vegetation in an image. Most of these indices are based on the rapid increase in reflectance that occurs at approximately 700 nm. This rapid increase is in fact typical of green vegetation; all the other covers show a gradual variation in the region of the spectrum around the same wavelength.

By representing the reflectance values in the infrared region compared to those in the red region, for all the pixels of a sample area, it is possible to obtain a distribution where the points of dense vegetation are highlighted due to the high reflectance values in the Near-Infrared Region (NIR) and the corresponding values in red. The points of uncovered soil, in relation to their specific condition, are distributed along the lower

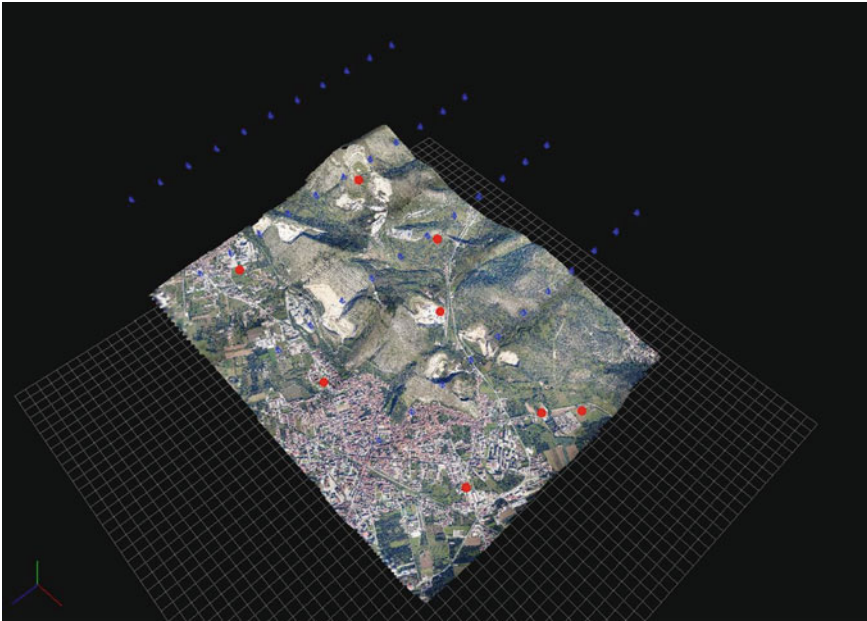


Fig. 4 Georeferenced Digital Surface Model (DSM) model of the area being monitored through remote sensing

right diagonal. This line results as a linear approximation between the reflectance values in the NIR and in the red, it is called the soil line and shows dry soils with high reflectance values in both bands, while wet soils are characterized by low reflectance values in both bands.

Vegetation indices are particularly useful for detecting changes in land use, for assessing the density of vegetation cover, and for crop discrimination. In fact, the different cultures can be identified by detecting the spectral differences with respect to the other elements. The visible radiation in the red (630–690 nm) is absorbed by the chlorophyll while the radiation in the near infrared (760–900 nm) is strongly reflected by the cell structures foliar.

In examining the general vegetation reflectance curve, the observed deviation between red and near infrared constitutes a sensitive variable to the presence of green vegetation. The spectral response of the red vegetation is strongly correlated to the concentration of chlorophyll while the spectral response in the near infrared is controlled by the leaf area index and the density of the green vegetation. In this study, a vegetative index was adopted, an index which reprocesses the hyperspectral remotely sensed images in shades of red in relation to the presence of vegetation for the representation of the vegetated areas (shades from red). This false color representation maps each of the monitor's red, green, and blue channels to channels at 789, 675, and 542 nm, respectively. In this way, the greater or lesser density of the vegetation is enhanced with shades of red (Fig. 5).

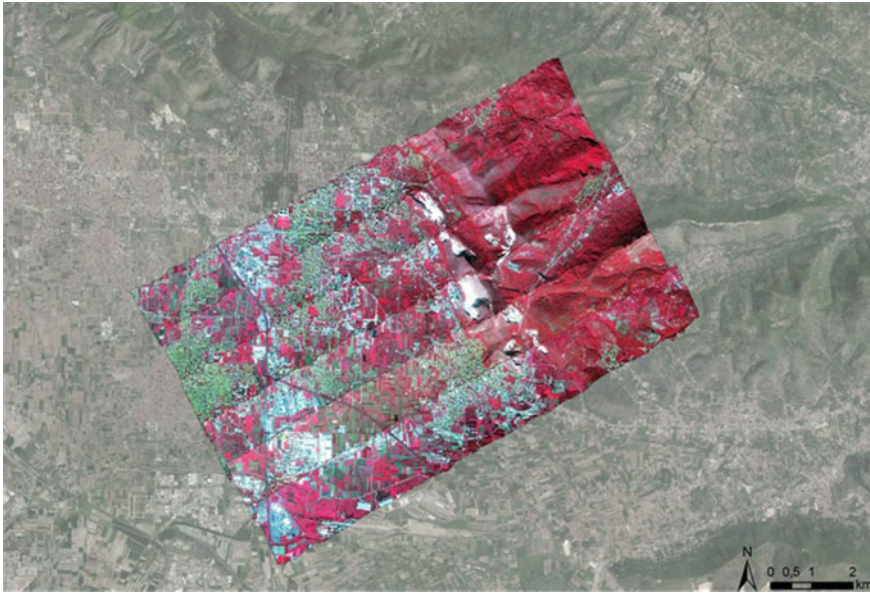


Fig. 5 False color representation of the CASI hyperspectral scan acquired in flight

Another index widely used by the scientific community is the Modified Soil Adjusted Vegetation Index (MSAVI), which samples the remotely sensed images with increasing gradations from white to black in relation to the greater or lesser chlorophyll activity of the vegetation [26]. In this way, the plant areas with the highest concentration of chlorophyll activity are highlighted. The MSAVI is correlated to the density of the vegetation cover which, compared to other spectral indices, has the advantage of not being influenced by the variability of the underlying soil (Fig. 6).

The analysis of a vast territory with different spectral signatures poses the problem of reducing the dimensionality of the features. In this context, the Principal Components Analysis (PCA) is applied. In the case of an $n \times p$ data matrix we can think of n points in a p -dimensional space. If p is high, the goal is to reduce the representation of the points to a space of smaller dimensions, but which maintains the structure of the original points in the best possible way [27, 28]. The new dimensions are identified by the main components: if the reduced space is two-dimensional, there will be two main components; if three-dimensional, three components, and so on. In other words, the PCA technique looks for a first linear combination with maximum possible variance, given some constraints. Then it looks for a second linear combination that maximizes the variability given the constraints and the first component, and so on. In this way the differences between the objects are highlighted to identify spectral anomalies (Fig. 7).

From all these elaborations we can appreciate the environmental remediation works on the northwest side of the quarry, with terraces covered with vegetation. Leaning at the base of the west flank and in a central position, there is a vegetated

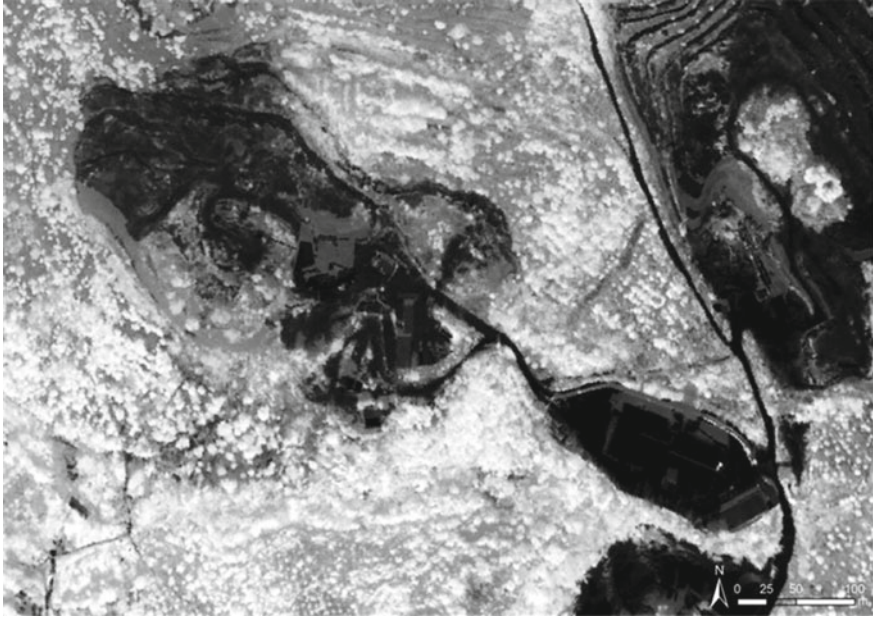


Fig. 6 MSAVI vegetation index calculated from data collected on the site of interest with a CASI sensor

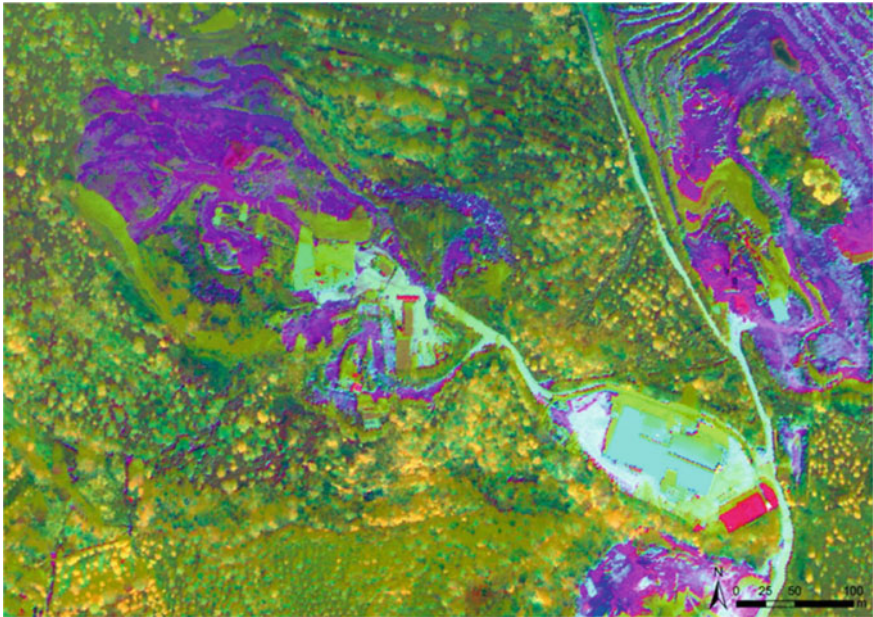


Fig. 7 PCA processing on the site of interest. With the three colors red, green, and blue, components 5, 3, and 1 are, respectively, represented

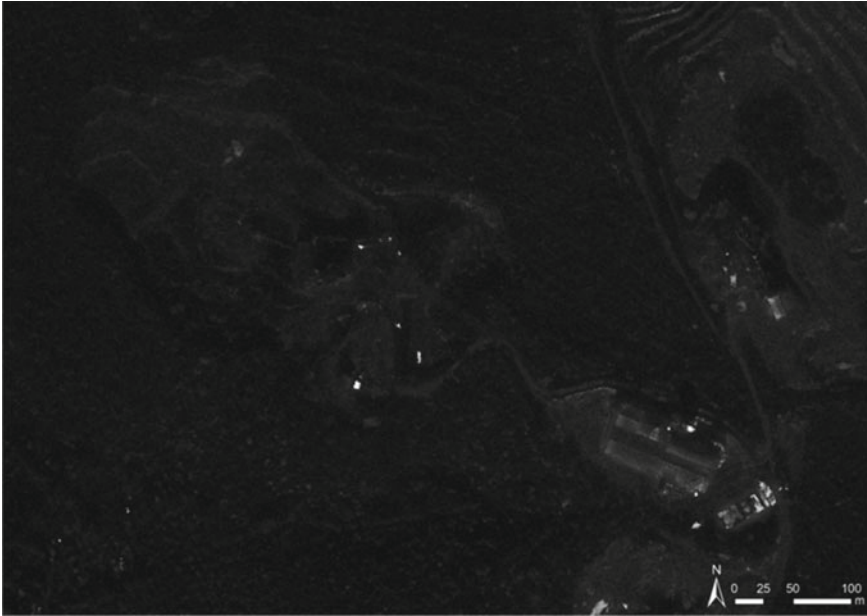


Fig. 8 RXD analysis for the identification of spectral anomalies in the site of interest

area inside the quarry, extending about 40 m from east to west and 55 m from north to south. The area appears to be higher than the area immediately to the south, where there is an accumulation of water, probably due to the rains of the previous days. The analysis of the main components (Fig. 7) does not highlight any further areas to be considered. The RXD image (Fig. 8) shows a limited number of spectral anomalies above the 99.9 percentile threshold [29–32]. The found anomalies do not seem justifiable as normal objects found in a quarry (sheds, moving machines, work tools). However, a group of anomalies corresponding to the accumulation of regular blocks present in the center of the quarry must be highlighted; the blocks appear visually identical. They present a certain spectral heterogeneity which then gave rise to the anomalous value. Taken together, the blocks cover an area of $25 \times 50 \text{ m}^2$.

4 Conclusions

In this work, images detected in a specific hyperspectral aerial remote sensing campaign with Itres CASI 1500 sensor, suitable for recording the electromagnetic reflection of the territory in the segment between 350 and 1080 nm, such as, from ultraviolet (UV), to visible (VIS), and near infrared (NIR). The measurements were stored in a georeferenced image with 36 levels, one for each detected wavelength. The hyperspectral images were post-processed using vegetation indices (Red-Veg

and MSAVI), PCA and RXD algorithms. The survey methodology made it possible to detect a mound covered with vegetation at the base of the western flank of the quarry. The mound extends approximately 40 m from east to west and 55 m from north to south. From the photogrammetric survey it can be inferred that the heap has a height of about 4 m with respect to the ground level. The visual absence of this accumulation in the orthophotos dating back to previous periods is noted. Furthermore, a set of blocks of regular shape was found in the center of the quarry. This mass has a quadrangular shape with dimensions of approximately $25 \times 50 \text{ m}^2$, and a height of approximately 2 m with respect to the ground level. A part of the blocks has a high value of the RXD spectral anomaly indicator that is, the set of blocks has a fair degree of spectral heterogeneity within it and with respect to the context. The reported observations are based on a remote aerial survey in the reflective optical domain, from the visible to the near infrared. The highlighted points of attention are subject to further investigation, by means of georadar and/or magnetometric scanning, and core drilling.

References

1. Rodríguez-Eugenio N, McLaughlin M, Pennock D (2018) Soil pollution: a hidden reality. *FAO*
2. Mishra RK, Mohammad N, Roychoudhury N (2016) Soil pollution: causes, effects and control. *Van Sangyan* 3(1):1–14
3. Fayiga AO, Saha UK (2016) Soil pollution at outdoor shooting ranges: health effects, bioavailability and best management practices. *Environ Pollut* 216:135–145
4. Ciaburro G (2021) Recycled materials for sound absorbing applications. *Mater Sci Forum* 1034:169–175
5. Milgrom T (2008) Environmental aspects of rehabilitating abandoned quarries: Israel as a case study. *Landsc Urban Plan* 87(3):172–179
6. Tsolaki-Fiaka S, Bathrellos GD, Skilodimou HD (2018) Multi-criteria decision analysis for an abandoned quarry in the Evros Region (NE Greece). *Land* 7(2):43
7. Ren X, Cai T, Wang X (2010) Effects of vegetation restoration models on soil nutrients in an abandoned quarry. *J Beijing For Univ* 32(4):151–154
8. Kivell PT (2021) Land reclamation through waste disposal. In: *Waste location: spatial aspects of waste management, hazards and disposal*, p 12
9. Gambardella C, Parente R, Ciabrone A, Casbarra M (2021) A principal components analysis-based method for the detection of cannabis plants using representation data by remote sensing. *Data* 6(10):108
10. Campbell JB, Wynne RH (2011) *Introduction to remote sensing*. Guilford Press
11. Weiss M, Jacob F, Duveiller G (2020) Remote sensing for agricultural applications: a meta-review. *Remote Sens Environ* 236:111402
12. Itres Research Limited homepage. <https://itres.com/>. Accessed 2022/05/10
13. Acheroy M (2007) Mine action: status of sensor technology for close-in and remote detection of anti-personnel mines. *Near Surf Geophys* 5(1):43–55. <https://doi.org/10.3997/1873-0604.2006017>
14. Cavalli RM, Colosi F, Palombo A, Pignatti S, Poscolieri M (2007) Remote hyperspectral imagery as a support to archaeological prospection. *J Cult Herit* 8(3):272–283. <https://doi.org/10.1016/j.culher.2007.03.003>
15. Maathuis BHP, van Genderen JL (2004) A review of satellite and airborne sensors for remote sensing based detection of minefields and landmines. *Int J Remote Sens* 25(23):5201–5245. <https://doi.org/10.1080/01431160412331270803>

16. Robledo L, Carrasco M, Mery D (2009) A survey of land mine detection technology. *Int J Remote Sens* 30(9):2399–2410. <https://doi.org/10.1080/01431160802549435>
17. Rowlands A, Sarris A (2007) Detection of exposed and subsurface archaeological remains using multi-sensor remote sensing. *J Archaeol Sci* 34(5):795–803. <https://doi.org/10.1016/j.jas.2006.06.018>
18. Motohka T, Nasahara KN, Oguma H, Tsuchida S (2010) Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens* 2(10):2369–2387
19. Qi J, Chehbouni A, Huete AR, Kerr YH, Sorooshian S (1994) A modified soil adjusted vegetation index. *Remote Sens Environ* 48(2):119–126
20. Ready P, Wintz P (1973) Information extraction, SNR improvement, and data compression in multispectral imagery. *IEEE Trans Commun* 21(10):1123–1131
21. Ciaburro G (2022) Time series data analysis using deep learning methods for smart cities monitoring. In: *Big data intelligence for smart applications*. Springer, Cham, pp 93–116
22. Reed IS, Yu X (1990) Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans Signal Process* 38:1760–1770
23. Ciaburro G (2021) Security systems for smart cities based on acoustic sensors and machine learning applications. In: *Machine intelligence and data analytics for sustainable future smart cities*. Springer, Cham, pp 369–393
24. Shen L, Stopher PR (2014) Review of GPS travel survey and GPS data-processing methods. *Transp Rev* 34(3):316–334
25. Kobayashi S, Sanga-Ngoie K (2008) The integrated radiometric correction of optical remote sensing imageries. *Int J Remote Sens* 29(20):5957–5985
26. Gilabert MA, González-Piqueras J, García-Haro FJ, Meliá J (2002) A generalized soil-adjusted vegetation index. *Remote Sens Environ* 82(2–3):303–310
27. Ciaburro G (2021) An ensemble classifier approach for thyroid disease diagnosis using the AdaBoostM algorithm. In: *Machine learning, big data, and IoT for medical informatics*. Academic Press, pp 365–387
28. Ciaburro G (2020) Sound event detection in underground parking garage using convolutional neural network. *Big Data Cogn Comput* 4(3):20
29. Imani M (2017) RX anomaly detector with rectified background. *IEEE Geosci Remote Sens Lett* 14(8):1313–1317
30. Mehmood A, Nasrabadi NM (2011) Kernel wavelet-Reed–Xiaoli: an anomaly detection for forward-looking infrared imagery. *Appl Opt* 50(17):2744–2751
31. Ranganathan G (2020) Real time anomaly detection techniques using PySpark frame work. *J Artif Intell* 2(01):20–30
32. Shakya S, Pulchowk LN, Smys S (2020) Anomalies detection in fog computing architectures using deep learning. *J Trends Comput Sci Smart Technol* 1(2020):46–55

A Naïve Bayes Approach for Predicting the Skin Allergy Diseases



B. Geluvaraj, K. Santhosh, T. Sandhya, V. Akshay Reddy, S. V. Bhaskar, and N. Sasikala

Abstract Machine learning (ML) plays a significant part in data mining, image identification, natural language processing, and disease diagnosis in medical industry. ML provides precise outputs in many fields discussed above (Geetha and Arunachalam in Evaluation based approaches for liver disease prediction using machine learning algorithms. Dr. MGR University, Chennai, India, 2021 [1]). This research study has analyzed various diseases such as impetigo, fungal infections, allergy, and other diseases. For analyzing the data, ML techniques are used to enhance over time and work efficiently but models require error-free data (Saboji and Ramesh in A scalable solution for heart disease using classification mining technique. CMR Institute of Technology, India, 2017 [2]). The researches have previously utilized different models to solve different challenges in the field of healthcare and diagnosis. The maximum accuracy is shown by naïve Bayes (Sneith Raja and Anurag in Machine learning based heart disease prediction system. Vardhaman College of Engineering, India, 2021 [3]). The algorithm that can compete with various precise models is random forest as it delivers a highly error-free prediction. Therefore, the proposed application requires highly precise results and it does not need a human readable model, which can make use of random forest (RF) algorithm (Chowdhury and Ahmed in Heart disease prognosis using machine learning classification techniques. Metropolitan University, Bangladesh, 2021 [4]). The decision tree algorithm delivers the second highest accuracy rate. This algorithm is convenient and operate fast forecastable class of test data.

Keywords Machine learning (ML) · K-nearest neighbor (KNN) · K-means · Logistic regression (LR) · Decision tress classifier (DSC) · Random forest classifier (RFC) · Naïve Bayes (NB) · Convolutional neural network (CNN) · Recurrent neural networks (RNN) · Support vector machine (SVM) · Artificial neural network (ANN)

B. Geluvaraj (✉) · K. Santhosh · T. Sandhya · V. Akshay Reddy · S. V. Bhaskar · N. Sasikala
Department of Computer Science and Engineering, New Horizon College of Engineering,
Bangalore, India
e-mail: geluvaraj999@gmail.com

1 Introduction

Machine learning focuses on recognizing different patterns, structures, and data analysis models in order to leverage a valid learning, decision-making, and interpretation other than human interactions. ML techniques used for data analysis improves over time and work more efficiently, but these models require error-free data. This system allows users to obtain instant predictions on health problems online via an automated design [5]. This methodology is applied to a wide range of symptoms and diseases. Diseases such as typhoid, cancer, heart disease, tuberculosis, and others are fatal if not treated [6]. The medical industry employs ML techniques to obtain accurate forecasting and decision-making. These techniques have a huge architecture because they acquire the patterns and relationships that are hidden in the architecture [7]. This system employs ML techniques which can be correlated with patient symptoms to precisely estimate the diseases. ML algorithms such as KNN, K-means, DSC, and RFC are used to predict the corrective diseases [8]. The dataset used in this model includes 95 symptoms (to avoid overfitting) and 41 diseases based on combinations or permutations. As a result, the proposed model aims to provide accurate results based on the symptoms [9].

2 Related Work

The most important techniques are the classification techniques, which are convenient and accurate ML algorithms applied in disease prediction [10]. For further research studies and analysis, various ML algorithms are used to interpret heart diseases, diabetes, etc. [11]. DSC, NB, and KNN classification models are utilized to diagnose the diseases correlated to the symptoms fed by the users [12]. The practitioners and researchers have previously utilized these models in the healthcare and diagnosis fields. ML algorithms such as DSC, NB, RFC, and KNN were experimented by using the hybrid intelligence out of which NB has provided a better output [13]. This experiment on a dataset delivered a model and the results revealed that the integrated hybrid and intelligent technique has improved the prediction accuracy. Analysis of reviewed paper is shown in Table 1.

3 Algorithm

3.1 Decision Tree Classifier

In reality, a tree has a large number of analogs, which would result in affecting a wide range of ML algorithms, which include differentiation and regression as shown in Fig. 1. DST is used for evaluating results in order to depict and take decisions

Table 1 Analysis of reviewed papers

Algorithms	Accuracy (%)	Year of published
DSC	93.29	2020
RFC	93.29	
NB	93.61	
LR	73.47	2018
RFC	92.67	
RNN	92.17	
KNN	99	2019
CNN	97	
NB	86.41	2021
SVM	79.67	
RFC	71.87	
ANN	90.14	2018
NB	95	
DSC	96.17	
ANN	73.23	2018
LR	76.13	
DSC	77.87	
DSC	98.18	2019
RFC	98.05	
NB	98.55	
Logistic model tree	61.38	2020
J48 DSC	44.55	
ANN	97.5	2020
NB	87.19	
DSC	91.4	
DSC	94.29	2020
NB	91.29	
Neural network	92.61	
LR	73.47	2018
RFC	92.67	
RNN	92.17	
RFC with linear model	88.7	2019
J48 DSC	77.5	2017
NB	83.49	
DSC	74	2019
SVM	82	
NB	80	
LR	73.23	2021
SVM	75.04	
DSC	97.5	2021
RFC	99.6	
DSC	82.45	2021
KNN	81.1	
NB	84	

(continued)

Table 1 (continued)

Algorithms	Accuracy (%)	Year of published
RFC	88	2017
NB	44	
DSC	77	2017
KNN	91	
SVM	77	2017
RFC	72	
NB	47	
CNN	83	2019
SVM	89	2021
LR	90	
KNN	55	2017
NB	96.2	2014
SVM	96.97	
RFC	96.63	
DSC	85.1	2020
NB	90.4	

evidently [14]. It has a tree structure, which has nodes and corresponding decisions. It is a most used technique in ML to fetch few techniques to accomplish specific goals [15].

- Decision node: It has various pairs of edges. In this model, the symptoms are described as decision nodes.
- Leaf node: Leaf nodes depicts a branch decision. In this model, diseases are described as leaf nodes [16].

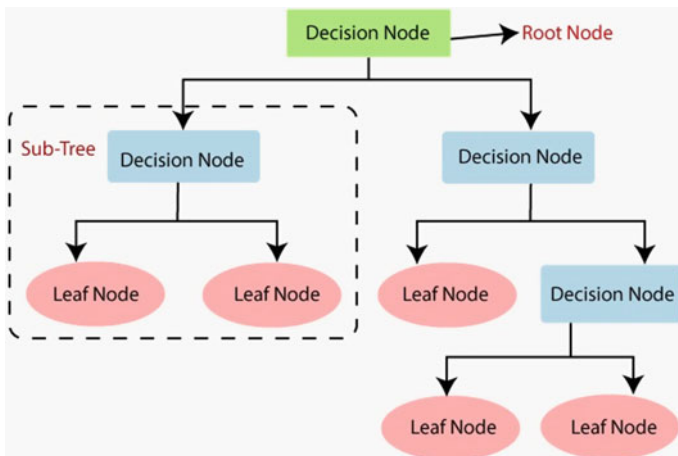


Fig. 1 Decision tree classifier

Random Forest Classifier

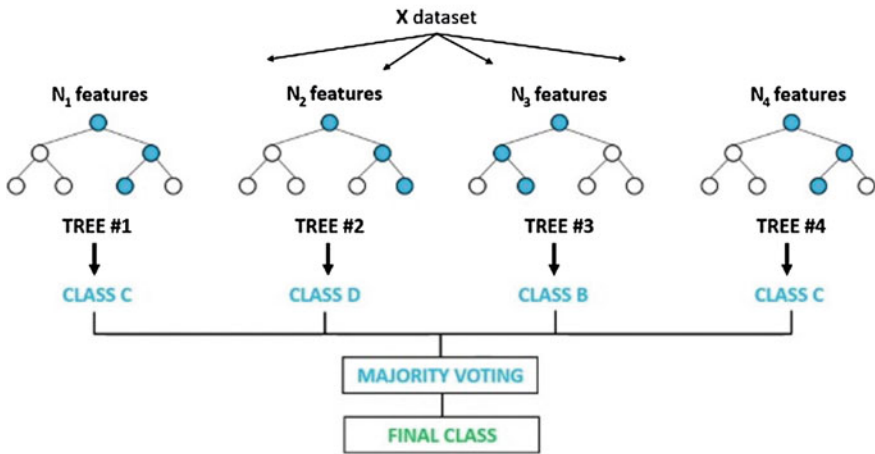


Fig. 2 Random forest classifier

3.2 *Random Forest*

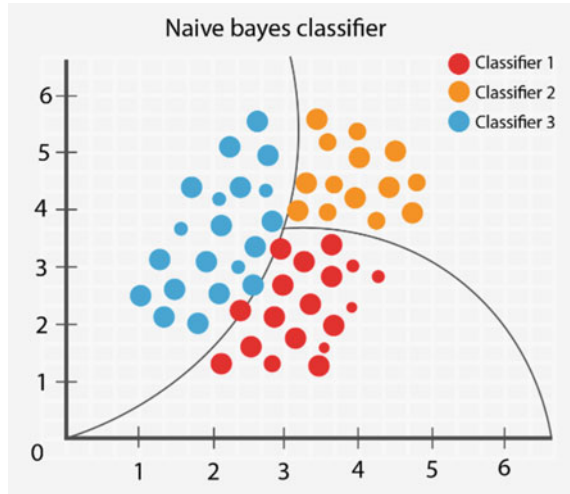
RFC is a conventional machine learning algorithm which remains as a part of supervised learning. It is used to solve complex challenges with the differentiation and regression techniques in ML [17]. It is more focused on group modeling methodology, where various classifiers are consolidated to direct a specific complication and improve the effectiveness of the schema.

RFC is a classification technique that includes a group of decision trees and to improve the predictive efficiency of the datasets on various subgroups of the provided dataset, the averages are considered [18] as shown in Fig. 2. RFC obtains the prediction from all trees and is based on the majority voting of the prediction rather than depending on one decision tree, which forecasts the finishing outcomes.

3.3 *Naïve Bayes*

Based on the Bayes theorem, NB is a supervised learning technique [19], wherein the categorization issues can be solved by using this technique. It is significantly used in classification of text, for which high dimensional training datasets are required [20]. The NB being the easiest and most dominant classification technique [14]. This technique is utilized to produce quick models of ML which can produce speedy forecasting. NB is a specific classifier which indicates on the likelihood object, it

Fig. 3 Naïve Bayes



forecasts [16] shown in Fig. 3. Credit scoring, classification of medical data, and classification of papers are some general instances of the NB algorithm.

4 Mathematical Construction

4.1 Decision Tree

Entropy: The amount of randomness and measures the unpredictability [15]. Further, in other words, we can say that entropy describe forecastability of an event particularly. In order to construct the decision tree, two kinds of entropy must be calculated by making use of each attribute of the frequency table as given below:

- $E(C)$: Using the frequency table this is one of the attributes, C is present state (current results) [21].
- $P(h)$: It is a probability of the state C of an event h .

$$E(C) = \sum_{h \in H} -P(h) \log_2 P(h) \tag{1}$$

- $E(C, A)$: C and A are the two attributes using the frequency table, C is the present state with the attribute named A .
- $P(h)$: It is a probability of attribute named A of an event h .

$$E(C, A) = \sum_{h \in H} [p(h) * E(C)] \tag{2}$$

$E(C)$ is the whole set entropy, on the other hand the next term is $E(C, A)$ correlated to attribute named A .

Information gain: It is described by $IG(C, A)$ for C state reduction in entropy post finalizing the attribute named A . It is measured by comparing entropy before and after dataset transformation, as follows [22].

$$IG(C, A) = E(C) - E(C, A) \quad (3)$$

4.2 Naïve Bayes

NB technique is formed on Bayes theorem as follows:

$$P(s/h) = (P(h/s)P(s))/P(h) \quad (4)$$

where,

$P(s/h)$ Posterior probability

$P(h/s)$ Likelihood

$P(s)$ Class prior probability

$P(h)$ Predictor Prior probability

From the terms mentioned above 's' corresponds to class and 'h' corresponds to features [23]. The denominator $P(h)$ is the only term which is the function of features. This term is not a function of the class which is presently being dealt with it [24]. Therefore, features will be same for every classes. In NBC, this denominator is ignored because it does not make any changes to the outcome of the classifier which does the prediction [25].

$$P(s/h) \propto P(h/s)P(s) \quad (5)$$

5 Datasets

To train the model for processing generous actions, the training dataset is utilized. To train the model, detailed characteristics are extracted from the training set [26]. These models are then combined to form the prototype. The test data is a dataset that is independent of the training dataset and has the same probability distribution as the training dataset [27]. The application is based on the medical records of 4920 patients, who are susceptible to 41 diseases caused by a combination of symptoms. In this application, 95 symptoms were considered [28].

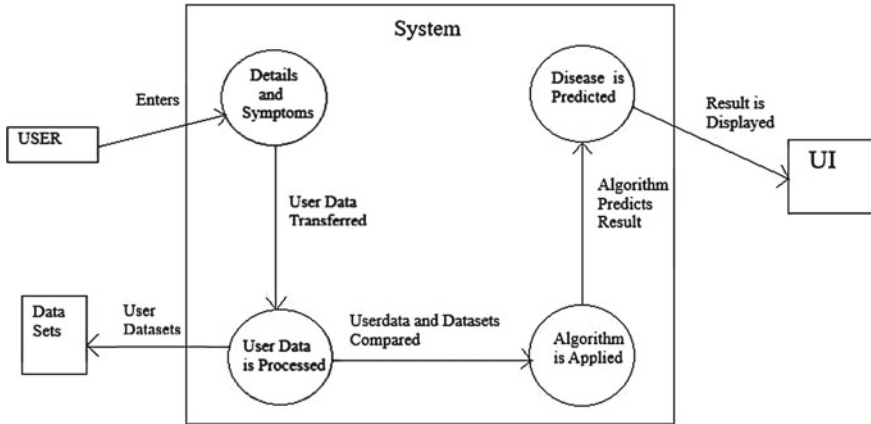


Fig. 4 Block diagram for patient login

6 Experimental Setup

- The system consists of a login page using which the user/patient can login with their credentials.
- Once the credentials match, the system redirects the user to disease forecasting webpage where the user will be asked to specify few symptoms.
- If the credentials do not match, then system will display an error message that credentials does not match.
- Once the user specifies the symptoms, the system shows the predicted disease and accuracy of the algorithm used. Refer Fig. 4 represents the workflow of the system.

7 Result

Figure 5 depicts the accuracy rate of allergy predicted by the system based on the symptoms provided. NB has shown 98% accuracy rate, which is the highest accuracy rate among the other algorithms. RFC and DSC have shown 86% accuracy. Naïve Bayes is clearly the best classifier based on this result.

Figure 6 depicts the accuracy rate of fungal infection predicted by the system based on the symptoms provided. NB has shown 92% accuracy rate, which is the highest accuracy rate among the other algorithms. DSC has shown 90% and RFC has shown 77% accuracy. Again, naïve Bayes is clearly found as the best classifier based on this result.

Fig. 5 Accuracy rate of allergy

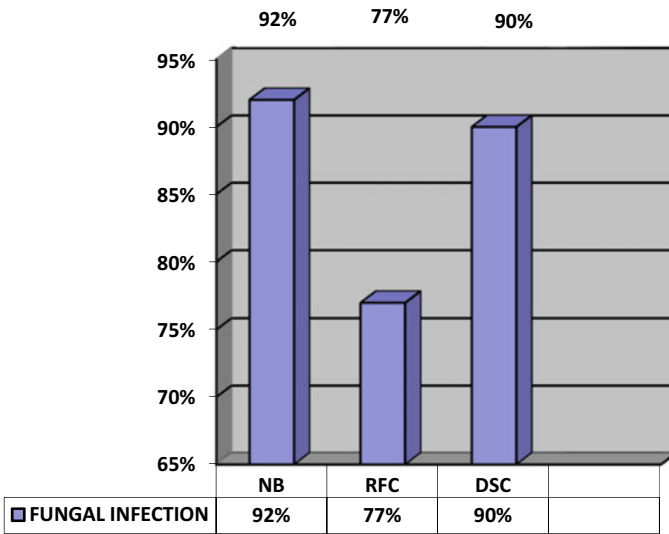
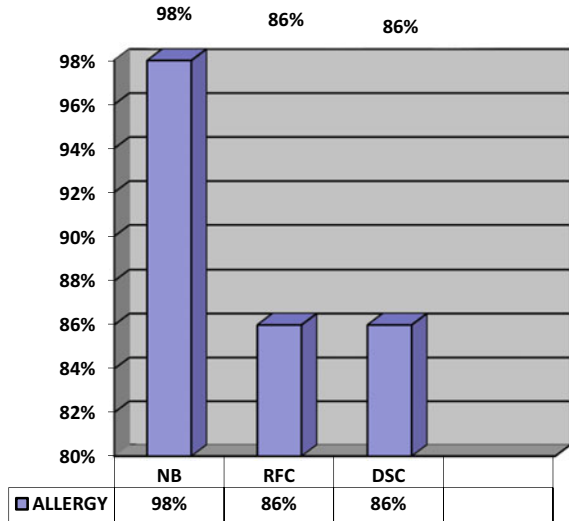


Fig. 6 Accuracy rate of fungal infection

8 Conclusion

This research work has successfully developed a novel disease forecasting system by using the ML technique [29]. With the goal of using existing data to predict precise disease based on symptoms and the recent increase in medical data, the RFC, DSC, and NB algorithms are used to classify diseases based on the symptoms

entered by the end user [30]. Machine learning plays a significant role in the medical industry especially in the process of disease diagnosis. In many of the fields discussed above, ML produces precise results. In all of these areas, ML produces accurate results [31]. This research paper characterizes various ML techniques for processing and predicting many diseases such as the heart diseases and diabetes. Most of the structures deliver excellent results since they precisely describe the characteristics. According to the results provided by these algorithms, DST and NB had the highest accuracy rate [32].

References

1. Geetha C, Arunachalam AR (2021) Evaluation based approaches for liver disease prediction using machine learning algorithms. Dr. MGR University, Chennai, India
2. Saboji RG, Ramesh PK (2017) A scalable solution for heart disease using classification mining technique. CMR Institute of Technology, India
3. Sneeth Raja M, Anurag M (2021) Machine learning based heart disease prediction system. Vardhaman College of Engineering, India
4. Chowdhury MNR, Ahmed E (2021) Heart disease prognosis using machine learning classification techniques. Metropolitan University, Bangladesh
5. Sharma V, Yadav S (2020) Heart disease prediction using machine learning techniques. Banaras Hindu University, India
6. Geluvaraj B, Santhosh K, Akshay Reddy V, Sandhya T, Bhaskar SV (2022) A hybrid approach for predicting diseases using clustering and classification techniques. In: 2022 international conference on advances in computing, communication and applied informatics—ACCAI
7. Grampurohit S, Sagarnal C (2020) Disease prediction using machine learning algorithms. K.L.E Institute of Technology Hubli, India
8. Hashi EK, Uz Zaman MS (2017) An expert clinical decision support system to predict disease using classification techniques. Rajshahi University of Engineering & Technology, Bangladesh
9. Wibamanto W, Das D, Chelliah S (2020) Smart health prediction system with data mining. Asia Pacific University of Technology & Innovation, Malaysia
10. Gavhane A, Kokkula G, Pandya I, Devadkar K (2018) Prediction of heart disease using machine learning. Sardar Patel Institute of Technology, Mumbai, India
11. Kunjir A, Sawant H (2017) Data mining and visualization for prediction of multiple diseases in healthcare. Modern Education Society's College of Engineering, Pune
12. Christensen T, Frandsen A, Glazier S, Humpherys J, Kartchner D (2018) Machine learning methods for disease prediction with claims data. Brigham Young University, Provo, Utah
13. Dahiwade D, Patle G, Meshram E (2019) Designing disease prediction model using machine learning approach. Abha-Gaikwad Patil College of Engineering, Nagpur, Maharashtra, India
14. Hamsagayathri P, Vigneshwaran S (2021) Symptoms based disease prediction using machine learning techniques. Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamil Nadu, India
15. Vijaya Shetty S, Karthik GA, Ashwin M (2019) Symptom based health prediction using data mining. Nitte Meenakshi Institute of Technology, Bangalore
16. Deepika M, Kalaiselvi K (2018) A empirical study on disease diagnosis using data mining techniques. Vels Institute of Science Technology and Advanced Studies, Chennai, India
17. Mohan S, Thirumalai C, Srivastava G (2019) Effective heart disease prediction using hybrid machine learning techniques. School of Information Technology and Engineering, VIT University, Vellore, India
18. Sonar P, JayaMalini K (2019) Diabetes prediction using different machine learning approaches. Mumbai University, Mumbai, India

19. Mohapatra S, Pati B, Patra PK, Mohanty S (2018) Smart health care system using data mining. College of Engineering and Technology, Bhubaneswar, India
20. Johari NAAM, Mohamad N, Isa N (2020) Smart self-checkup for early disease prediction. University Teknologi MARA Cawangan Terengganu, Kampus Kuala Terengganu, Malaysia
21. Ferdous M, Debnath J (2020) Machine learning algorithms in healthcare: a literature survey. Daffodil International University, Dhaka, Bangladesh
22. Geluvaraj B, Sundaram M (2021) A matrix factorization technique using parameter tuning of singular value decomposition for recommender systems. *Turk J Comput Math Educ* 12(2):3313–3319
23. Geluvaraj B, Satwik PM, Ashok Kumar TA (2019) The future of cybersecurity: major role of artificial intelligence, machine learning, and deep learning in cyberspace. In: Smys S, Bestak R, Chen JZ, Kotuliak I (eds) International conference on computer networks and communication technologies. Lecture notes on data engineering and communications technologies, vol 15. Springer, Singapore
24. Geluvaraj B, Sundaram M (2021) An implementation and combining of hybrid and content based and collaborative filtering algorithms for the higher performance of recommended systems. In: Chaubey N, Parikh S, Amin K (eds) Computing science, communication and security. COMS2 2021. Communications in computer and information science, vol 1416. Springer
25. Geluvaraj B, Sundaram M. A novel hybrid approach for machine learning based recommended systems. *Int J Grid Distrib Comput*
26. Geluvaraj B, Sundaram M (2022) A hybrid approach to resolve data sparsity and cold start hassle in recommender systems. In: Ranganathan G, Bestak R, Palanisamy R, Rocha Á (eds) Pervasive computing and social networking. Lecture notes in networks and systems, vol 317. Springer, Singapore
27. Suresh A, Suresh G (2017) Prediction of cardiac arrhythmia type using clustering and regression approach. Amrita School of Engineering, India
28. Sundaram M, Geluvaraj B (2020) An implementation and combining of hybrid and content based and collaborative filtering algorithms for the higher performance of recommended systems, May 2020
29. Perna D, Tagarelli A (2019) Deep auscultation: predicting respiratory anomalies and diseases via recurrent neural networks. University of Calabria, Italy
30. Paul R, Sayed A (2017) Clustering medical data to predict the likelihood of diseases. Bangladesh University of Engineering and Technology, Bangladesh
31. Sumana BV, Santhanam T (2014) Prediction of diseases by cascading clustering and classification. Department of Computer Science, Vijaya College, Jayanagar, India
32. Khan P, Kader MF (2021) Machine learning and deep learning approaches for brain disease diagnosis. University of Chittagong, Bangladesh

Internet of Things with Deep Learning Driven Disease Detection and Classification Model



C. Nithyeswari and G. Karthikeyan

Abstract Currently, Internet of Things (IoT) is extensively adopted in number of applications that its prominence is extending in day-to-day lives. Also, it is being utilized as E-health application on divergent characteristics like earlier diagnosis of medical issues, computer-assisted rehabilitation, and emergency notification. Due to misdiagnoses of disease that increased recently in a scarily manner, many research workers dedicated their effort and deployed technology to enhance the medical diagnoses process and reduce the resulted risk. This paper majorly concentrates on the design of artificial fish swarm algorithm (AFSA) with deep belief network (DBN) model for disease detection and classification in the IoT environment. The proposed AFSA-DBN model recognizes the presence or absence of disease on the healthcare data. The presented AFSA-DBN model encompasses three major subprocesses. At the initial stage, the min–max normalization approach is applied to scale the input data into useful format. Followed by, the DBN method can be exploited for the identification and classification of diseases. At last, the AFSA is utilized to optimally choose the hyperparameter values related to the DBN method. To assure the enhanced performance of the AFSA-DBN model, a wide-ranging experimental analysis was conducted. The comparative results assured the improved outcomes of the AFSA-DBN algorithm over recent techniques with maximum accuracy of 0.9111.

Keywords Disease diagnosis · Internet of Things · Deep learning · Machine learning · Data classification

1 Introduction

In the last decade, the progression of sensor and IoT technology based on wearable medical gadgets has improvised the maintenance quality of patients via smart remote health monitoring mechanisms [1]. Currently, the cloud-related IoT networks were implied broadly in medical monitoring and smart remote health mechanisms [2]. The

C. Nithyeswari (✉) · G. Karthikeyan
Department of Computer Science, Periyar Arts College, Cuddalore, Tamil Nadu, India
e-mail: varshnicy2012@gmail.com

grouping of IoT and cloud consist numerous aids from source organization prospects like source spreading, influential processing, ignoring data fragmentation over several files, and guiding user flexibility in monitoring mechanisms [3]. The latest remote healthcare monitoring mechanisms in cloud-related IoT atmosphere involve a background where the biological information of patients can be forwarded and saved in clouds, as well as shared for acquiring analytics from anytime and anywhere [4]. For obtaining diagnostic data to predict the health abnormal changes of patients, data mining techniques were extensively utilized in medical monitoring system which includes clustering and classification techniques, neural networks (NN), and other methods on the basis of distinct machine learning (ML) methods [5].

With the advent of Artificial Intelligence (AI) methods, surgical gadgets, as well as mixed real applications, disease treatment, and diagnosis were more powerful [6]. Through utilization of AI, particular results were gained from clinical decision support system (CDSS) like the prognosis of skin cancer, hepatitis, and lung cancer. Additionally, the accurateness of AI prognosis has surpassed the accuracy established physically [7]. Moreover, ML-related methods were accurate when compared with well-trained doctors, especially imaging experts and pathologists. Thus, an extraordinary and illustrative invention in CDSS has been bowled by IBM's Watson [8]. This invention contains an effectual intellectual system and was utilized to provide the optimal solution by using in-depth analysis of literature and medical particulars. Therefore, a drastic impact was encountered by medical professionals in prognosing cancer as well as diabetes [9]. The application of CDSS was more effective and guides the doctors in advancing the diagnostic procedures, restraining the incidence of unexploited prognosis and mis prognosis, and allowing the users get prompt and appropriate medical treatment. As per smart prognosis, patient's disease severity and health state could be described precisely by following a personalized medication process [10].

Bharathi et al. [11] introduce an Energy Effective PSO related Clustering (EEPSOC) method for selecting the cluster head (CH) effectively among various IoT gadgets. The IoT gadgets employed to sense medical data were merged to cluster and a CH would be selected by considering the usage of EEPSOC. And then, an artificial neural network (ANN)-related classifier method was implemented to diagnose the medical dataset in the cloud server for identifying the disease severity. In [12], a medical recommender mechanism was granted for identifying and treating chronic diseases with the help of IoT gadgets. Mansour et al. [13] suggested a technique employs crow search optimization method related cascaded long short-term memory (CSO-CLSTM) method for diagnosing diseases. To reach superior categorization of the healthcare data, CSO was implied for tuning both 'bias' and 'weights' parameters of CLSTM method. Also, isolation forest (iForest) approach was leveraged in the study for eliminating the outliers.

Praveen et al. [14] provide an oppositional glowworm swarm optimization (OGSO) method related clustering using deep neural network (DNN) termed OGSO-DNN technique for distributed medical systems. The OGSO technique has been implied in this research work for selecting CHs from the existing IoT gadgets. The CHs which are selected transfer the data to cloud server, and after which perform

DNN-related classifying process for medical diagnosis. Anupama et al. [15] offers a new skin lesion diagnosis method that is DL with evolutionary algorithm oriented image segmentation (DL-EAIS) for cloud and IoT-based smart medical ambiances. Mainly, the dermoscopy image was captured by making use of IoT gadgets that can be transferred afterward to cloud server for further diagnosis.

This paper majorly concentrates on the design of artificial fish swarm algorithm (AFSA) with deep belief network (DBN) model for disease detection and classification in the IoT environment. The proposed AFSA-DBN model recognizes the presence or absence of disease on the healthcare data. The presented AFSA-DBN model encompasses three major subprocesses. At the initial stage, the min–max normalization approach is applied to scale the input data into useful format. Followed by, the DBN algorithm can be exploited for the identification and classification of diseases. At last, the AFSA is utilized to optimally choose the hyperparameter values related to the DBN method. To assure the superior performance of the AFSA-DBN model, a wide-ranging experimental analysis is carried out.

2 The Proposed Model

In this study, a new AFSA-DBN model has been established for disease detection and classification in the IoT environment. The proposed AFSA-DBN model recognizes the presence or absence of disease in the healthcare data. The presented AFSA-DBN model encompasses three major subprocesses. At the initial stage, the min–max normalization approach is applied to scale the input data into useful format. Followed by, the DBN model is exploited for the identification and classification of the diseases. At last, the AFSA is utilized to optimally choose the hyperparameter values related to the DBN model.

2.1 Data Pre-processing

At this point, the data is preprocessed in two ways such as data normalization and conversion. At the initial phase, the input dataset in .xls format is transformed to .csv format. In addition, data normalization is implemented by the min–max technique from which the minimum and maximum values are regraded from the presented data. It means the normalization of sample to a high value of one and low value of zero. It is characterized by the following expression.

$$\text{Min – Max. Norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

2.2 DBN Based Disease Detection and Classification

Once the medical data is pre-processed, the DBN module is exploited for the identification and classification of the diseases. DBN belonging to the deep neural network (DNN) using huge amount of hidden layers and several hidden units in each layer [16]. Conventionally, DBN is like restricted Boltzmann machine (RBM) containing output units. In addition, DBN utilizes robust, supervised finetuning scheme for altering the system and greedy unsupervised learning technique to train RBM by labeled dataset. The RBM encompasses v visible layer, h hidden layer, and linked to undirected weight. For accumulating RBM in DBN, hidden state of RBM is regarded as visible layer of future RBM. Figure 1 depicts the framework of DBN. The subset parameter of RBM as (w, b, a) , whereby w_{ij} indicates the weights among v_i and h_j . b_i and a_j are defined by bias of layer. Also, it determines corresponding energy as follows:

$$E(v, h|\theta) = - \sum_i b_i v_i - \sum_j a_j h_j - \sum_i \sum_j w_{ij} v_i h_j \tag{2}$$

The joint possibility distribution of h and v are determined by,

$$p(v, h|\theta) = \frac{\exp(-E(v, h|\theta))}{\sum_{v,h} \exp(-E(v, h|\theta))} \tag{3}$$

The marginal possibility distribution of v is established as follows,

$$p(v|\theta) = \frac{\sum_h \exp(-E(v, h|\theta))}{\sum_{v,h} \exp(-E(v, h|\theta))}. \tag{4}$$

For gaining an optimum value θ for v individual data vector, gradient of log possibility assessment is measured according to the following formula,

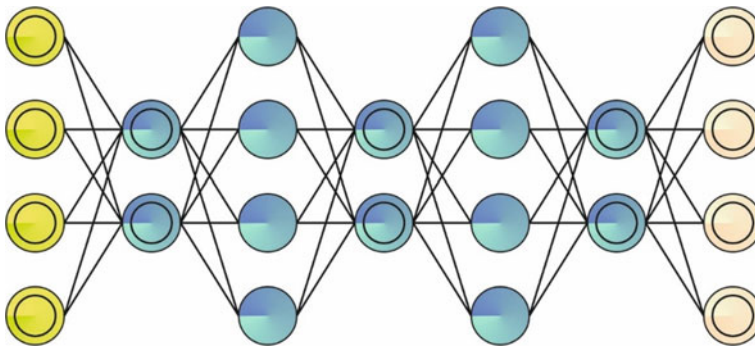


Fig. 1 Framework of DBN

$$\begin{aligned}
\frac{\partial \log p(v|\theta)}{\partial w_{ij}} &= \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}, \\
\frac{\partial \log p(v|\theta)}{\partial a_j} &= \langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{model}}, \\
\frac{\partial \log p(v|\theta)}{\partial b_i} &= \langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{model}}.
\end{aligned} \tag{5}$$

In Eq. (5), $\langle \cdot \rangle$ indicates the expectation using the distribution of particular subscripts. Owing to the nonexistence of interconnections among units in comparable layers, $\langle \cdot \rangle_{\text{data}}$ is easily accomplished by estimating the condition likelihood distribution and signified as follows

$$\begin{aligned}
p(h_j|v, \theta) &= \frac{1}{1 + \exp(-\sum_i w_{ij} v_i - a_j)}, \\
p(v_i|h, \theta) &= \frac{1}{1 + \exp(-\sum_j w_{ij} h_j - b_i)}.
\end{aligned} \tag{6}$$

The activation function can be characterized by using sigmoid activation function. For $\langle \cdot \rangle_{\text{model}}$, Contrastive Divergence (CD) learning mechanism is exploited through reconstructing to decrease the difference of 2 Kullback–Leibler divergences (KL). At first, CD learning is effectual in real-time application and constraint the computation cost in comparison with Gibbs sampling methodology. Consequently, weight in DBN undergoes training via unlabeled dataset by faster and greedy unsupervised methodology. For calculation, supervised layer is incorporated in DBN to modify the learned feature using labeled dataset in the applications of up–down finetuning approach. Currently, the full connection layer serves as the topmost layer, as well as sigmoid function was employed.

2.3 AFSA Based Hyperparameter Tuning

In this study, the AFSA is utilized to optimally choose the hyperparameter values related to the DBN model. The AFSA is composed of two key components [17]. Parameter and function are associated with nearby factors and fish behavior. The parameter involves the visual distance of individual fish (Visual), (δ) the crowd factor of fish, (Step) the size of fish motion, and distance among the two fish indicates the X_i and X_j ($d_{ij} = \|X_i - X_j\|$), whereby $X = (X_1, X_2, X_3, \dots, X_n)$ and $Y = f(x)$. X characterizes a separate state in fish populace as well as Y denotes objective functions or feed focus value. Visual and Step implemented a significant portion in the middle of the four variables. The superior the value of parameter, AFSA moves fast to global optima for the reason that the fish investigate larger space closer to them and move large phase in each iteration. The conditional vector of fish swarm is represented as

follows

$$X = (X_1, X_2, X_3, \dots, X_n) \quad (7)$$

In Eq. (7), X represents the visual, and the fish locations are represented as follows

$$X_v = (X_{v1}, X_{v2}, X_{v3}, \dots, X_{vn}) \quad (8)$$

In Eq. (8), X_v indicates the fish location in Visual. The above-mentioned process (7) has performed by the following equation:

$$X_{vi} = X_{vi} + \text{Visual} \times \text{Rand}(); \quad i \in (0, 1] \quad (9)$$

In Eq. (9), X_{vi} denotes the condition of fish in Visual.

$$X_{\text{next}} = X + [(X_v - X) / (\|X_v - X\|)] \times \text{Step} \times \text{Rand}() \quad (10)$$

In Eq. (10), X_{next} signifies the subsequent fish in Visual. In above-mentioned, Eq. (7) denotes the state of fish, as well as Eq. (8), represents the fish location in Visual. Equation (9) specifies how Eqs. (7) and (8) works collectively, and it demonstrates the coward factor in AFSA. Equation (10) demonstrates the fish is determined by the distance among the two fishes and the Step value. Then, the fish activities denote the function in AFSA is handled by the following expression.

$$X_i = X_i + \text{Visual} \times \text{Rand}(); \quad \text{Search Function} \quad (11)$$

In Eq. (11), Search Function represents search behavior. If $Y_i < Y_j$ if (12) is implemented. Next, Y_i indicates the existing food concentration and Y_j represents the subsequent food concentration.

$$X_i^{(t+i)} = X_i^t + \left[\left(X_i - X_i^{(t)} \right) / \left(\|X_i - X_i^{(t)}\| \right) \right] \times \text{Step} \times \text{Rand}() \quad (12)$$

Or else, implement (11) by arbitrarily choosing a state X_i and inspect the results. While it doesn't fulfill after that try_number, t ($t < \text{Search Function}$) then take a step forward making them flee from the local extrema.

$$X_i^{(t+i)} = X_i^t + \text{Visual} \times \text{Rand}() \quad (13)$$

$$X_i^{(t+i)} = X_i^t + \left[\left(X_j - X_i^{(t)} \right) / \left(\|X_c - X_c^{(t)}\| \right) \right] \times \text{Step} \times \text{Rand}(); \quad \text{Swarm Function} \quad (14)$$

Here, Swarm Function characterizes swarm behavior. Implement (14) if it fulfilled the whole conditions. The existing state of point is X_i ($d_{ij} < \text{Visual}$), or else implement Search Function:

- $(n_f/n) < \delta$
- $Y_c > Y_i$; Y_c represents the central food concentration

$$X_i^{t+1} = X_i^t + [(X_j - X_i^t) / (\|X_j - X_i^t\|)] \times \text{Step} \times \text{Rand}(); \text{ Follow Function} \tag{15}$$

However, Follow Function represents follow behavior.

3 Results and Discussion

In this section, the classification performance of the suggested model is tested using the Cleveland Heart Disease (CHD) dataset [18]. The dataset holds 297 samples with two class labels as displayed in Table 1. A set of 160 samples exist under absence class and 137 samples comes under presence class.

The confusion matrices created by the proposed model is depicted in Fig. 2. On entire dataset, the proposed model has identified 152 samples into absence and 115 samples into presence. Meanwhile, on 70% of training (TR) dataset, the proposed method has identified 111 samples into absence and 74 samples into presence. Eventually, on 30% of testing (TS) dataset, the proposed approach has identified 41 samples into absence and 41 samples into presence.

Table 2 and Figs. 3, 4 highlights the overall classification outcomes of the proposed model under distinct aspects. The experimental values highlighted that the proposed model has shown enhanced performance under each aspect. For instance, on entire dataset, the proposed model has attained average acc_y of 89.90%, sens_y of 89.47%, spec_y of 89.47%, and F_{score} of 89.74%. Concurrently, on 70% of TR data, the proposed technique has obtained average acc_y of 89.37%, sens_y of 88.55%, spec_y of 88.55%, and F_{score} of 89.02%. Next to that, on 30% of TS data, the proposed algorithm has reached average acc_y of 91.11%, sens_y of 91.29%, spec_y of 91.29%, and F_{score} of 91.11%.

The validation accuracy (VA) and training accuracy (TA) gained by the proposed technique on test dataset is demonstrated in Fig. 5. The experimental outcome denoted

Table 1 Dataset details

Class	No. of samples
Absence	160
Presence	137
Total number of samples	297

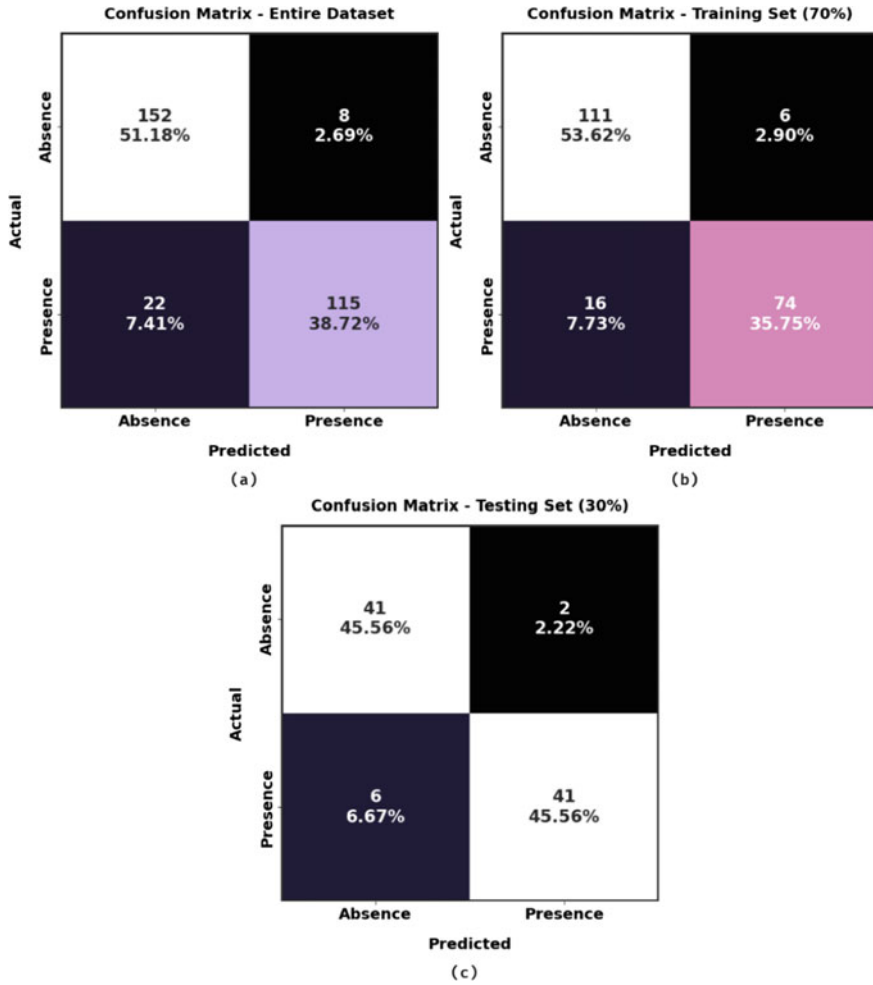


Fig. 2 Confusion matrices of proposed method. **a** Entire dataset, **b** 70% of TR data, and **c** 30% of TS data

the SOSDCNN-HAR model has reached maximum values of TA and VA. To be specific, the VA is greater than TA.

The validation loss (VL) and training loss (TL) acquired by the proposed method on test dataset are exhibited in Fig. 6. The experimental outcome represented the proposed method has accomplished least values of TL and VL. Particularly, the VL is lesser than TL.

Table 2 Result analysis of proposed method with various measures

Class labels	Accuracy	Sensitivity	Specificity	F-score
<i>Entire dataset</i>				
Absence	89.90	95.00	83.94	91.02
Presence	89.90	83.94	95.00	88.46
Average	89.90	89.47	89.47	89.74
<i>Training set (70%)</i>				
Absence	89.37	94.87	82.22	90.98
Presence	89.37	82.22	94.87	87.06
Average	89.37	88.55	88.55	89.02
<i>Testing set (30%)</i>				
Absence	91.11	95.35	87.23	91.11
Presence	91.11	87.23	95.35	91.11
Average	91.11	91.29	91.29	91.11

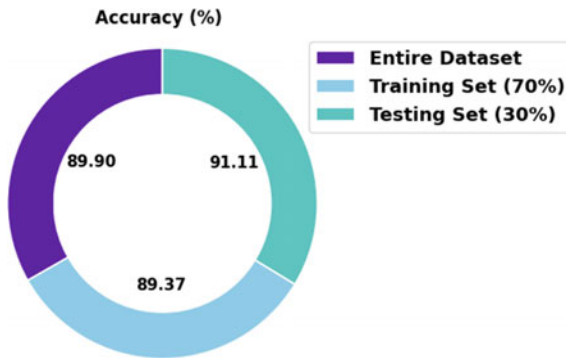


Fig. 3 Accuracy analysis of proposed method with distinct datasets

Table 3 and Fig. 7 illustrate a comparative accuracy inspection of the proposed model with other existing models. The results implied that the NN-Fuzzy, NN-GA, and DT models have demonstrated lower accuracy values of 0.8000, 0.8099, and 0.8068, respectively. Followed by, the ELM, SVM, and DT-GR models have demonstrated reasonable outcomes with accuracy of 0.850, 0.8676, and 0.8410 correspondingly. Though, the proposed model has obtained maximal accuracy of 0.9111. Therefore, the proposed model is found to be effective on medical data classification process.

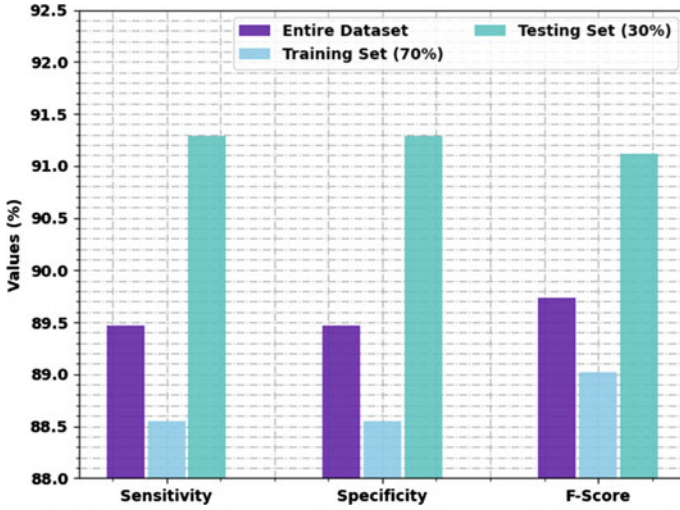


Fig. 4 Result analysis of proposed method with various measures

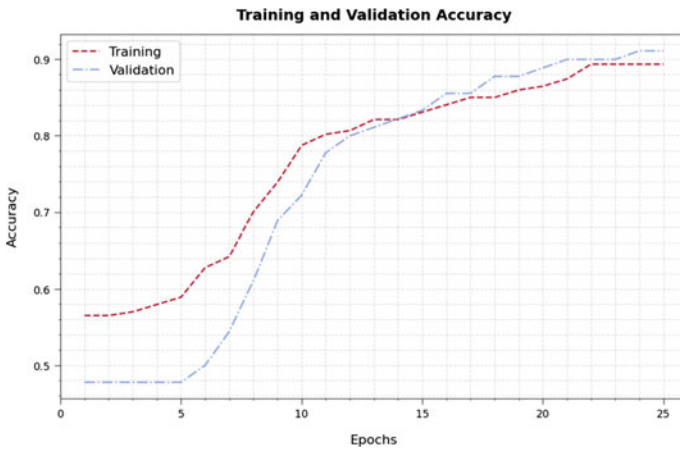


Fig. 5 TA and VA analysis of proposed method

4 Conclusion

In this study, a new AFSA-DBN model has been established for disease detection and classification in the IoT environment. The suggested AFSA-DBN method recognizes

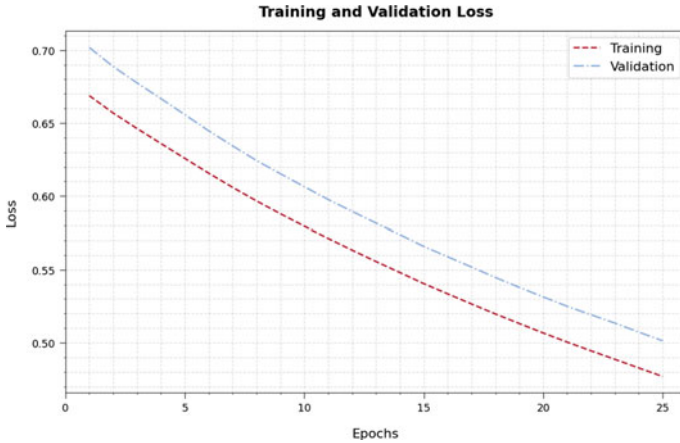


Fig. 6 TL and VL analysis of proposed method

Table 3 Comparative analysis of proposed method with existing approaches

Measures	Accuracy values
The proposed model	0.9111
VNB-LR	0.8741
NN-fuzzy	0.8000
Decision tree	0.8068
ELM model	0.8650
SVM	0.8676
NN-GA	0.8099
DT-GR	0.8410

the presence or absence of disease in the healthcare data. The presented AFSA-DBN model encompasses three major subprocesses. At the initial stage, the min-max normalization approach is applied to scale the input data into useful format. Followed by, the DBN module is exploited for the identification and classification of the diseases. At last, the AFSA is utilized to optimally choose the hyperparameter values related to the DBN module. To assure the enhanced performance of the AFSA-DBN model, a wide-ranging experimental analysis is carried out. The comparative results proved the improved outcomes of the AFSA-DBN model over recent methods with maximum accuracy of 0.9111. In the future, hybrid DL models can be utilized to boost the classifier results.

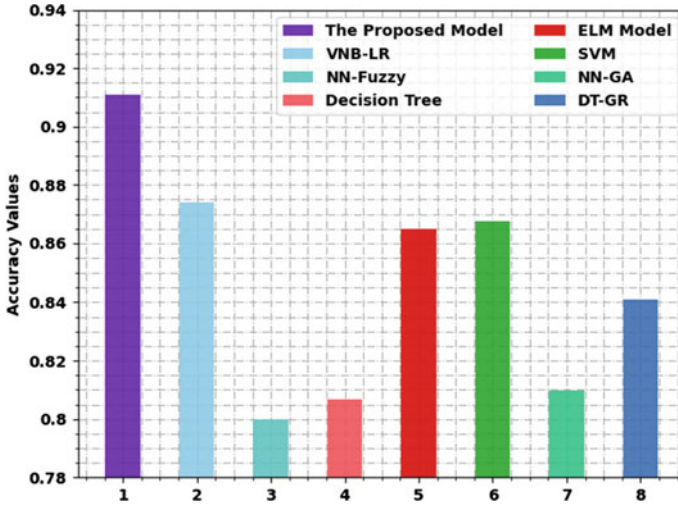


Fig. 7 Comparative analysis of proposed method with existing approaches

References

1. Verma P, Sood SK (2018) Cloud-centric IoT based disease diagnosis healthcare framework. *J Parallel Distrib Comput* 116:27–38
2. Nagarathna CR, Kusuma M (2022) Automatic diagnosis of Alzheimer’s disease using hybrid model and CNN. *J Soft Comput Paradigm* 3(4):322–335
3. Kumar PM, Lokesh S, Varatharajan R, Babu GC, Parthasarathy P (2018) Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier. *Future Gener Comput Syst* 86:527–534
4. Tuli S, Basumatary N, Gill SS, Kahani M, Arya RC, Wander GS, Buyya R (2020) HealthFog: an ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments. *Future Gener Comput Syst* 104:187–200
5. Muthumanjula M, Bhoopalan R (2022) Detection of white blood cell cancer using deep learning using CMYK-moment localisation for information retrieval. *J IoT Soc Mob Anal Cloud* 4(1):54–72
6. Abdali-Mohammadi F, Meqdad MN, Kadry S (2020) Development of an IoT-based and cloud-based disease prediction and diagnosis system for healthcare using machine learning algorithms. *IAES Int J Artif Intell* 9(4):766
7. Latif G, Shankar A, Alghazo JM, Kalyanasundaram V, Boopathi CS, Arfan Jaffar M (2020) I-CARES: advancing health diagnosis and medication through IoT. *Wireless Netw* 26(4):2375–2389
8. Satpathy S, Mohan P, Das S, Debbarma S (2020) A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA. *J Supercomput* 76(8):5849–5861
9. Verma P, Tiwari R, Hong WC, Upadhyay S, Yeh YH (2022) FETCH: a deep learning-based fog computing and IoT integrated environment for healthcare monitoring and diagnosis. *IEEE Access* 10:12548–12563
10. Alotaibi M, Alotaibi SS (2022) Optimal disease diagnosis in internet of things (IoT) based healthcare system using energy efficient clustering. *Appl Sci* 12(8):3804
11. Bharathi R, Abirami T, Dhanasekaran S, Gupta D, Khanna A, Elhoseny M, Shankar K (2020) Energy efficient clustering with disease diagnosis model for IoT based sustainable healthcare systems. *Sustain Comput Inform Syst* 28:100453

12. Nanekaran YA, Licai Z, Chen J, Zhongpan Q, Xiaofeng Y, Navaei YD, Einy S (2022) Diagnosis of chronic diseases based on patients' health records in IoT healthcare using the recommender system. *Wireless Commun Mob Comput* 2022
13. Mansour RF, El Amraoui A, Nouaouri I, Díaz VG, Gupta D, Kumar S (2021) Artificial intelligence and internet of things enabled disease diagnosis model for smart healthcare systems. *IEEE Access* 9:45137–45146
14. Praveen K, Prathap PM, Dhanasekaran S, Punithavathi IS, Duraipandy P, Pustokhina I, Pustokhin DA (2021) Deep learning based intelligent and sustainable smart healthcare application in cloud-centric IoT. *CMC Comput Mater Contin* 66(2):1987–2003
15. Anupama CSS, Natrayan L, Laxmi Lydia E, Wahab Sait AR, Escorcía-Gutiérrez J, Gamarra M, Mansour RF (2021) Deep learning with backtracking search optimization based skin lesion diagnosis model. *Comput Mater Contin* 70(1):1297–1313
16. Geng Z, Li Z, Han Y (2018) A new deep belief network based on RBM with glial chains. *Inf Sci* 463:294–306
17. Zheng ZX, Li JQ, Duan PY (2019) Optimal chiller loading by improved artificial fish swarm algorithm for energy saving. *Math Comput Simul* 155:227–243
18. <https://archive.ics.uci.edu/ml/datasets/heart+disease>

OCR-Based Number Plate Identification Using MATLAB



P. Pooja, G. Maha Lakshmi, R. S. V. S. Vinith, and Siva Sankar Yellampalli

Abstract Licence plate detection is a fully automated real-time approach that has been broadly used for identification, robbery control, and protection validation of motors. Image processing using Matlab is used for number plate detection by following some methods. This paper uses the optical character recognition (OCR) method to read the image of a number plate. Many automobile industries are urging smart detection on vehicles, such as in parking systems, where parking authorities use this system to allow vehicles to park in their area. Previous works went with detection on type and model of vehicle but not the authorized vehicle. Our research work is giving methodology for finding the authorized vehicle by number plate recognition system. OCR process the captured image and read each character in image for recognition by changing the letters in image to text and that can be converted later. The automatic number plate recognition (ANPR) system employs image processing technology. It is one of the systems required for detecting the vehicle number plate, as well as for template matching and precision (result).

Keywords Optical character recognition (OCR) · Number plate recognition · Thresholding · Template matching

P. Pooja (✉) · G. Maha Lakshmi · R. S. V. S. Vinith · S. S. Yellampalli
SRM University-AP, Andhra Pradesh, Mangalagiri, India
e-mail: pola_pooja@srmap.edu.in

G. Maha Lakshmi
e-mail: gurram_maha@srmap.edu.in

R. S. V. S. Vinith
e-mail: ramisetty_sri@srmap.edu.in

S. S. Yellampalli
e-mail: Sivasankar.y@srmap.edu.in

1 Introduction

Today's digital world is full of digital visual communication. This increased the information in image format and has urged to have the efficient robust object recognition methods. Amongst those number plate identification system is emerging and fundamental security system. Image processing was employed in this paper in order to develop a number plate recognition system that has become increasingly important in today's fast-paced environment [1]. The number of automobiles being stolen, disobeying traffic laws, and entering restricted areas is on the rise, as is the number of cars being given a payslip and then being allowed to drive on the road [2]. They also enable the vehicle to park in their territory through this method. With the increasing number of cars on the road in today's world, it's impossible to keep track of them all manually. A person must be present all over a day to keep track of the quantity. It's a time-consuming process that necessitates manpower. Furthermore, manually stored data becomes unreadable over time. So, in order to overcome some of these limitations, we built a machine that could mechanically locate the variety plate and store it in its database. Later on, when the records are needed, one can obtain and use them [3]. The escalating growth of cutting-edge city and countrywide road networks over the past 3 a long time emerged the want of green tracking and control of street traffic. Conventional strategies for traffic measurements, consisting of inductive loops, sensors or EM microwave detectors, suffer from critical shortcomings, costly to install, they call for site visitors disruption in the course of set up or preservation, they may be cumbersome and they are unable to locate gradual or transient prevent automobiles [4]. On the opposite, structures that are based on video are clean to install, use the prevailing infrastructure of visitors surveillance. Furthermore, they may be effortlessly upgraded and that they provide the power to redesign the device and its functionality with the aid of surely converting the gadget algorithms [5]. There is an extensive sort of structures primarily based on video and image processing using different methodologies to detect cars and items. Traffic surveillance device is an lively studies topic in PC vision that tries to hit upon, apprehend and music automobiles for a different types of some images and it helps in adding and making changes and explain the behaviour, vehicle activity through changing the getting older old traditional method of tracking cameras by using human operators [6]. These structures are proving to be useless for busy large locations as the number of cameras exceeds the capability of human specialists.

As a result, it is always beneficial to automate the process of vehicle licence plate recognition. Autonomous Number Plate Recognition (ANPR) identifies the licence plate information of a vehicle from a picture or sequence of images of vehicles and is the best solution to deal with this concern [7]. The use of ANPR for vehicle inspection has risen significantly in recent years. When using an ANPR system, there are three basic steps to follow: (1) Number plate region detection, (2) Character breakdown, and (3) Optical character recognition (OCR). Only beneficial figures and information are obtained for recognition by separating characters on the number plate. ANPR-centered applications were gleaned from a slew of peer-reviewed articles. Artificial

neural networks (ANNs) and probabilistic neural networks (PNNs) artificial neural networks and probabilistic neural networks optical character recognition (OCR), MATLAB, and a configurable method are some of the methods that can be used to create an ANPR system [8]. This work uses a template matching technique to construct an ANPR system for vehicle number plate identification. For this technique, it's all about matching up the vehicle's licence plate number to a template. The modern era of technology necessitates accurate automatic vehicle plate identification [9]. Toll collecting, parking management, access control, and crime investigation are just a few of the concerns it solves. Many experts throughout the world are excited about the potential of ANPR research [10]. Due to the wide variety of licence plates and non-uniform lighting circumstances when taking images of vehicles, this is a difficult challenge to solve.

The next part of the paper is organized as following. A brief explanation of different existing methods were explained in Sect. 2. Different types of algorithms and various techniques for identifying the location of number plate were discussed in Sect. 3. Simulation results from MATLAB followed by experiments were explained in Sect. 4. Conclusion and future scope were explained in Sect. 5.

2 Literary Review

So, many researches went to find and locate the moving objects. Amongst those some studies focused to find the type to vehicle by using the Sobel filtering approach in helps to precisely identify by detecting the vehicle edges [11]. The Contour Let Transform and Support Vector Machine detect vehicle type and model in addition to vehicle detection. These methods were put to the test in the real world and thoroughly examined before being deemed reliable. According to [12], various statistical methodologies were used to investigate the automobile category. In a cluttered/messy situation, the MACH filter was used to reveal the area of interest. All the previous researches worked on finding the vehicle type and model but not the authorized vehicles and motivated to find the particular vehicle. So, finding licence plate helps to detect the authorized vehicle. For Licence plate detection few steps plays major role like character identification, certified plate recognition. Camera, edge-capture device, PC and ANPR image processing and analysis software can be used to perform these tasks.

It is possible to edit and search information from several types of documents thanks to OCR technology, such as scanned paper documents, PDF files, and photos taken with a digital camera. Computers can read handwritten, typed, or printed text by converting scanned images into a form that can be read by OCR systems. Using this method, we are able to save storage space, change the content, and index it. Optical character recognition (OCR) systems play a critical role in document image processing because they help convert electronic photos into electronic text documents that can be altered, searched, indexed, and stored effectively and efficiently. It is possible to extract text from an image using generic OCR systems, although there

are several restrictions to this capability. In most cases, OCR systems are unable to read text from document photos due to technical restrictions, and as a result, the images must be properly prepared before they can be read by OCR systems. In this study, the goal is to establish a few preprocessing processes for document images so that OCR systems can more easily read and understand them [13].

Optimizing the images for recognition helps the OCR software do its job better. Using two non-intersecting pictures data sets, the neural network was able to mimic real-world conditions. In the field of pattern recognition, artificial neural networks (ANNs) are widely employed. To categorize inputs into a collection of target groups, the multilayer feed-forward neural network (ANN) is the most commonly utilized type of algorithm. When it comes to neural network implementation, the most fundamental way is information mining, which is employed in both and to handle the contributions of neural network contributions independently. A two-stage hybrid OCR technique can boost recognition rates. Using four statistical sub-classifiers, each of which recognizes input characters on its own, the Bayes' method is used to integrate the results. An operational stage can be used for further differentiation if a character documented in the first step matches a group of similar characters (e.g., A/4, B/8, and S/5). Jaya Lakshmi et al. [14] used a computer with a Dual Core 2 GHz processor and 8 GB of RAM to test the method. Neural networks can also be created using this method. Six thousand and forty-five distinct binary images of varying tenacity were employed. Initial step is to resize the binarized images of the characters such that they are all of the same size [15]. In order to determine the correct size, a variety of input photos were employed. By employing larger character images, by the complexity of neural network increasing there is possibility to increase the rates of recognition. For the best results, the final neural network is scaled to the appropriate size [16]. All of the methods for identifying automobiles and recognizing number plates covered in the study have their own advantages and disadvantages.

3 Working Method and Design

The goal of this study is how to notice a licence plate in the snapshot you took. It is not uncommon for ANPR systems to make use of cameras with both a monochrome and a colour sensor. Become familiar with the area around the number plate before applying for a number plate ID. The methods we use to locate the position of the number plate in images can be classified into three processing groups. Some programmes use a pattern image, grayscale, and colour to aid in character recognition. We can identify characters by separating them from each other, but we can also achieve this by matching a template or by learning about them. Figure 1 illustrates the various methods in which plate numbers can be recognized.

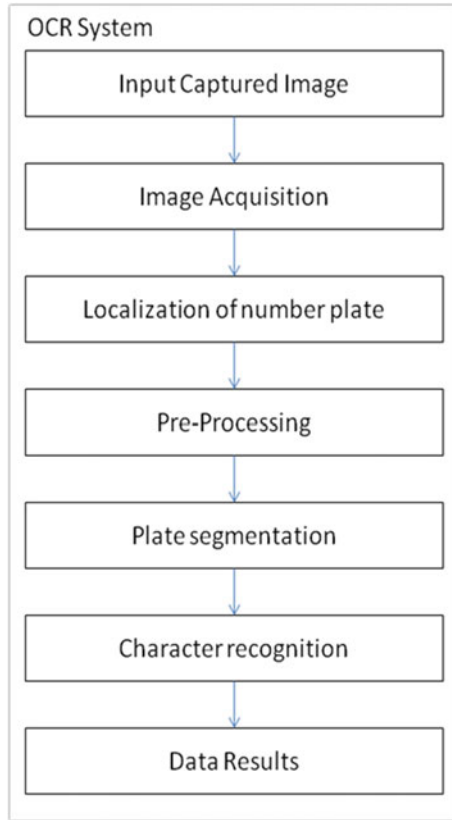


Fig. 1 Block diagram of a car numbering system

3.1 Binary Image Processing

Using edge statistics and morphological approaches, our strategy extracts licence plate regions from backdrop pictures. From a total of 9745 photos, our method has been able to correctly identify 98% of them, assuming that the borders of the number plate frame are perfectly straight. To find the no plate region, this approach of extracting characters from a binary image requires a lengthy analysis of all binary objects. The result will be incorrect if there is any other text in the image.

3.2 Grey-Level Processing

A grayscale image contains only one value for each pixel and is used to record only intensity data. To differentiate between black and white or monochrome, these

images are typically greyscale because the intensity of the colours is separated, with black presenting a low level of intensity and white an extremely high one. To begin, a colour image is converted to a greyscale image. This can be stated as follows: $R = \text{rgb2grey}(p)$, where R is the greyscaled image and p is the colour image.

3.3 Colour Processing

Plate identification relies heavily on colour processing, as most nations mandate that the letters on licence plates be black on a white backdrop. This is particularly true in India, where vehicles are required to adhere to certain colour standards. We need colour processing to get a more accurate character retrieval with more efficiency, but, because of the bad lighting circumstances and plate placement.

3.4 Adaptive Thresholding

The images must first be turned to grayscale before thresholding can begin. A binary image is created through the process of thresholding. As an illustration, the values of each pixel are compared to a predetermined constant (threshold). Pixels of black or white are used to replace values that fall below or exceed the constant value. Threshold values are computed by averaging pixels' local values. In order to formulate the adaptive threshold, the local mean intensity of pixels is taken into account: $O(X, Y) = 255$ if $I(X, Y) < \text{threshold}$, $O(X, Y) = 0$ if $I(X, Y) > \text{threshold}$. Images that are used as input and output, respectively, are referred to as I and O . The window size parameters are chosen based on the character size in the region, m and n .

3.5 Contrast Extension

Histogram equalization is necessary to increase the contrast in an image. The sharpness of the image is enhanced by the contrast extension technique. Here, we are using pixel to find and represent the brightness of that we can use image's grey level histogram. To increase the quality of an image with weak contrast, histogram equalization is used. In total, there are four stages: (i) totalling the data in the histogram in order to normalize the data, (ii) multiply these figures by the image's pixel count, and (iii) to enhance and improve the values we can use or we can increase the grey level value.

3.6 Median Filtering

The noise in an image can be reduced using a median filter. A 3×3 matrix is passed into the image using this way. It is possible to alter these parameters in response to changes in noise levels. Sorting the pixel values and then replacing the pixel that's under consideration with the median value accomplishes this task.

3.7 Character Segmentation

MATLAB's region props function separates the alphanumeric characters from the output image in the resulting output plate region into their own described boxes. The Region props method gives the output as the smallest described box with an individual character. This is the only method can be used here to create and generate a list of all the individual characters from the plate.

3.8 Feature Extraction

In order to extract and find all the image features of a number plate segmented, the Feature Extraction technique is used. The zonal density feature is used to identify characters on licence plate images. Each pixel of an image is counted using the Zonal density function, which divides the image into sections. Each area's pixel density is equal to the object's total pixel density. The image's total area is equal to the number of features it contains. We divide a 32×32 image into 16 zones of different densities, resulting in 16 distinct features per image. In order to have 16, 64, 128 or 256 zones, the pixel must be 32×32 .

3.9 Template Matching for OCR

One character recognition employs a technique known as template matching. Locating a subphotograph known as a template, which is embedded in an image, is the goal of this method. Matching a template to other windows of the same size in a picture requires identifying similarities and differences between the two. As the template is moved about, this approach looks at each individual image pixel for any differences between that pixel and the template's location. Using a database of characters or templates, this technique works. For each and every one of the input characters, a template exists. For all alphanumeric characters, the "normal" font type is utilized (from A–Z and 0–9). Table 1 shows a selection of alphabetic characters with their corresponding template. A possible match or a layout that comes closest

Table 1 Template matching results

Actual plate	Predicted plate	Mismatched character	Accuracy (%)
HR26DK8337	HR26DK8337	0	100
MH12DE14##	MH12DE14##	0	100
HR26DK83##	HR26DK83##	0	100

to portraying the current information character must be found in order for recognition to take place. Moving the standard layout allows it to capture the character's ideal position and perform the correct match. Target characters are coordinated using conventional character format from eight directions: up, down, left and right, as well as upper and lower right. Layout coordination for character identification is shown in Table 1 using still pictures of Indian licence plates. The photos of licence plates shown in Fig. 2 are those that were utilized for format coordination.



Fig. 2 Number plates used for template matching and output images

4 Results

We typically conduct our experiments on a variety of vehicle prototypes of varying shapes and sizes in a variety of environments to measure our methods and precision. The algorithm's accuracy was limited because the segmentation method did not work for plates at a certain angle and plates at the edges of the picture.

A software application is created in this system to detect the licence plate numbers of vehicles by reading their licence plates. In the beginning, the location of the plate is determined through morphological operations, followed by segmentation to separate the plate characters. Plate characters are finally recognized using correlation in the final step of template matching.

5 Conclusion and Future Scope

There has been a thorough analysis of the new technologies and algorithms proposed in this paper for determining the number and type of vehicles with no licence plates. Because an off-the-shelf ANPR device that meets our needs isn't readily available, we've taken it upon ourselves to create one specifically for educational institutions. An average Accuracy has been reached by using template matching on static plates. Table 1 shows the degree of precision with which each character was generated (number 1–9, alphabet A–Z, and a–z). Using two layers of neural networks and positioning the camera such that it captures the exact image of the body, this degree of accuracy may be substantially increased. Multi-level evolutionary algorithms can be used to advance the execution of the stated technique in order to boost the popularity of multiple vehicles with quantity number plates in a single photo body. By using a stationary clip to record images, selecting the best vehicle border for the vehicle category, and applying neural networks to recognize number plates, the device's complexity can be reduced.

Acknowledgements First and foremost, we sincerely thank our esteemed institution SRM University AP, Andhra Pradesh for giving us the opportunity to work on the project. We express our heartfelt thanks to the faculty for their encouragement and supervision provided during the project. We bestow all these credits and feel happy to work under the esteemed guidance.

References

1. Qadri MT, Asif M (2009) Automatic number plate recognition system for vehicle identification using optical character recognition. *IEEE*
2. Puranic A, Deepak KT, Umadevi V (2016) Vehicle number plate recognition system: a literature review and implementation using template matching. *Int J Comput Appl* 134(1). 0975-8887
3. Han B-G, Lee JT, Lim K-T, Chung Y (2015) Real-time license plate detection in high resolution videos using fastest available cascade classifier and core patterns. *ETRI J* 37(2)

4. Ansari NN, Singh AK (2016) License number plate recognition using template matching. *Int J Comput Trends Technol (IJCTT)* 35(4)
5. Shidore MM, Narote SP (2011) Number plate recognition for Indian vehicles. *IJCSNS Int J Comput Sci Netw Secur* 11(2)
6. Swaroop P, Sharma N (2016) An overview of various template matching methodologies in image processing. *Int J Comput Appl* 153(10). 0975-8887
7. Kodwani L, Meher S (2013) Automatic license plate recognition in real time videos using visual surveillance techniques. *ITSI Trans Electr Electron Eng* 1(6). ISSN (PRINT): 2320-8945
8. Islam R, Sharif KF, Biswas S (2015) Automatic vehicle number plate recognition using structured elements. In: *IEEE conference on systems, process and control*, Dec 2015, pp 44–48
9. Chen JIZ (2021) Automatic vehicle license plate detection using K-means clustering algorithm and CNN. *J Electr Eng Autom* 3(1):15–23
10. Erdinc Kocer H, Kursat Cevik K (2011) Artificial neural networks based vehicle license plate recognition. *Procedia Comput Sci* 3:1033–1037
11. Roy A, Ghoshal DP (2011) Number plate recognition for use in different countries using an improved segmentation. In: *2nd national conference on emerging trends and applications in computer science (NCETACS)*, pp 1–5
12. Öztürk F, Özen F (2012) A new license plate recognition system based on probabilistic neural networks. *Procedia Technol* 1:124–128
13. Prabuwo AS, Idris A (2008) A study of car park control system using optical character recognition. In: *International conference on computer and electrical engineering*, pp 866–870
14. Jaya Lakshmi C, Jhansi Rani A, Sri Ramakrishna K, Kanti Kiran M (2011) A novel approach for Indian license recognition system. *Int J Adv Eng Sci Technol* 6(1):10–14
15. Zhai X, Bensaali F (2013) Standard definition ANPR system on FPGA and an approach to extend it to HD. In: *2013 IEEE GCC conference and exhibition, Doha, Qatar, 17–20 Nov 2013*, p 214
16. Jiao J, Ye Q, Huang Q (2009) A configurable method for multi-style license plate recognition. *Pattern Recogn* 42(3):358–369

Machine Learning-Based Classification Between Block Cipher and Stream Cipher



Shivank Kumar Dadhwal and Girish Mishra

Abstract Rivest investigated the relationship between Machine Learning and crypt-analysis in 1991 by emphasizing how ideas from each field benefit the other. Encryption in block cipher is done block by block, whereas encryption in stream cipher is done digit by digit, preferably (bit by bit). The proposed research study has utilized a ML-based approach to distinguish between two types of ciphers. To validate the proposed system, block cipher AES and stream cipher Grain are used. The proposed experiment shows that these encryption algorithms are well randomized and do not provide any regularities in the cipher, which a ML method can use to differentiate between them.

Keywords AES · Grain · Block cipher · Stream cipher · Deep learning

1 Introduction

Deep Learning is a topic of study in which learning methods are based on a framework. Artificial Neural Networks, Recurrent Neural Networks (RNN), Deep Belief Networks and Convolution Neural Networks (CNN) are all the examples of deep learning or deep neural network architectures. These architectures help to solve a variety of real-world challenges with the advanced technologies such as natural language processing (NLP), computer vision (CV), and medical image processing (MIP) [1, 2].

In some applications and areas, these architectures have even outperformed humans, resulting in remarkable results. Deep Learning algorithms offer an advantage over typical ML algorithms due to their ability to do representation learning. Deep learning techniques discover the representation required for feature detection directly from raw data using a class of algorithms. Representation Learning is the

S. K. Dadhwal (✉)

Maharaja Surajmal Institute of Technology, Janakpuri, Delhi 110058, India
e-mail: shivankdadhwal@gmail.com

G. Mishra

Defence Research and Development Organization, Metcalfe House, Delhi 110054, Delhi, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes
in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_38

531

name given to this method. For this technique to work, the data must be computationally convenient and meaningful for processing. Between the input and output layers, a deep learning network has numerous hidden layers. Large number of hidden layers help getting high level features from raw data [3].

The use of Machine Learning (ML), particularly Deep Learning, in conjunction with cryptography has been studied in a limited way. Rivest discussed how each of these fields offer ideas to each other in his paper [4]. Rivest referred to machine learning and cryptography as sibling fields since they both aim to learn unknown functions from input-output pairs. Abadi and Anderson [5] investigated whether neural networks may learn to employ secret keys to secure information during communication from adversarial neural networks in a landmark paper. A side-channel study of block cipher AES was performed by Wang using a deep learning approach. Wang [6] and Gohr [7] recently introduced improved attacks on the round reduced lightweight block cipher Speck32/6410 [8]. Benamira [9] presented a detailed analysis and thorough explanations of the intrinsic workings of the neural distinguisher stated by Gohr, building on Gohr's findings. According to the literature review, the cryptography community has just lately begun to investigate the application of profound learning in cryptography.

Both Machine Learning (ML) and Cryptography deal with a considerable amount of data and a vast search space. Though the application of ML in cryptography has been made in a very confined manner in the past, with time, it has become much more critical to apply ML techniques in cryptography, especially in cryptanalysis.

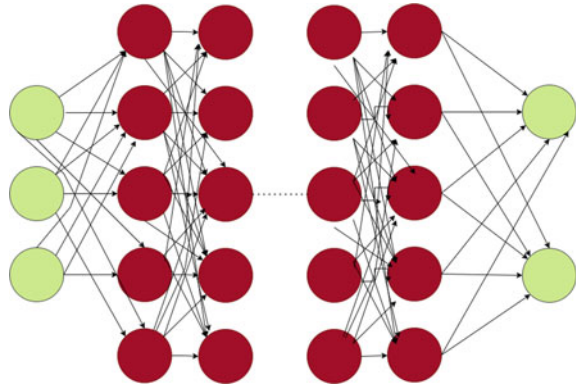
In our work, we try to distinguish between a block cipher and a stream cipher using a machine learning approach. We prepare the cipher data in hex format with the respective class label. If the cipher data is AES encrypted, the class label is taken as 0, and if the cipher data is encrypted using Grain, the class label is taken as 1. In both cases, the cipher data is obtained by encrypting a large plaintext data set. Our experiments show that the two classes do not provide any regularities or patterns to be exploited by machine learning techniques.

Section 2 briefly describes both encryption algorithms used in this paper, i.e., Grain stream cipher and AES block cipher. Also, it talks about deep learning fundamentals. Section 3 discusses about dataset formation, encryption of plain text data into 161 length hex codes and generating dataset with their labels, different modes of encryption used. Sections 4 and 5 explain the experiment results and their conclusions.

2 Preliminaries

This section explains the principles of deep learning, which are required to comprehend the proposed strategy. The AES block cipher and the Grain stream cipher are both described in detail (Fig. 1).

Fig. 1 General deep learning architecture



2.1 Deep Learning Methods

Deep Learning is a subdivision of ML [3, 10–12] which employs large number of artificial neurons layers to get higher-level properties from raw input. In image processing, for example, lower layers discern minor features like edges, whereas higher layers perceive more important items like digits, characters, and faces [11].

A neural network basically has three layers, they are input, hidden and output layer. The input layer gets the raw data, the intermediate hidden layers uses an activation function to process the data (generally, ReLU). The output layer utilizes the Softmax function for final classification [11, 13, 14]. Softmax function is given as:

$$F(x) = e^{x_i} / \sum_{i=0}^k e^{x_i} \quad (1)$$

2.2 Brief Description of AES

AES is the most abundantly used and popular symmetric-key algorithm.

The main features of AES are:

- 128-bit plaintext data, 128/192/256-bit keys
- Symmetric key block cipher
- Designers give design details and full specifications
- Stronger and faster than Triple-DES
- Analyzed comprehensively by the cryptanalysis community.

The input plain text size for all three variants is 128 bits, with critical dimensions being 128, 192, and 256 bits. Its a symmetric block cipher as only a single key is used for encryption and decryption. It is well-known that the form of encryption in which only one key (private key) used for encryption and decryption both is defined as Symmetric Encryption. Its counterpart, Asymmetric Encryption, is a type of encryption in

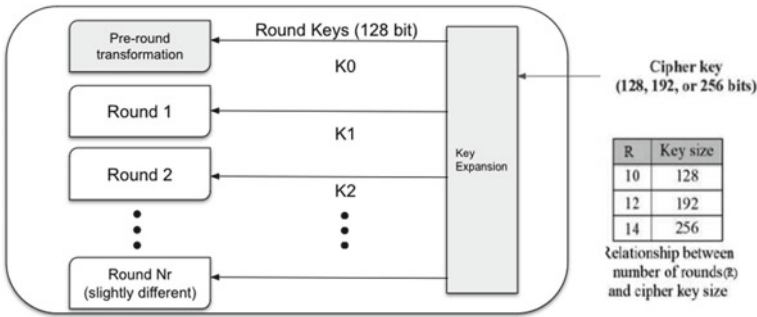


Fig. 2 Schematic of AES

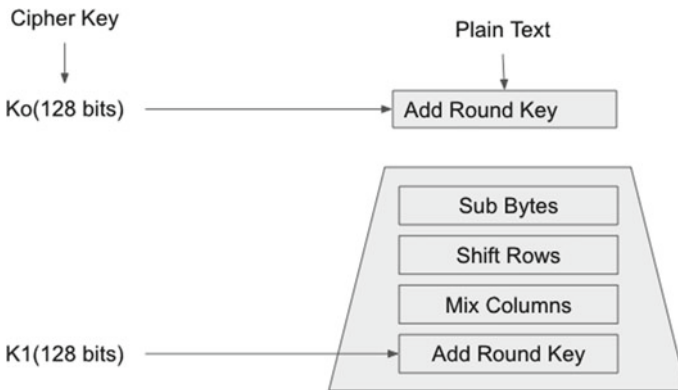


Fig. 3 First round processes

which two different keys (private and public key) are used for encryption and decryption. AES was developed to serve as substitution for Data Encryption Standard which was used prior. DES needed to be phased out as it was increasingly more vulnerable to brute-force attackers. The original development goal of AES is to protect sensitive government information. Still, it was utilized by most organizations in their security process due to its security and ease of implementation. It's comprises of a series of joined operations, some of them involve moving bits about (permutations) and others which involve substituting certain inputs with specific outputs (substitutions). All AES computations are performed on byte data not bits. Resulting, AES divides the plain text data block's 128 bits into 16 bytes. The 16 bytes are organized in a four by four matrix. Figure 2 shows the AES structure [15].

The AES encryption process is divided into rounds, each of which is further divide into sub-processes. Figure shows first round 3 [15].

– BYTE SUBSTITUTION (SubBytes)

The procedure entails looking up a fixed table and substituting 16 input bytes (S-box). Final product is four by four matrix

– SHIFTRROWS

Each of the row is shifted to left. If an entry falls from the right side, it is inserted again.

- No change in the 1st row.
- The 2nd row is shift toward left by one.
- The 3rd row shift toward left by two.
- The 4th row shift toward left by three.

– MIXCOLUMNS

Each column is changed using a special function. The function replaces the old column with new bytes by taking 4 bytes of a column as input. This step is omitted during the last round.

– ADDROUNDKEY

Sixteen byte matrix taken as 128 bits is xor with 128 bit round key. In the final round, this output is called ciphertext. We keep repeating the process till the final round. The AES Structure is seen schematically in Fig. 2. AES encryption in ECB mode is shown in Fig. 4, while AES encryption in CBC mode is shown in Fig. 5.

```

{
expanded_key= [None] * 176
state= new_matrix(4, 4)
rounds= 10
fill_state(state, msg)
key_expansion(key, expanded_key)
add_round_key( msg , key )
for i: round do
    {
        sub_bytes(state)
        shift_rows(state)
        mix_columns(state)
        add_round_key(state, expanded_key[16 * i:])
    }
end
}
sub_bytes(state)
shift_rows(state)
add_round_key(state, expanded_key[160:176])
}

```

Algorithm 1: ECB-AES algorithm

2.3 Brief Description of Grain

Grain is a hardware-oriented stream encryption [15]. Stream ciphers are symmetric ciphers, meaning they encrypt and decrypt with the same key. The key must only be

```

encrypt( plaintext, iv) {
plaintext = pad(plaintext)
blocks = []
previous = iv
for plaintext_block : split_blocks( plaintext ) do
    {
    block = encrypt_AES (xor_bytes( plaintext_block, previous ))
    blocks.append(block)
    previous = block
    }
end
}
return(blocks)
}

```

Algorithm 2: CBC-AES algorithm

known by the entities in between the data that must be transferred. Grain was created with hardware applications in mind. It features two feedback shift registers, linear and non-linear, each with an output of 80 bits. Grain-128 was created as a result of additional improvements, the capacity of feedback registers was increased. Making grain-128 a new target for cryptanalysts as a result. Grain-128a [12] was presented as a new version of Grain-128 that included Message Authentication Code (MAC). We have utilized the 80-bit grain cipher throughout this article. For each bit of ciphertext given by a non-linear filter function grain changes 1 bit of lfsr and 1 bit of nlfsr state. The 80-bit nlfsr is updated using a non-linear five to one Boolean function and a 1 bit linear input selected from the lfsr. The non-linear five to one function takes five bits of the nlfsr state as input. The 80-bit lfsr is updated using a six to one linear function. During keying, the cipher's output is also sent back as linear inputs to both the nlfsr and lfsr update functions.

```

encrypt_GRAIN() {
init(iv,key)
init_clock()
plain = string2bits(plain)
stream = gen_key_stream()
for ix : range(len(plain)) do
    {
    cipher = str(int(bool(int(plain[ix]))xor((next(stream)))).
    cipher = ' ' .join(cipher)
    }
end
}
}

```

Algorithm 3: Grain algorithm

3 DataSet Formation

3.1 Plaintext Collection

In this work, we need a large amount of plaintext data for creating the ciphertext data by using two encryption algorithms mentioned in Sect. 2. This requirement is necessary for passing the ciphertext with their respective labels (0-AES, 1-Grain) to a neural network so that the model. We use the online freely available text edition of the novel *The Art of War* by Sun Tzu [16] as the plaintext data.

3.2 Encryption

To encrypt the plaintext dataset, we utilize the block cipher AES and the stream cipher grain. Different encryption mechanisms, such as AES, are used to implement block ciphers. For our investigations, we used the Electronic codebook (ECB) and Cipher block chaining (CBC) modes.

– Electronic Code Book (ECB)

It is the simplest mode, int it each block of unencrypted data is encrypted by a single key. In case of any key, the plaintext is mapped to its corresponding ciphertext. This is similar to looking for plaintext in a huge codebook and reading the ciphertext that matches it. It helps in encrypting many blocks at the same time. However, because t unencrypted information is revealed in this method, it is hardly used in real life applications.

– Cipher Block Chaining (CBC)

The ciphertext of a block in CBC mode is determined not only by the key, but also by the ciphertext retrieved in the preceding block. It limits parallelization of the encryption process because each block of ciphertext is dependent on all previous blocks of plaintext.

4 Experimental Setup and Results

For experiments, we create two datasets, one consisting of AES implementation in ECB mode and the other consisting of AES implementation in CBC mode. Each dataset is of size $14,094 \times 162$, where 14,094 denotes the number of ciphertext samples and 162 indicates each sample having 161 features (in hexadecimal) and one class label (0-AES, 1-Grain). Both data sets consist of 7047 AES encrypted samples and 7047 Grain-encrypted samples. Hence basically the data sets consist of

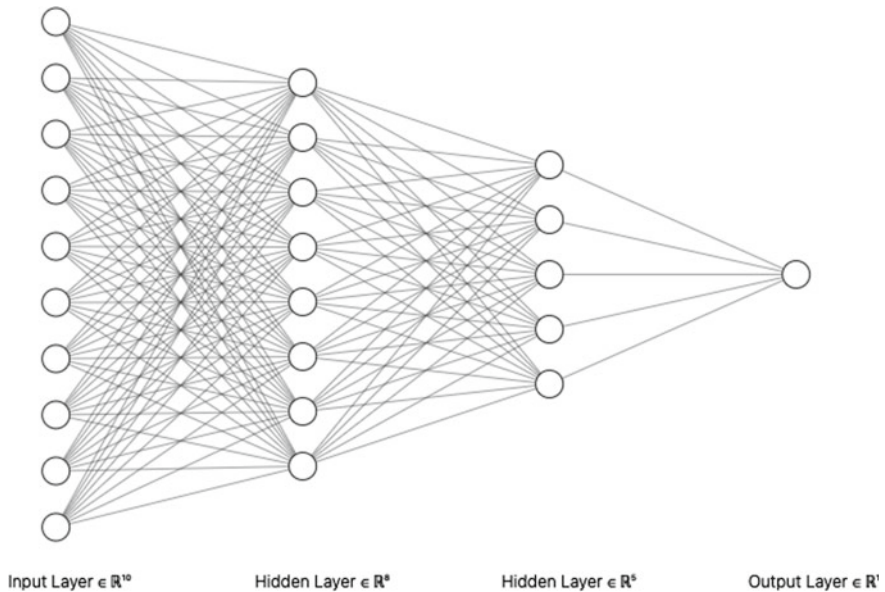


Fig. 4 Neural network architecture

the encrypted text and the label stating which algorithm is used to encrypt it. The encryption text for each row is 161 character long, and each character is represented as their decimal representation.

Google Colab [17] is used to carry out the coding implementation for the specified experiment. Collaboratory (short for Google Collaboratory) is a Google cloud service in the Jupyter notebook environment, here the availability of GPUs and TPUs is leveraged to construct problem-solving codes. It enables the creation of deep learning algorithms utilizing well-known libraries like PyTorch, TensorFlow, and Keras.

4.1 Neural Network Architecture

The neural network architecture chosen for implementation consists of four layers (1 input layer, two hidden layers, and one output layer). The input layer contains ten neurons, the two hidden layers have eight and five neurons, respectively, and the final layer has a single neuron because it is considered as a binary classification task. The architecture of the neural network utilized here is shown in Fig. 4.

Fig. 5 Loss in ECB mode

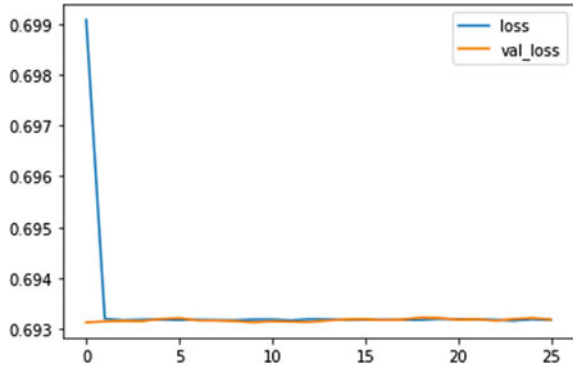


Fig. 6 Loss in CBC mode

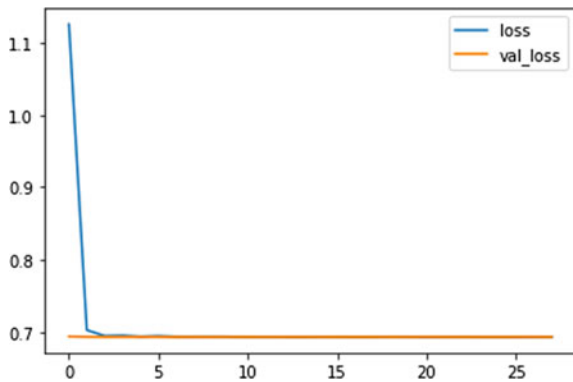


Table 1 Results

Dataset	Accuracy	Precision	Recall	F1 score
EBC	0.50	0.50 (0) and 0.53 (1)	0.92 (0) and 0.09 (1)	0.65 (0) and 0.15 (1)
CBC	0.50	0.50 (0) and 0.00 (1)	1.00 (0) and 0.00 (1)	0.66 (0) and 0.00 (1)

4.2 Results

The neural network is trained using Early stopping with the patience of 25 (For 25 iterations if the value of validation loss is same the model will stop training) and monitor being ‘val_loss’.

Figures 5 and 6 show the graph of training loss and validation loss corresponding to ECB and CBC modes respectively.

It can be inferred from the figures that the results are identical in both the cases (ECB and CBC).

Table 1 shows complete results on ECB and CBC datasets.

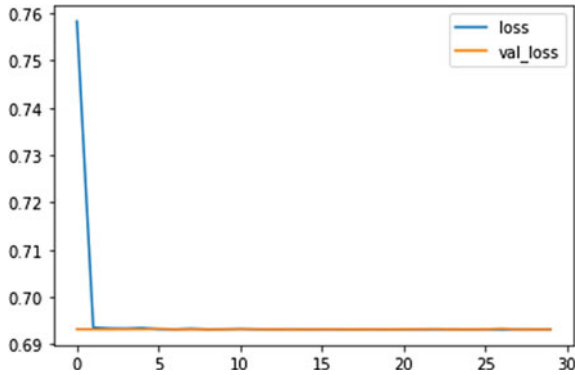


Fig. 7 Loss on larger dataset

	precision	recall	f1-score	support
0	0.50	1.00	0.67	12061
1	0.00	0.00	0.00	12065
accuracy			0.50	24126
macro avg	0.25	0.50	0.33	24126
weighted avg	0.25	0.50	0.33	24126

Fig. 8 Classification report on larger dataset

Table 2 Results

Algorithm	Accuracy
Neural network	0.500
Naive bayes	0.511
Logistic regression	0.515
K-neighbors	0.507
Random forest	0.505

We also perform the classification task using pre-trained algorithms like naive Bayes, random forest, etc. The results of all the algorithm including the Neural network are listed in Table 2.

Table 2 shows that the average accuracy is around 0.50.

In order to confirm our results, we performed the above experiments on a larger dataset with 2^{17} samples.

The larger dataset is formed by combining the text from several novels. This text data was pre-processed and encrypted using CBC-AES and Grain Algorithm.

Figure 7 shows the loss graph and Fig. 8 shows the classification report on the larger dataset.

5 Conclusion

The accuracy of the classification conducted using the neural network is around 0.50 in the classification report in Fig. 8. Table 2 also shows that all the known and well used pre-trained models like Naive Bayes, Logistic Regression, and Random Forest also give an accuracy of around 0.50 while classifying the two algorithms. This level of precision shows that the two algorithms, AES and Grain, have good random properties and are resistant to our machine learning-based attack. Hence both block cipher and stream cipher are robust to Neural Network attacks making them well randomized, Both the small and large dataset experiments support similar observations.

References

1. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473), <https://arxiv.org/abs/1409.0473>. Accessed on 1 Aug 2020
2. Chen C, Seff A, Kornhauser A, Xiao J (2015) Deepdriving: learning affordance for direct perception in autonomous driving. *IEEE* 2722–2730. <https://doi.org/10.1109/iccv.2015.312>
3. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473), <https://arxiv.org/abs/1409.0473>
4. Rivest RL (1991) Cryptography and machine learning. In: Proceedings of the international conference on the theory and application of cryptology. Springer, Berlin, pp 427–439
5. Abadi M, Andersen DG (2016) Learning to protect communications with adversarial neural cryptography. arXiv preprint [arXiv:1610.06918](https://arxiv.org/abs/1610.06918), <https://arxiv.org/abs/1610.06918>. Accessed on 31 July 2020
6. Wang H (2019) Side-channel analysis of aes based on deep learning. <https://www.diva-portal.org/smash/get/diva2:1325691/FULLTEXT01.pdf>. Accessed on 7 Aug 2020
7. Gohr A (2019) Improving attacks on round-reduced speck32/64 using deep learning. In: Proceedings of annual international cryptology conference, Springer, pp 150–179. https://doi.org/10.1007/978-3-030-26951-7_6
8. Beaulieu R, Shors D, Smith J, Treatman-Clark S, Weeks B, Wingers L (2013) The SIMON and SPECK families of lightweight block ciphers. *IACR Cryptol. ePrint Arch*, p 404. <https://eprint.iacr.org/2013/404.pdf>. Accessed on 07 Aug 2020
9. Benamira A, Gerault D, Peyrin T, Tan QQ (2021) A deeper look at machine learning-based cryptanalysis. In: Canteaut A, Standaert FX (eds) *Advances in cryptology—EUROCRYPT 2021*. EUROCRYPT 2021. Lecture Notes in Computer Science, vol 12696. Springer, Cham. https://doi.org/10.1007/978-3-030-77870-5_28
10. Yadav T, Kumar M (2020) Differential-ML distinguisher: machine learning based generic extension for differential cryptanalysis. <https://eprint.iacr.org/2020/913.pdf>. Accessed on 14 Oct 2020
11. Jarrett K, Kavukcuoglu K, Ranzato MA, LeCun Y (2009) What is the best multi-stage architecture for object recognition? In: Proceedings of the international conference on computer vision, *IEEE*, pp 2146–2153
12. Jamil T (2004) The Rijndael algorithm. *IEEE Potentials* 23(2):36–38. <https://doi.org/10.1109/MP.2004.1289996>
13. Karuppusamy P (2021) Machine learning approach to predictive maintenance in manufacturing industry—a comparative study. *J Soft Comput Paradigm* 2(4):246–255

14. Pasumpon Pandian A (2019) Review of machine learning techniques for voluminous information management. *J Soft Comput Paradigm* 1(2):103–112
15. Martin H, Thomas J, Willi M (2007) Grain: a stream cipher for constrained environments. *IJWMC* 2:86–93. <https://doi.org/10.1504/IJWMC.2007.013798>
16. Sun-tzu Griffith SB (1964) *The art of war*. Clarendon Press, Oxford
17. Bisong E (2019) Google colabatory. In: *Building machine learning and deep learning models on Google cloud platform*. Apress, Berkeley. https://doi.org/10.1007/978-1-4842-4470-8_7

Comparative Study of Various Machine Learning Algorithms with MDLDPTS for Plant Leaf Disease Analysis



N. Venkatakrisnan and M. Natarajan

Abstract Agricultural productivity can be much affected because of different plant leaf diseases. Hence, prediction of plant diseases accurately at the earliest is urgent and necessary. Prediction of leaves diseases with lab tests or naked eye is inaccurate, time-consuming and vague. Several classifiers are presented. But, it is unclear that which is intelligently predicted the different leaf diseases of different plants. This study presents a study to evaluate the performance of most commonly used classifiers, namely fuzzy k-NN, naïve Bayes (NB), support vector machine (SVM), fuzzy SVM (FSVM), decision tree (DT), radial basis function neural network, random forest, and self organizing map using various criterion so as to decide which classifier provides the betterment outcome for plant leaves diseases, namely potato, apple, corn, cherry, pepper, tomato, and grapes. This study reveals new insights into the classifiers for leaves diseases identification. The features of leaves are extracted using multi-direction and location distribution of pixels in trend structure (MDLDPTS) which represents a matrices of quantized values of colour/texture/edge against orientations of equal\small\large local trends structure and quantized values of colour/texture/edge against average location of distribution of pixels of equal\small\large local trends structure for orientations 0° , 45° , 90° and 135° respectively on HSV colour space. The experimental outcome evident that fuzzy SVM provides betterment outcome and the attained outcome in terms of accuracy, area under curve (AUC) and receiver operating characteristic (ROC) curve will assist the farmers to take early steps to prevent their plants from becoming irrevocably damage which in turn increases crop production.

Keywords Multi-direction and location distribution of pixels in trend structure · Self-organizing map · Support vector machine · Random forest neural network

N. Venkatakrisnan (✉) · M. Natarajan
Department of Computer and Information Science, Annamalai University, Annamalai Nagar,
Chidambaram, Tamil Nadu 608 002, India
e-mail: narayanav175@gmail.com

1 Introduction

Nowadays, agriculture productivity feeding major sources to economy of any country and the leading factor affecting the agriculture productivity is plant leaves diseases. Lot of problems experienced by the farmers in detecting plant leaf diseases as there are varieties of diseases with similar visual spots and streaks. In addition, the colour, size and shape of visual spots and streaks changes continuously when the disease progress. Therefore, lacking of depth knowledge in leaf diseases fails to make right effective control towards the preventive measures or treatment of leaf diseases at the earliest which affects agriculture production [1]. Hence, the plant leaves need to be monitored continuously against fungi, insect pests, bacteria, viruses, climate and environmental conditions with the help of the agriculture experts. However, a monitoring and detecting plant leaf disease with an expert is subjectiveness, time-consuming, laborious, and less accuracy. Therefore, automated approach is needed to identify different plant leaves diseases. The important aspect of automated system based plant leaf diseases detection is providing proper diagnosis and knowledge regarding the preventive measures, and to select right treatment for each kind of leaf diseases which in turn increases agriculture production and saves time and cost.

Over the decade, several techniques pertaining to computer vision (CV) and machine learning (ML) were suggested to identify the plant leaf diseases. Automatic leaf disease recognition has been suggested in [2] using colour co-occurrence and neural networks (NN). Kulkarni et al. [3] reported Gabor features with NN for different plant leaves diseases. Histogram matching for colour co-occurrence features and edges are performed by Smita and Niket [4] to identify the diseases in plant leaves where the edges are computed separately for R, G and B images, and colour co-occurrence feature was estimated using spatial grey level dependence matrix. In [5], the leaf and lesion region are segmented by applying triangle and simple threshold techniques then by estimating the quotient of lesion and leaf area, leaf diseases are identified and it is exposed that the approach is more accurate and fast. Wang et al. [6] presented wheat and grape leaves disease identification system which performed image resizing followed by cropping, applied median filtering and transform images to $L^*a^*b^*$ from RGB followed by segmentation using k-means clustering then applied morphological operations to fill the holes and estimates discriminative colour, shape and texture features then employed backpropagation neural network for leaf disease prediction.

Recently, Anjna et al. [7] reported capsicum leaf diseases identification system in which k-means approach is employed for extracting the disease infected area from the leaf, grey level co-occurrence matrix is incorporated for calculating texture descriptors of disease infected area, classifiers namely k nearest neighbour (k-NN), SVM, linear discriminant, and decision tree are used and authors revealed that SVM provides betterment results for capsicum leaves. Shrivastava et al. [8] suggested a fully automated system for soya leaf diseases using the parameters namely lesion colour index, ratio for infected area, infection per region and severity index for damage. The authors performed YC_bC_r conversion of RGB image of soya leaf then

segment the leaf from the segregated Y component then the segmented leaf is inverted afterwards segregated Y, C_b and C_r with inverted image are concatenated to form the intermediate image which is then transformed to RGB and is concatenated with inverted image and results in background separated image on which colour filtering followed by morphological operations are performed to compute the disease parameters. Prediction of cotton leaf diseases using the energy of wavelet and SVM is presented in [9]. In [10], region of interest (ROI) is obtained by performing clipping operation then smoothed and contrast enhanced and uses genetic algorithm for segmentation then transform the segmented image into HSI and segregates it into H, S and I afterwards entropy, local level homogeneity, contrast and energies are estimated for colour co-occurrence matrix of H, S and I respectively and used SVM and minimum classification criterion with k-means for classification. The authors revealed that SVM shown betterment result for the prediction of bacterial and fungal diseases in beans, bacterial diseases in rose leaves, sun burn diseases in lemon leaves, and early scorch diseases in banana leaves. In [11], dimensionalities reduced SIFT using the distribution called generalized extreme value (GEV) is suggested for tomato leave disease identification. Subsequently, the authors used Johnson SB distribution instead of GEV in [12]. In both [11, 12], the authors performed segmentation by colour thresholding and incorporated SVM for detection of diseases. Pantaz et al. [13] presented leaf diseases identification system for 18 plants in which GrabCut approach employed in segmentation phase, LBP histograms are estimated for segmented image afterwards classification is performed by SVM. Savita and Parul [14] employed GLCM and performed comparative analysis of various classification approaches namely SVM, k-NN, probabilistic neural network and self-organizing map for plant leaves disease identification and suggested that k-NN seems to be more appropriate. In line with this, many handcrafted features with NN based systems have been reported for leaf disease prediction of various plants [15–18].

On contrary to handcrafted features based systems, many deep learning approaches were revealed in literature for detecting various diseases of leaves. Multiple convolutional neural networks (CNN), namely ResNet50 and InceptionV3 are integrated by Ji et al. [19] for the prediction of leaves diseases in grape plant. In [20], CNN is used to diagnose herbicide damage and nutritional deficiencies in apple plant leaves. Sladojevic et al. [21] employed CaffeNet for diagnosing the leaf diseases of grape, apple, pear, cherry, and peach plants. Brahimi et al. [22] used AlexNet, DenseNet169, ResNet34, Inceptionv3, VGG13 and SqueezeNet for predicting 26 diseases of 14 plants leaves. Recent advances in leaves diseases identification by deep learning approaches were reported in [15–17, 23, 24]. Although deep learning approaches performed betterment than the combination of handcrafted features with NN based approaches, large and diverse datasets are required for efficient training of deep learning approaches [15]. Regrettably, such a large and diverse datasets are so far not available in the domain of leaf disease detection [15]. Hence, handcrafted features based approaches are still the choice of researchers in this domain. Although several studies have been suggested, sufficient accuracy is not attained due to the failure in selection of appropriate feature vector and classifier.

Keeping this in mind, discriminative feature vector has been chosen more carefully and appropriate classifier has been identified after performing exhaustive performance analysis for various classifiers like FkNN [25, 26], NB [27], SVM [28, 29], fuzzy SVM [30], decision tree (DT) [26], radial basis function neural network (RBFNN) [31], random forest (RF) [32], and self organizing map (SOM) [33] in the proposed work. It is revealed in the literature that MDLDPTS [34] and its earlier version [26, 30] significantly outperforms state-of-the-art feature vectors for heterogeneous photo and histological images classification and retrieval, and it might be due to the effective estimation of colour, shape and texture based on correlation of local level structure in form of direction oriented trends. The outcome of the proposed study evident that leaf diseases detection accuracy for various plants is enhanced using MDLDPTS and fuzzy SVM, and fuzzy SVM is found to be more appropriate classifier for leaf diseases of various plants.

The paper is structured as follows. Section 2 deals proposed MDLDPTS and various machine learning approaches. Benchmark dataset and results and discussion are expressed in Sect. 3. Section 4 deals conclusion and research directions for future.

2 Proposed Methodology

2.1 Feature Computation

The proposed work incorporated multi-direction and location distribution of pixels in trend structure descriptor (MDLDPTS) [34] which effectively exemplifies the information change of pixels, spatial arrangement of local structure and correlation among local structure in form of equal\small\large trends for texture, colour and shape descriptors where equal \small \large trends depicts pixel values are same, large to small and small to large, respectively. Besides, MDLDPTS also encodes average location of distribution of pixel values of each trend in local level structure [34].

In order to exploit MDLDPTS, images in RGB are transformed into HSV colour space then colour quantization on H, S and V, texture quantization on V and edge quantization on Sobel operator applied V component image are performed [34] and the level of quantization are set to 108, 20 and 9 for colour, texture and edge information respectively [34]. Consequently, a 3×3 non-overlapping window is moved in left to right direction from left-top most corner to right-bottom most corner for each colour/texture/edge quantized information. Local structures for each 3×3 non-overlapping sub-image are calculated for each texture\colour\edge quantized information in form of trends [34].

MDLDPTS is represented as matrices of colour/texture/edge quantized values versus orientations of equal\small\large trends structure and colour/texture/edge quantized values versus average location for distribution of pixels for equal\small\large trends structure for every orientation. The computed colour, texture

and edge features for trends structure and its orientations (i.e. 0°, 45°, 90° and 135°) is illustrated as follows.

$$F_{\theta} = \{(\theta_E^{Q_c}, \theta_S^{Q_c}, \theta_L^{Q_c}), (\theta_E^{Q_e}, \theta_S^{Q_e}, \theta_L^{Q_e}), (\theta_E^{Q_t}, \theta_S^{Q_t}, \theta_L^{Q_t})\} \tag{1}$$

where orientation (θ) $\in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$, quantized colour value (Q_c) $\in \{1, 2, \dots, 108\}$, quantized texture value (Q_t) $\in \{1, 2, \dots, 20\}$, quantized edge value (Q_e) $\in \{1, 2, \dots, 9\}$, E, S and L symbolizes equal, small and large trends structure correspondingly and $\theta_E^{Q_c}$, $\theta_S^{Q_c}$ and $\theta_L^{Q_c}$ represents the orientation of equal trend structure for Q_c , Q_e and Q_t , respectively and vice-versa. Therefore, the dimension of the matrices $\theta_E^{Q_c}$, $\theta_S^{Q_c}$ and $\theta_L^{Q_c}$ are 108×4 , 9×4 and 20×4 correspondingly and vice versa. The computed location of distribution of pixel for local level trend structure is illustrated as follows

$$F_{\mu} = \left\{ \left(\mu_{E_{\theta}}^{Q_c}, \mu_{S_{\theta}}^{Q_c}, \mu_{L_{\theta}}^{Q_c} \right), \left(\mu_{E_{\theta}}^{Q_e}, \mu_{S_{\theta}}^{Q_e}, \mu_{L_{\theta}}^{Q_e} \right), \left(\mu_{E_{\theta}}^{Q_t}, \mu_{S_{\theta}}^{Q_t}, \mu_{L_{\theta}}^{Q_t} \right) \right\} \tag{2}$$

where μ portrays location of distribution for pixel values, $\mu_{E_{\theta}}^{Q_c}$, $\mu_{S_{\theta}}^{Q_c}$ and $\mu_{L_{\theta}}^{Q_c}$ symbolizes the location of distribution for pixels at θ orientation of equal trends structure for Q_c , Q_e and Q_t respectively and vice-versa. The estimation of μ in each local level trend structure is illustrated as below

$$\mu = \frac{1}{M} \sum_{i=0}^M P_i \tag{3}$$

where P and M symbolizes the pixels and number of pixels in trends structure. Accordingly, MDLDPTS is defined as

$$F = \{F_{\theta}, F_{\mu}\} \tag{4}$$

2.2 Classifiers

Generally classification is performed by machine learning algorithms because of its quite simplistic and accuracy. In some circumstances, classifiers also used as supporting tool to diminish the browsing space. The intention of this study is identifying appropriate classifier for the analysis of different leaves diseases. The classifiers which are significantly performing fine for complicated, vague, and gigantic patterns, namely FkNN, NB, SVM, fuzzy SVM, DT, RBFNN, and SOM are considered in this study.

2.2.1 Fuzzy k-NN Classifier

The choice of classifier is of significant importance to plant leaf disease diagnosis problem. Thus, FkNN [25, 26] is employed as it is simple, results in uncertainty, and easily interpretable. It is evident that FkNN has been studied effectively in various domain and it employed fuzzy set theory to k-NN, named as FkNN. Instead of assigning feature vectors $F = \{F_1, F_2, \dots, F_N\}$ to individual classes, fuzzy memberships of feature vectors F are assigned to C number of classes based on its distance from its k neighbors which results in a fuzzy membership matrix (D) of N versus C where N portrays number of feature vectors and $C = 1, 2, \dots, M$, and each element in D is the degree of fuzzy membership (D_{ij}) in Class C_i for feature vector F_j where $j = 1, 2, \dots, N$ and $i = 1, 2, \dots, M$. In matrix D , value of element D_{ij} is in between 0 to 1 for all i and j and FkNN assigns D_{ij} for each predicted class in accordance with the distance of F from its neighbors. When D_{ij} is 1 for C_i then D_{ij} for C_j (where $j \neq i$) are 0. Always summation of D_{ij} for feature vector F_j should be greater than 0 for class C_i if not, it means class C_i does not exist, and class C_i with the highest D_{ij} is taken as the winner and is estimated using the Eq. (6), and D_{ij} of F is influenced by inverse of distances from nearest neighbors and its class memberships. In FkNN, fuzzy membership of feature vector (F) is exemplified as below

$$D_i(F) = \frac{\sum_{j=1}^k \left(\frac{D_{ij}}{\|F - F_j\|^{\frac{2}{(w-1)}}} \right)}{\sum_{j=1}^k \left(\frac{1}{\|F - F_j\|^{\frac{2}{(w-1)}}} \right)} \quad (5)$$

$$C(F) = \arg_{i=1}^M (D_i(F)) \quad (6)$$

where k symbolizes nearest neighbour numbers and it is chosen deterministically, $\|F - F_j\|$ exemplifies the distance of F and k nearest neighbour F_j using Euclidean metric and w stands for weight parameter for fuzzy distance function. D_{ij} is calculated a-priori from training data and it is a degree of fuzzy membership for j th neighbor to i th class. The best numbers of nearest neighbour (k) is selected through a trial and error process

2.2.2 Naïve Bayes (NB) Classifier

NB [27] algorithm is an intuitive and simple approach amongst other algorithms for classification. Generally, NB is employed for multiple class and large data volume problems, and is frequently incorporated in the field of text, sentiment, and medical image analysis. NB is supervised learning algorithm and rooted in Bayes' theorem with naive assumption which uses strong independence assumptions among each feature in each class for prediction, i.e. even the features are depends on each

other or upon the existence of other features, NB algorithm assumes that all the features are independently contributes to probability. Bayes’ theorem is elucidated mathematically with Eq. (7) as follows

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{7}$$

where A and B were events and $P(A|B)$, $P(B|A)$, $P(A)$ and $P(B)$ represents posterior, likelihood, prior and marginal probabilities and are elucidated as probability of A (Class) on observed B (features), Probability of evidence (B) given the probability of a hypothesis is true i.e. A , Probability of hypothesis i.e. A before observing evidence B , and Probability of Evidence respectively.

2.2.3 Self Organizing Map

SOM [33] is feed-forward network and unsupervised learning algorithm. It uses competitive and cooperation learning to map input space to output space and conserves topological relations in high-dimensional input patterns to exploit its underlying structure and is an inherent feature for the problem.

Let $I = \{I_1, I_2, \dots, I_N\}$ be the input patterns which are connected with neurons $J = \{J_1, J, \dots, J_M\}$ in computational layer and the initial weights for each connections are randomly assigned in the range 0 to 1 where M should be equal to or less than half number of classes in dataset. The algorithm is elucidated as follows.

For each I in training data.

- While training progress, competitive learning is utilized to adjust weights of connections among nodes of input and computational layers.
- The winning neuron (ωN) i.e. neuron whose weight best matches with input pattern is estimated for each input pattern using the following equations.

$$\omega N = \arg \min\{D_j(I), j1, 2, \dots, N\} \tag{8}$$

where

$$D_j(I) = \sqrt{\sum_{i=1}^N (I_i(t) - \omega_{ji}(t))^2} \tag{9}$$

where $I_i(t)$ and $\omega_{ji}(t)$ represents i th input pattern and weight for j th neuron and i th input pattern at t time respectively and Eq. (9) is a discriminant function namely Euclidean distance.

- Gaussian or Maxican Hat are used to determine the wN ’s neighbourhood in order to offer cooperation process among wN and its neighbourhood such that neurons very close and far away to wN are scaled towards wN and least respectively. The

equation for w adjustment is

$$\omega_{ji}(t + 1) = \omega_{ji}(t) + \eta(t) \cdot h_{ji}(t) \cdot (I_i - \omega_{ji}(t)) \quad (10)$$

where $\eta(t)$ depicts learning rate and is as follows

$$\eta(t) = \eta_0 \left(1 - \frac{t}{T}\right) \quad (11)$$

where η_0 and T depicts initial learning rate (always 1) and maximum iteration time. η_0 will become 0 after T .

- Most commonly used Gaussian function is incorporated in this study as neighbourhood function and is given in Eq. (12).

$$h_{ji}(t) = \exp\left[-\frac{D_j(I)^2}{2\sigma^2(t)}\right] \quad (12)$$

where t , $D_j(I)$ and σ represents time, Euclidean distance among I and J and width of Gaussian function respectively and $\sigma(t)$ is expressed as

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau}\right) \quad (13)$$

where σ_0 and τ represents initial value of σ and time constant. σ is steadily decreased during learning progress in order to attain best learning rate.

End For.

2.2.4 Radial Basis Function Neural Network

RBFNN [31] falls on supervised learning, and is more familiar because of its efficiency in classification, simple architecture and fast convergence of optimization. Like SOM, RBFNN also feed forward method. RBFNN has three layers which correspond to input, hidden with activation function and output. The input layer has neurons equals to dimension of input patterns, k-means method decides number of neurons for hidden layer and number of nodes for output layer equals to class number in the dataset. The bias and neurons for input layer is connected to all hidden layer neurons with no weights. Commonly Gaussian function has been used as RBF which is radially symmetric and be an activation function of hidden layer, and expressed as in Eq. (14)

$$\varphi_j = \exp\left(-\frac{\|x_i - c_j\|^2}{2\sigma^2}\right) \tag{14}$$

where x, c, σ, φ_j and $\|x_i - c_j\|$ corresponds to input neuron, centre of hidden neuron, width of hidden neuron, output of hidden neuron and Euclidean distance, and $i = 1, 2, \dots, n; j = 1, 2, \dots, m$. In Eq. (14), σ has been assigned to utmost distance between two hidden neurons.

Subsequently, hidden layer neurons are connected to all output layer nodes with weight. Weighted sum for activation value for each hidden layer neuron with bias is expressed in Eq. (15) as follows and the highest Y_o is used in decision-making.

$$Y_o = b_o + \sum_{j=1}^m \omega_{jo} h_j \tag{15}$$

where Y, b, ω_{jo} and h_j corresponds to output, bias, w_{ij} weight of j th hidden neuron to O th output and j th hidden neuron correspondingly and $o = 1, 2, \dots, p$.

Let $H = [h_1, h_2, \dots, h_m]$ be the hidden neurons of hidden layers and the weights between hidden and output layer is $\omega = [\omega_{11}, \dots, \omega_{m1}; \omega_{12}, \dots, \omega_{m2}; \dots, \omega_{1p}, \dots, \omega_{mp}]$. Therefore, ω is calculated as follows in Eq. (16).

$$\omega = H^+ o = (H^T H)^{-1} H^T o \tag{16}$$

Where H^T is a pseudo-inverse of H .

2.2.5 Decision Tree Classifier

DT [27] is tree structured map of series of related choices for possible solutions and falls in supervised learning. In DT, internal nodes designate feature vector, branches designate decision rules and leaf nodes designate solution, i.e. class label. As DT recursively break down complex data to exploratory knowledge, easy to construct, interpret and understand, it is often used for classification in the field of medicine, marketing, fraud detection, etc. The algorithm for DT is given below.

- Step 1:** Begin with complete data set as root node, say R .
- Step 2:** Compute value for each attribute using any one of the criterion namely entropy, gini index (GI), chi-square, gain ratio (GR), information gain and reduction in variance (RV).
- Step 3:** Select the attribute based on criterion. i.e. attribute with highest value become the root for the information gain criterion whereas attribute with lowest value become the root for the entropy criterion.
- Step 4:** R is splitted into subset of data using selected attribute.

Step 5: On each subset of data, algorithm continues recursively still reaching the leaf nodes.

In this study, entropy criterion is incorporated [] and is estimated as follows in Eq. (17).

$$E(D) = \sum_{i=1}^n -p_{c(i)}(p_{c(i)}) \tag{17}$$

where $p_{c(i)}$, $E(D)$ and D represents probability of class $c(i)$ in a node, entropy for D , measure for disorder of considered data samples, respectively.

2.2.6 Support Vector Machine

SVM [28, 29] falls on supervised learning approach and it has been incorporated in various domains because of its heftiness to high-dimensional data, high generalization capacity, reasonable classification accuracy, and response time. SVM is elucidates as follows [28, 29].

Let training samples $T = \{(y_i, z_i)\}_{i=1}^n$, $y_i \in \mathfrak{R}^d$ is a feature vector, d portrays feature vector dimension and z_i corresponds to class label of y_i , is either -1 or $+1$. In SVM, hyper plane is elucidated as $\omega^T y_i + b = 0$ when training data are linearly separable. It assures $z_i(\omega^T y_i + b) \geq 0$, where ω , b and y_i corresponds to weight vector, offset and data, respectively, and ω is orthogonal to hyperplane. Thus, linearly separable problem could be portrayed as follows [28, 29].

$$\min_{\omega, b} \frac{1}{2} \omega^T \omega \tag{18}$$

Subject to

$$z_i(\omega^T y_i + b) \geq 0$$

Equation (18) is generalized by comprising slack variable τ_i to y_i , while data are nonlinearly separable, The optimal hyperplane problem transformed by τ_i into [28, 29]

$$\min \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \tau_i \tag{19}$$

Subject to

$$z_i(\omega^T y_i + b) \geq 1 - \tau_i \text{ and } \tau_i \geq 0, i = 1, \dots, n$$

where regulation parameter (C) inflicts balance among the minimization for error function and maximization for margin for optimal hyperplane and parameter τ_i measures error. Substituting Lagrange multiplier α_i to y_i and in accordance to Karush–Kuhn–Tucker theorem, dual form of (18) is created as [28, 29].

$$\begin{aligned} \max_{\alpha} \omega(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n z_i z_j \alpha_i \alpha_j k(y_i, y_j) \\ \text{Subject to} \quad &\sum_{i=1}^n z_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1, \dots, n \end{aligned} \tag{20}$$

At last, decision function $h(y)$ of SVM is elucidated as

$$f(y) = \text{sign}(h(y)) = \text{sign}\left(\sum_{i=1}^n \alpha_i z_i k(y, y_i) + b\right) \tag{21}$$

2.2.7 Fuzzy Support Vector Machine

In FSVM [30], each data point is allocated with a membership value, derived from its relative importance in class. Further, with membership value, data points are rightly assigned various degrees of significance to its own classes [30] that evades the phenomenon of over-fitting owing to outliers and noise [30]. It assigns low membership when data point y_i is identified as an outlier. Therefore, its role to total error-term declines. In accordance with [30], fuzzy training set is delineated as $\{(y_i, z_i, \mu_i)\}_{i=1}^n$, where μ_i portrays membership value assigned to y_i .

To decline the sensitivity of less important data points [30], FSVM fuzzifies the penalty term which will be a function of membership values [30], and hence SVM is coined as FSVM [30]. Classification problem [30] is delineated as

$$\min \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \mu_i^m \tau_i \tag{22}$$

Subject to

$$y_i (\omega^T y_i + b) \geq 1 - \tau_i$$

and $\tau_i \geq 0, i = 1, \dots, n$

where μ_i is a membership.

Akin to SVM, optimization problem is resolved with Lagrange multiplier α_i to y_i then in accordance with theorem called Karush–Kuhn–Tucker, dual form of Eq. (22) could be formed as [30]

$$\begin{aligned} \max_{\alpha} \omega(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n z_i z_j \alpha_i \alpha_j k(y_i, y_j) \\ \text{Subject to } \sum_{i=1}^n z_i \alpha_i &= 0, 0 \leq \alpha_i \leq \mu_i^m C, i = 1, \dots, n \end{aligned} \quad (23)$$

It is obviously clear from Eqs. (22) to (23) that upper-bound of Lagrange multipliers is the only change among SVM and FSVM [30]. In FSVM, data points having Lagrange multipliers with equal value possibly will expose different kind of support vectors in the FSVM because of μ_i [30].

2.2.8 Random Forest Algorithm

RF [32] builds set of DTs on randomly chosen subset of training data and takes the votes of DTs to determine the final outcome. RF is ensemble approach and thus it uses a technique called bagging which is known as bootstrap aggregation. In bagging approach, randomly choosing the subsets of training data is known as bootstrap and combining the votes of DTs to determine the majority voting is known as aggregation. RF is easy to use and more flexible approach and it falls on supervised learning. RF is apt for high-dimensional data as it is working with subsets of data and it decreases over-fitting problem with result averaging. The algorithm for RF is elucidated as follows.

Step 1: Samples selected randomly from training data.

Step 2: Build DT for each sample.

Step 3: Compute prediction outcome from each DT.

Step 4: Voting is accomplished for each predicted outcome.

Step 5: Prediction outcome with majority votes become final result.

2.3 Performance Measure for Assessment

The measures employed in this study for evaluating the performance are classification accuracy (CACC), AUC, and ROC curve [35]. ROC curve is plotted using sensitivity and specificity and it shown the logistic model of predictive accuracy, whereas AUC is described as area under ROC curve [35]. The CACC, AUC, sensitivity (SE) and specificity (SP) are estimated as

$$\text{CACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \times 100\% \quad (24)$$

$$\text{SE} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (25)$$

$$SP = \frac{TN}{FP + TN} \times 100\% \tag{26}$$

where TP, FP, TN and FN stands for true-positive, false-positive, true-negative, and false-negative correspondingly.

3 Experimental Results and Discussion

All images utilized in this study were collected from PlantVillage dataset [35] that has 54,303 leaves images of 14 plants. The images of PlantVillage dataset are grouped into 38 classes and all the leaves images have ground truth. In this study, we considered the images of leaves of plants, namely potato, pepper, corn, grapes, cherry, tomato, and apple. The experimental dataset has 554, 199 221 and 99 images of healthy, black rot, scab and cedar apple rust of apple leaves, 409 and 334 images of powdery mildew and healthy of cherry leaves, 187, 432, 354, and 433 images of cercospora-LS, common rust, healthy and northern leaf blight of corn plant leaves, 407, 465, 146, and 379 images of black rot, leaf blight, healthy and esca of grape leaves, 1248 and 2022 images of bacterial spot and healthy of pepper leaves, 1377, 1372 and 193 images of early blight, late blight and healthy of potato leaves, 3404, 2886, 516, 3769, 2195, 1331, 2251, 2195, 2411, 2144 images for target and bacterial spot, mosaic virus, yellow leaves curl virus, late and early blight and healthy of tomato leaves, respectively. Totally, 39,937 images are considered in the proposed study. Few example images of each class of plant leaves are depicted in Fig. 1.



Fig. 1 Example images from different leaf diseases of seven plants

In accordance with [35], absorption light scattering model is incorporated for leaf images enhancement then images transformed to $L^*a^*b^*$ from RGB colour space so as to have expanded spectral range next global thresholding is accomplished to separate plant leaves from background afterwards primary seed points was chosen so as to perform region growing, and then global standard deviation for image is defined as a constraint for grouping the neighbouring pixels with the seed points then leaf boundaries are refined using random sample consensus to have precise leaf region.

The effectiveness and efficiency of MDLDPTS motivated us to choose it as a feature of leaves images in order to identify the leaves diseases effectively. The MDLDPTS typify the information change of pixels and extract the correlation among local structure through trends for colour, texture and shape separately along 0° , 45° , 90° , 135° orientations in global and local levels. Besides, it also exploits local level distribution of pixels of each identified local structures using trends for colour, edge and texture separately along 0° , 45° , 90° , and 135° orientations at both global and local level. The MDLDPTS are exploited from the precise leaf region and are combined to form single feature descriptor.

In training, k -fold cross validation (CV) approach is incorporated where k is set 10. Therefore, experimental image feature vector dataset is divided randomly into 10 subsets. Every time, $k-1$, i.e. nine subsets has been collectively used in training and remaining 1 subset is utilized in testing. Therefore, every feature vector partakes in testing exactly once and partakes in training for $k-1$ times, i.e. nine times. The performance of all the classifiers considered in this study is assessed with test dataset during the execution of k -fold CV. The 10 times found outcome are averaged in order to provide the final result. We estimated four metrics namely CACC, AUC, sensitivity and specificity where sensitivity and specificity are used to plot ROC curve. Totally eight classifiers, namely FkNN, NB, DT, SVM, SOM, RF, fuzzy SVM, and RBFNN are assessed in this study for the task of leaves diseases diagnoses of all the plants considered in this study. During the training process, the parameters of all the aforementioned classifiers are initialized. The CACC and AUC attained after experimentally assessing all the above mentioned classifiers are elucidate in Table 1. It is obviously found from Table 1 that fuzzy SVM beats the other seven classifiers in terms of CACC and AUC, and performance of NB's approach is poorest, SVM, RF, and SOM are above moderate, DT is moderate and RBFNN is below moderate but defeats the performance of NB approach for the entire experimental dataset. In Table 1, it is apparent that the differences between the outcomes of all the machine learning approaches considered in this study are relatively small. However, such small differences shown significantly better outcome in the recognition of plant leaves diseases. The ROC plots using the sensitivity and specificity measures are depicted in Fig. 2 for all the classifiers considered in this study so as to predict the accuracy rate of leaves diseases diagnosis of all the plants considered in this study. It is also apparent that fuzzy SVM attained significantly best output followed by SVM and RF, and SOM stands in middle by defeating NB, DT, FkNN, and RBFNN. Again DT and FkNN are moderate; RBFNN is below moderate and NB offers lowest in terms of ROC for the entire experimental dataset.

Table 1 Classification accuracy (CACC) and AUC for various machine learning algorithms on entire experimental dataset

Classifier	Experimental dataset (leaves of all Plants considered in this study)	
	CACC	AUC
NB	0.917	0.917
RBFNN	0.926	0.926
FkNN	0.944	0.942
DT	0.948	0.947
SOM	0.953	0.951
RF	0.960	0.957
SVM	0.964	0.962
Fuzzy SVM	0.969	0.968

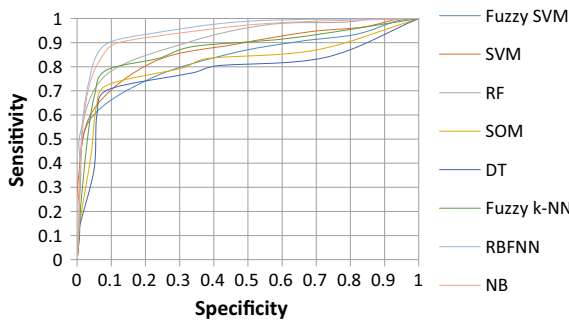


Fig. 2 The ROC curve for the comparison of the various machine algorithms

The proposed study also assessed the performance of all the eight classifiers for each plant leaves individually. The CACC and AUC attained for each plant leaves against eight classifiers are depicted in Table 2. It is very apparent from Table 2 that fuzzy SVM is the outperformer for the leaves of all plants. It is also obvious from Fig. 2 that ROC plot of fuzzy SVM algorithm has much greater area than the other algorithms. The average classification accuracy of all the classifiers for each plant leaves is exemplified in Fig. 3.

By considering all the three metrics into account, this study assert that most consistent, accurate, and repeatable outcome yielded by fuzzy SVM among all the other seven classifiers in the context of diagnoses of different leaves diseases of different plants. It is observed in the results that even SVM is superior then the RF, SOM, DT, RBFNN, and NB classifiers, the accuracy attained by SVM is not significantly better than fuzzy SVM owing to noises or outliers present in training data which roots to over-fitting problem in SVM owing to equally treating all the points of training data. Whereas fuzzy SVM avoided the over-fitting problem by assigning membership

Table 2 Comparative result analysis with existing methods

Classifier		NB	RBFNN	FkNN	DT	SOM	RF	SVM	Fuzzy SVM
Apple	CACC	0.901	0.924	0.959	0.965	0.971	0.978	0.984	0.989
	AUC	0.937	0.945	0.976	0.98	0.984	0.989	0.992	0.995
Cherry	CACC	0.951	0.959	0.977	0.979	0.985	0.989	0.992	0.995
	AUC	0.932	0.947	0.965	0.969	0.974	0.98	0.985	0.991
Corn	CACC	0.909	0.914	0.928	0.933	0.938	0.945	0.949	0.956
	AUC	0.918	0.922	0.937	0.943	0.947	0.955	0.961	0.967
Grapes	CACC	0.927	0.933	0.949	0.953	0.957	0.963	0.967	0.971
	AUC	0.906	0.918	0.927	0.933	0.937	0.944	0.951	0.959
Pepper	CACC	0.951	0.957	0.973	0.977	0.983	0.988	0.991	0.994
	AUC	0.908	0.924	0.935	0.939	0.945	0.951	0.955	0.962
Potato	CACC	0.905	0.913	0.928	0.932	0.936	0.941	0.946	0.951
	AUC	0.913	0.922	0.934	0.939	0.944	0.949	0.953	0.957
Tomato	CACC	0.872	0.881	0.891	0.898	0.904	0.913	0.919	0.924
	AUC	0.902	0.906	0.917	0.923	0.927	0.932	0.937	0.942

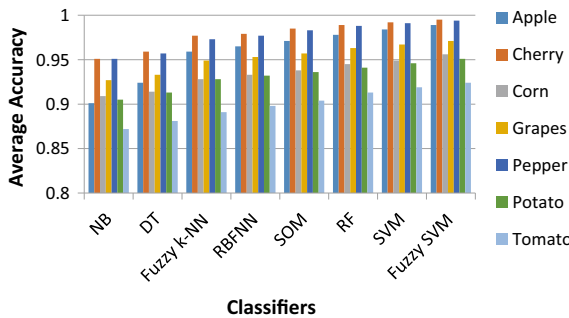


Fig. 3 The average accuracy comparison of various classifiers for different plant leaves

values to each data point of training dataset in accordance with its relative significance in the class. Therefore, data points recognized as outliers has low membership value. Hence, its role to total error term significantly decline. Though the outcome of SOM is very close to SVM, it took long epoch. On the other side, even though DT is faster and flexible then the other classifiers, DT pruning results in accuracy loss. At the same time, time cost of fuzzy SVM is less then SVM, SOM, RF, FkNN, RBFNN and higher than the DT and NB approaches. However, as accuracy is more important than the time cost and nowadays latest machine configurations provides support for speed, time constraint is not taken into consideration in this study.

4 Conclusion

This study performed comparative study of machine learning algorithms namely fuzzy SVM, SVM, RBFNN, SOM, FkNN, DT, and NB for the diagnosis task of different leaves diseases of plants namely potato, pepper, corn, grapes, cherry, tomato, and apple. The classifiers are tested using the effective feature namely MDLDPTS which capture direction oriented local structures in the kind of trends for colour, shape and texture in addition to distribution of pixels along direction oriented local structures in the kind of trends for colour, shape and texture. The experimental outcome shown that NB, RBFNN, FkNN, DT, SOM, RF, SVM, and fuzzy SVM attains 91.7%, 92.6%, 94.4%, 94.8%, 95.3%, 96.0%, 96.4%, and 96.9%, respectively, for overall experimental dataset. Notably, fuzzy SVM shown betterment results then SVM, RBFNN, SOM, FkNN, DT, and NB because of its generalization ability and avoiding over-fitting problem. In future, comparative study of deep learning approaches will be performed in order to assist the researchers in this domain.

References

1. Singh AK, Sreenivasu SVN, Mahalaxmi USBK, Sharma H, Patil DD, Asenso E (2022) Hybrid feature-based disease detection in plant leaf using convolutional neural network, Bayesian optimized SVM, and random forest classifier. *J Food Qual* 2022, <https://doi.org/10.1155/2022/2845320>
2. Badnakhe MR, Deshmukh Prashant R (2011) An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases. *Int Conf Adv Inf Technol* 20, IPCSIT
3. Manoharan JS (2021) Flawless Detection of herbal plant leaf by machine learning classifier through two stage authentication procedure. *J Artif Intell Capsule Netw* 3(2):125–139
4. Smita N, Niket A (2013) Advances in image processing for detection of plant diseases. *Int J Appl Innov Eng Manage* 2(11)
5. Dhaya R (2020) Flawless identification of fusarium oxysporum in tomato plant leaves by machine learning algorithm. *J Innovative Image Proc (JIIP)* 2(04):194–201
6. Wang H, Li G, Ma Z, Li X (2012) Image recognition of plant diseases based on backpropagation networks. In: 2012 5th international congress image signal process CISP 2012, pp 894–900. <https://doi.org/10.1109/CISP.2012.6469998>
7. Anjna MS, Singh PK (2020) Hybrid system for detection and classification of plant disease using qualitative texture features analysis. *Procedia Comput Sci* 167:1056–1065
8. Shrivastava S, Singh SK, Hooda DS (2014) Colour sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation. Springer, Berlin. pp 1–18
9. Bernardes AA et al (2013) Identification of foliar diseases in cotton crop. In: Tavares Joao Manuel RS, Jorge RN (eds) *Topics in medical image processing and computational vision. Lecture notes in computational vision and biomechanics*, pp 67–85
10. Singh V, Misra AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf Process Agric* 4(1):41–49
11. Hlaing CS, Zaw SMM (2017) Model-based statistical features for mobile phone image of tomato plant disease classification. In: *Distrib Comput Appl Technol PDCAT Proc 2018 PDCAT Proc 2018; 2017-Dec*, pp 223–239. <https://doi.org/10.1109/PDCAT.2017.00044>

12. Hlaing CS, Maung Zaw SM (2018) Tomato plant diseases classification using statistical texture feature and colour feature. In: Proceeding of the 17th IEEE/ACIS international conference computer and information science ICIS 2018, 439–444. <https://doi.org/10.1109/ICIS.2018.8466483>
13. Pantazi XE, Moshou D, Tamouridou AA (2019) Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. *Comput Electron Agric* 2019. <https://doi.org/10.1016/j.compag.2018.11.005>
14. Ghaiwat Savita N, Parul A (2014) Detection and classification of plant leaf diseases using image processing techniques: a review. *Int J Recent Adv Eng Technol* 2(3):2347–2812
15. Ngugi LC, Abelwahab M, Abo-Zahhad M (2021) Recent advances in image processing techniques for automated leaf pest and disease recognition—a review. *Inf Process Agric* 8(1):27–51
16. Sherly Puspha Annabel L, Annapoorani T, Deepalakshmi P (2019) Machine learning for plant leaf disease detection and classification—a review. In: International conference on communication and signal processing, 4–6 Apr 2019, India
17. Applalanaidu MV, Kumaravelan G (2021) A review of machine learning approaches in plant leaf disease detection and classification. In: Proceedings of the third international conference on intelligent communication technologies and virtual mobile networks (ICICV 2021), pp 716–724
18. Kaur P, Gautam V (2021) Plant biotic disease identification and classification based on leaf image: a review. In: Proceedings of 3rd international conference on computing informatics and networks, lecture notes in networks and systems 167, https://doi.org/10.1007/978-981-15-9712-1_51
19. Ji M, Zhang L, Wu Q (2020) Automatic grape leaf diseases identification via united model based on multiple convolutional neural networks. *Inf Process Agric* 7(3):418–426
20. Nachtigall LG, Araujo RM, Nachtigall GR (2017) Classification of apple tree disorders using convolutional neural networks. In: Proceeding of the 2016 IEEE 28th international conference tools with artificial intelligence ICTAI 2016 2017:472–476. <https://doi.org/10.1109/ICTAI.2016.75>
21. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D (2016) Deep neural networks based recognition of plant diseases by leaf image classification. *Comput Intell Neurosci* 2016. <https://doi.org/10.1155/2016/3289801>
22. Brahimi M, Arsenovic M, Laraba S, Sladojevic S, Boukhalfa KMA (2018) Deep learning for plant diseases: detection and saliency map visualisation. Springer International Publishing, <https://doi.org/10.1007/978-3-319-90403-0>
23. Lu J, Tan L, Jiang H (2021) Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture* 11:707. <https://doi.org/10.3390/agriculture11080707>
24. Konstantinos P (2018) Ferentinos, deep learning models for plant disease detection and diagnosis. *Comput Electron Agric* 145:311–318
25. Sivakumar S, Sathiamoorthy S (2019) Image retrieval using fuzzy k-NN on dominant rotated local binary patterns. In: 2019 International conference on smart systems and inventive technology (ICSSIT). Tirunelveli, India, pp 806–812. <https://doi.org/10.1109/ICSSIT46314.2019.8987898>
26. Natarajan M, Sathiamoorthy S (2019) Content based medical image retrieval using multi-trend structure descriptor and fuzzy k-NN classifier. In: 2019 international conference on communication and electronics systems (ICCES).Coimbatore, India, pp 2097–2102. <https://doi.org/10.1109/ICCES45898.2019.9002310>. (ISBN: 978-1-7281-1261-9)
27. Sudhir P, Suresh VD (2021) Comparative study of various approaches, applications and classifiers for sentiment analysis. In: Global transitions proceedings, 2(2):205–211
28. Karthikeyan M, Sathiamoorthy S, Vasudevan M (2020) Lane keep assist system for an autonomous vehicle using support vector machine learning algorithm. In: Raj J, Bashar A, Ramson S (eds) Innovative data communication technologies and application. ICIDCA 2019. Lecture notes on data engineering and communications technologies, vol 46. Springer, Cham (ISBN: 978-3-030-38040-3) ISSN: 2367-4512

29. Seetharaman K, Sathiamoorthy S (2014) A framework for colour image retrieval using full range Gaussian Morkov random field model and multi-class SVM learning approach. *Int J Innovative Res Adv Eng* 1(7):53–63 (ISSN: 2349-2163)
30. Natarajan M, Sathiamoorthy S (2019) Heterogeneous medical image retrieval using multi-trend structure descriptor and fuzzy SVM classifier. *Int J Recent Technol Eng (IJRTE)* 8(3):3958–3963, Sept 2019. ISSN: 2277-3878. <https://doi.org/10.35940/ijrte.C5332.098319>
31. Seetharaman K, Sathiamoorthy S (2014) Colour image retrieval using statistical model and radial basis function neural network. *Egypt Inform J* 15(1):59–68
32. Palimkar P, Shaw RN, Ghosh A (2022) Machine learning technique to prognosis diabetes disease: random forest classifier approach. In: Bianchini M, Piuri V, Das S, Shaw RN (eds) *Advanced computing and intelligent technologies. lecture notes in networks and systems*, vol. 218. Springer, Singapore. https://doi.org/10.1007/978-981-16-2164-2_19
33. Sivakumar S, Sathiamoorthy S (2019) Dominant feature descriptors with self organising map for image retrieval. In: book: *Proceeding of the international conference on computer networks, big data and IoT (ICCB—2019)*, Mar 2020. https://doi.org/10.1007/978-3-030-43192-1_60. ISSN: 2367-4512 ISBN: 978-3-030-43191-4
34. Sathiamoorthy S, Natarajan MM (2020) An efficient content based image retrieval using enhanced multi-trend structure descriptor. *SN Appl Sci* 2:217. <https://doi.org/10.1007/s42452-020-1941-y>
35. Kurmi Y, Gangwar S (2021) A leaf image localization based algorithm for different crops disease classification. *Inf Process Agric*
36. Vijai S, Mishara AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf Process Agric* 4:41–49

Arithmetic Optimization Algorithm with Deep Learning-Based Medical X-Ray Image Classification Model



T. Kumar and R. Ponnusamy

Abstract Recently, number of medical X-ray images being generated is increasing rapidly due to the advancements in radiological equipment in medical centres. Medical X-ray image classification techniques are needed for effective decision making in the healthcare sector. Since the traditional image classification models are ineffective to accomplish maximum X-ray image classification performance, deep learning (DL) models have emerged. In this study, an Arithmetic Optimization Algorithm with Deep Learning-Based Medical X-Ray Image Classification (AOADL-MXIC) model has been developed. The proposed AOADL-MXIC model investigates the available X-ray images for the identification of diseases. Initially, the AOADL-MXIC model executes the pre-processing step using the Gabor filtering (GF) technique to eliminate the presence of noise. In the next level, the Capsule Network (CapsNet) model is utilized to derive feature vectors from the input X-ray images. Furthermore, for optimizing the hyperparameters related to the CapsNet approach, the AOA is exploited. Finally, the bidirectional gated recurrent unit (BiGRU) model is employed for the classification of medical X-ray images. The experimental result analysis of the AOADL-MXIC technique on a set of medical images stated the promising performance over the other models.

Keywords X-ray images · Arithmetic optimization algorithm · Deep learning · Feature extraction · Hyperparameter tuning

1 Introduction

Medical image classification refers to a sub-topic of image classification. Several methods in image classification could also be utilized on it. Such as several image advanced methodologies for enhancing the discriminable features for classification [1]. Classifying the medical images effectively serves an important part in helping

T. Kumar (✉) · R. Ponnusamy
Department of Computer and Information Science, Annamalai University, Annamalai Nagar,
Chidambaram, Tamil Nadu, India
e-mail: kumarddeau@gmail.com

health care and treatment. For instance, an analysis X-ray considers an optimal technique in diagnosing pneumonia that accounts for around 50,000 individuals' death annually in the United States [2], but pneumonia classification from chest X-rays requires expert radiotherapists which was costly and scarce resource for certain areas. In medical image diagnosis, chest radiography becomes one of the main diagnostic imaging processes [3]. The ultimate goal was to assess and identify lung diseases like pleural thickening, cardiomegaly, and pneumothorax. This process was deeply reliable for the radiotherapists in some cases might adversely affect the productivity and excellence of radiotherapists' tasks [4]. Additionally, in those cases wherein there is an absence of radiotherapist, the analysis done by the corresponding medical official cannot be precisely examined and thereby needs a second opinion. The development of medical image analysis has evolved as computer-aided diagnosis (CAD) [5]. The first and foremost objective was to guide in discovering the lesion location and to predict the disease chances in the chest X-ray (CXR) image. One such main technological prospect for CAD is to classify and detect diseases on the fundamental of machine learning (ML) [6].

The utilization of the classical ML techniques, like support vector methods (SVMs), in medical image classification, was initiated previously [7]. But such methodologies comprise the subsequent drawbacks are the performances were not matched with the practical standard, and the advancing them was relatively slow in the last few years. Moreover, the feature extraction and selection were time taking and change in accordance with distinct objects [8, 9]. The deep neural networks (DNN), mostly the convolutional neural networks (CNNs), were broadly utilized in varying image classification errands and reached substantial outcomes in 2012. Certain research works on medical image classification by CNN attained outcomes competing human professionals. [10]

In this study, an Arithmetic Optimization Algorithm with Deep Learning-Based Medical X-Ray Image Classification (AOADL-MXIC) model has been developed. The proposed AOADL-MXIC model primarily employs image pre-processing using the Gabor filtering (GF) technique. Besides, the Capsule Network (CapsNet) model is utilized to derive feature vectors from the input X-ray images. Besides, for optimizing the hyperparameters related to the CapsNet approach, the AOA is exploited. Finally, the bidirectional gated recurrent unit (BiGRU) model is employed for the classification of medical X-ray images. The experimental result analysis of the AOADL-MXIC model takes place on a set of medical X-ray images.

2 Related Works

Gour and Jain [11] developed an uncertainty-aware CNN method, known as UA-ConvNet, for the automated identification of COVID-19 disease from CXR images, having a prediction of connected uncertainty from the estimations of the methods. The suggested technique uses the Monte Carlo (MC) dropout and EfficientNet-B3 method, wherein an EfficientNet-B3 method was finely tuned on the CXR images. In

[12], the researchers suggest a CXR image classifier technique on the basis of feature fusion of a visual geometry group network (VGG16) and a dense convolutional network (DenseNet). This article includes an attention system (category attention block global and attention machine block) to the method for extracting deep features. A residual network (ResNet) was utilized for segmenting efficient image information to rapidly attain precise categorization. Bhattacharyya et al. [13] suggest a novel technique to detect pneumonia and coronavirus infection with the help of CXR images. The suggested technique is defined as a three-step procedure. A primary step adds the raw X-ray images segmentation by making use of the conditional generative adversarial network (C-GAN) to acquire the lung images. During the next stage, providing segmented lung images to a new pipeline merging key points extracting techniques and well-trained DNNs to extract the discriminatory features.

Munusamy et al. [14] designed a new FractalCovNet architecture with the help of U-Net and fractal blocks for segmenting chest CT scan images for localizing the lesion area. The same FractalCovNet infrastructure was also utilized to classify CXR images with utilize of transfer learning (TL). A comparison was made for the segmenting outcomes with the help of several methods like FCN, U-Net, ResnetUNet, DenseUNet, and Segnet. Chen et al. [15] examine a new label co-occurrence learning (LCL) structure on the basis of graph convolution networks (GCNs) to explicitly discover the dependencies among pathologies for the multi-label CXR image classifying mission that can be called as “CheXGCN”. To be specific, the suggested CheXGCN is made up of 2 modules that are the LCL and image feature embedding (IFE) modules.

3 The Proposed Model

In this study, a novel AOADL-MXIC model was introduced for the effectual detection and classification of medical X-ray images. The presented AOADL-MXIC approach primarily applied GF approach to pre-process the images. Next, the feature extraction process was implemented by the use of AOA with CapsNet model. At last, the BiGRU model is utilized for the classification of medical X-ray images. Figure 1 depicts the block diagram of AOADL-MXIC approach.

3.1 Image Pre-processing

Firstly, the proposed AOADL-MXIC model primarily applied GF approach to pre-process the images. The GF technique can be modelled as a difficult sine function modulated using Gaussian function whose scale parameter is σ , aspect ratio is λ , whereas the X-axis is at an angle with major axis [16]. It is defined as follows:

$$h(x, y) = g(x', y') \exp [2\pi j(Ux + Uy)] \quad (1)$$

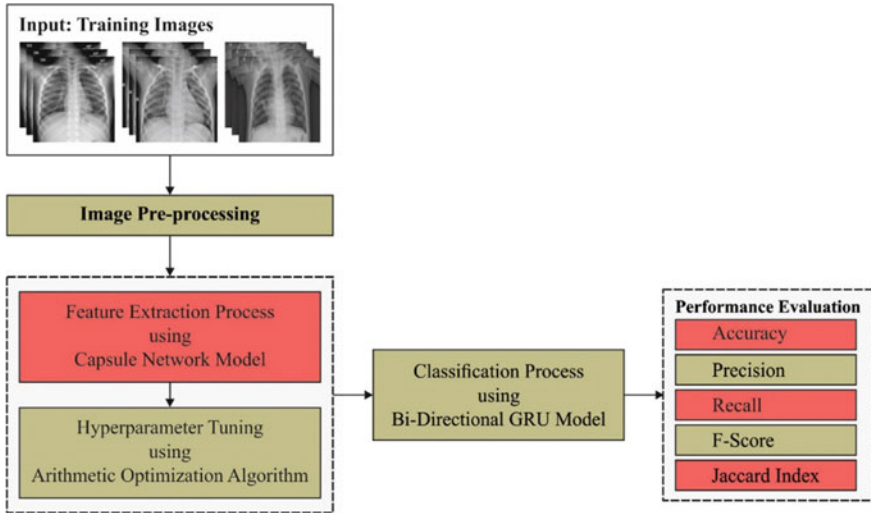


Fig. 1 Block diagram of AOADL-MXIC method

where $(x', y') = (x \cos \varphi + y \sin \varphi, -x \sin \varphi + y \cos \varphi)$ refers the rotation coordinate, and 2D Gaussian function was represented by

$$g(x, y) = \frac{1}{2\pi \lambda \sigma^2} \exp \left[-\frac{(x/\lambda)^2 + y^2}{2\sigma^2} \right] \tag{2}$$

In the Fourier transform defined in Eq. (1), the frequency-domain expression of the 2D Gabor function is given below:

$$H(u, v) = \exp \left\{ -2\pi^2 \sigma^2 \left[(u' - U')\lambda^2 + (v'V'^2) \right] \right\} \tag{3}$$

where $(u', v') = (u \cos \varphi + v \sin \varphi, -u \sin \varphi + v \cos \varphi)$, (U', V') is the same rotation complete by the central frequency (U, V) .

3.2 Feature Extraction Using CapsNet Model

At this stage, the pre-processed images are passed into the CapsNet model to generate feature vectors [17]. CapsNet uses dynamic routing between capsules, to overcome these shortcomings. CapsNet has the capacity to encode spatial datasets and differentiate between different poses, textures, and orientations. A capsule is a set of neurons, where every capsule has an activity vector related to it that captures many instantiating parameters for recognition of certain kind of object or its part. The orientation and length of the vector present the probability or the likelihood of existence

of that object and its generalized pose. That vector is passed onto the upper level capsule from lower layer capsules. A coupling coefficient exists among those layers of capsule. If the prediction by the lower level capsules matches the output of the existing capsules, then value of coupling coefficients among them rises, calculated by using softmax function. Specifically, if the existing capsule identifies a tight cluster of preceding prediction, which indicates the incidence of that object, it leads to a higher probability, called routing by agreement and it is expressed in the following equation,

$$\hat{u}_{j|i} = W_{ij}u_i, \tag{4}$$

In Eq. (4), $\hat{u}_{j|i}$ represents the prediction or output vector of the upper-level j -th capsules, W_{ij} and u_i indicates the weight matrixes and predictive vector of i -th capsules in the lower layer correspondingly. It might capture spatial relations and relationships among objects and sub-objects. In Eq. (5), depending on the degree of agreement among adjacent layer capsules, the coupling coefficient is calculated by using the softmax function,

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})}, \tag{5}$$

In Eq. (5), b_{ij} represents the log probabilities among the two capsules, initialized to zero, and k indicates the capsule count. The input vector s_j to the j -th layer capsules that a weighted sum of $\hat{u}_{j|i}$ vector learned through the routing technique is computed by using Eq. (6),

$$s_j = \sum_i c_{ij}\hat{u}_{j|i}, \tag{6}$$

At last, a squashing function which integrates squashing and unit scaling is employed for confining the value of output in the ranges among zero and one, thus calculating the probability as follows,

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|}, \tag{7}$$

The loss function is related to the capsule in the final layer, where $m+$ and $m-$ are fixed to 0.9 and 0.1.

$$l_k = T_k \max(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k) \max(0, \|v_k\| - m^-)^2, \tag{8}$$

In Eq. (8), the value T_k is 1 for accurate label and 0 or else, λ refers to the constant that value is 0.5.

3.3 Hyperparameter Tuning Using AOA

For optimal hyperparameter tuning of the CapsNet model, the AOA is utilized [18]. During the AOA approach, four fundamental arithmetic operators were established for optimizing the exploration stage and development stage of the technique correspondingly.

Exploration stage: Based on the arithmetic operator, the mathematical computation executed by the division as well as multiplication operators attains extremely distributing values. It can be precise due to their extremely distributing nature which can be impossible for approaching the target simply. Noticeably, the features of these 2 operators were very appropriate that utilized in the searching phases. It is demonstrated as:

$$X(t + 1) = \begin{cases} X_{\text{best}}(t) \div (\text{MOP} + \varepsilon) \times ((UB - LB) \times \mu + LB) \\ r_2 < 0.5 \\ X_{\text{best}}(t) \times \text{MOP} \times r_2 < 0.5((UB - LB) \times \mu + LB) \\ \text{otherwise} \end{cases} \quad (9)$$

$$\begin{cases} \text{MOP}(t) = 1 - \frac{t^{\frac{1}{\alpha}}}{T^{\frac{1}{\alpha}}} \\ \text{MOA}(t) = \text{Min} + t \times \left(\frac{\text{Max} - \text{Min}}{T} \right) \end{cases} \quad (10)$$

whereas ε refers the smaller integer and μ signifies the control parameter in the searching procedure that is set at 0.5. The math optimizer probability (MOP) is a coefficient, α represents the sensitive coefficient and is utilized for defining the progress accuracy, set at 5. Math optimizer accelerated (MOA) has utilized for selecting the searching phase, and Max and Min were the maximal and minimal values of acceleration functions.

Exploitation stage: Due to arithmetic, subtraction, and addition, operators attain higher-density outcomes with mathematical computations, while these 2 operators are minimal dispersion and are simple for approaching the target that is undoubtedly in line with the features of exploitation phase. Thus, these 2 operators were established for exploitation phase.

$$X(t + 1) = \begin{cases} X_{\text{best}}(t) - \text{MOP} \times ((UB - LB) \times \mu + LB) \\ r_3 < 0.5 \\ X_{\text{best}}(t) + \text{MOP} \times ((UB - LB) \times \mu + LB) \\ \text{otherwise} \end{cases} \quad (11)$$

The AOA model mainly determines a fitness function (FF) for optimal parameter tuning process. The fitness function involves the minimization of classification error rate and thereby maximizes the classification accuracy. It can be defined as follows.

$$\begin{aligned} \text{fitness}(x_i) &= \text{Classifier Error Rate}(x_i) \\ &= \frac{\text{number of misclassified samples}}{\text{Total number of samples}} * 100 \end{aligned} \quad (12)$$

3.4 Image Classification

Lastly, the BiGRU approach was employed for the classification of medical X-ray images. GRU is a variant of long short-term memory (LSTM), which combines input and forget gates into a novel gate, named an update and reset gate [19]. The GRU does not utilize hidden and unit layers for transmitting data; hence, two gated units are utilized, an update and reset gates. The reset gate function is to control the amount of disregarding the data at the preceding moment. Small the value is great the amount of disregarding data. The update gate function is to control the amount where the data of preceding moment is transferred to the present state. Large values indicate that more data from the preceding moment is transferred. In comparison with LSTM, GRU has one lesser gating function; hence, the parameter amount is lesser when compared to LSTM, and the computation speed is fast when compared to LSTM.

$$z_t = \sigma(W_i x_t + U_i h_{t-1}), \quad (13)$$

$$r_t = \sigma(W_j x_t + U_j h_{t-1}), \quad (14)$$

$$h_t = (1 - z_t) \circ \tanh(r^t \circ U h_{t-1} + W x_t) + z_t \circ h_{t-1}, \quad (15)$$

From the equation, r_t denotes the reset gate, z_t refers to the update gate, W_i , U_i and W_j , U_j indicates the weight of update and reset gates. The output data is mapped together on the hidden state for obtaining the concluding outcome that is utilized as the input of following layer and it is formulated by:

$$h_t = \left[\vec{h}_t, \overleftarrow{h}_t \right]. \quad (16)$$

4 Performance Validation

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600 k, GeForce 1050Ti 4 GB, 16 GB RAM, 250 GB SSD, and 1 TB HDD. The experimental result analysis of the AOADL-MXIC approach is tested utilizing a benchmark COVID-19 radiography database from Kaggle repository [20]. The dataset holds a set of 1500

Table 1 Dataset details

Class name	No. of images
COVID-19	500
Normal	500
Viral Pneumonia	500
Total number of images	1500

images with a total of 500 images under three classes, namely COVID-19, Normal, and Viral Pneumonia as exhibited in Table 1. A few sample images of three classes are depicted in Fig. 2. For experimental validation, 70% of training dataset and 30% of testing data is used.

Figure 3 depicts the confusion matrix generated by the AOADL-MXIC model on 70% of training (TR) data. The figure represented that the AOADL-MXIC system has identified 330 samples under COVID-19 class, 342 samples under Normal class, and 350 samples under Pneumonia class.

Table 2 and Fig. 4 depict an overall X-ray classification performance of the AOADL-MXIC model under distinct classes on 70% of TR data. The experimental values exhibited that the AOADL-MXIC approach has reached maximal performance in each class. For instance, the AOADL-MXIC model has recognized X-ray images under COVID-19 class with $accu_y$ of 97.43%, $prec_n$ of 97.06%, $reca_l$ of

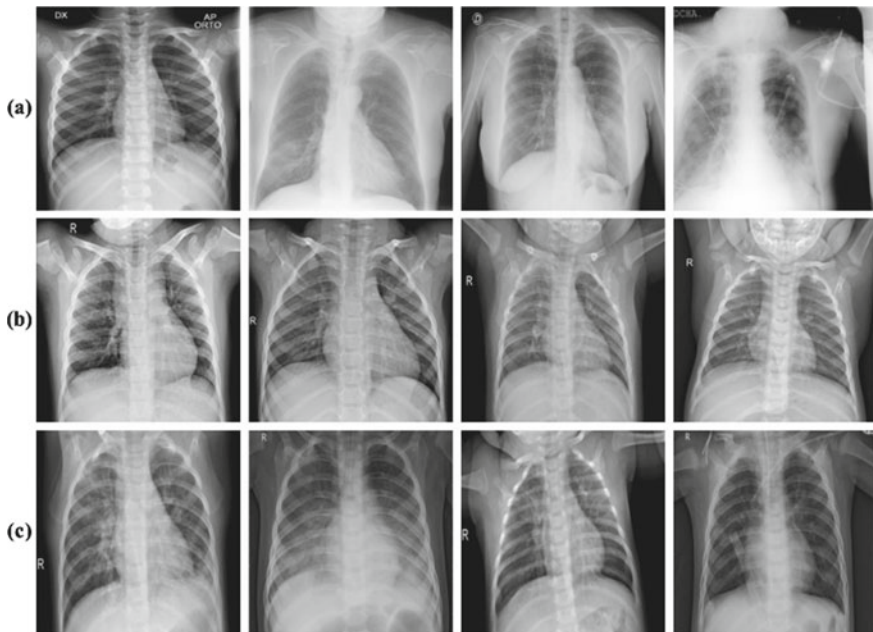


Fig. 2 Sample images **a** COVID-19, **b** Normal, and **c** Viral Pneumonia

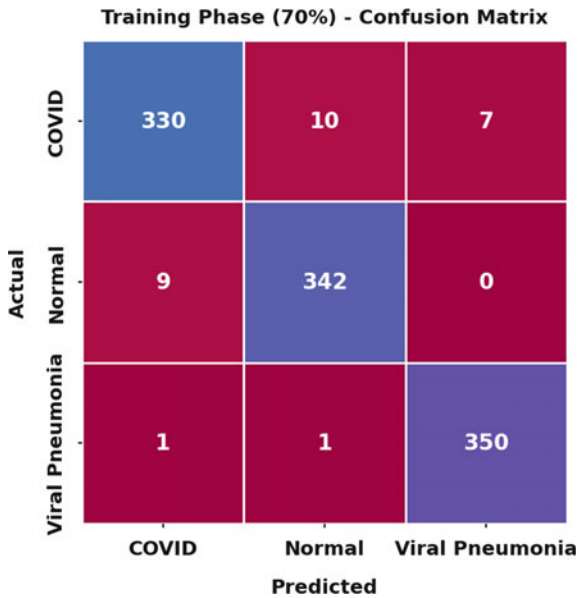


Fig. 3 Confusion matrix of AOADL-MXIC approach under 70% of TR data

95.10%, F_{score} of 96.07%, and $Jaccard_{Index}$ of 92.44%. Moreover, the AOADL-MXIC approach has recognized X-ray images under normal class with $accu_y$ of 98.10%, $prec_n$ of 96.88%, $reca_l$ of 97.44%, F_{score} of 97.16%, and $Jaccard_{Index}$ of 94.48%. Meanwhile, the AOADL-MXIC system has recognized X-ray images under viral pneumonia class with $accu_y$ of 99.14%, $prec_n$ of 98.04%, $reca_l$ of 99.43%, F_{score} of 98.73%, and $Jaccard_{Index}$ of 97.49%.

An average X-ray image classification performance of the AOADL-MXIC model under 70% of TR data is provided in Fig. 5. The figure pointed out that the AOADL-MXIC model has resulted in average $accu_y$ of 98.22%, $prec_n$ of 97.33%, $reca_l$ of 97.32%, F_{score} of 97.32%, and $Jaccard_{Index}$ of 94.80%.

Figure 6 depicts the confusion matrix generated by the AOADL-MXIC approach on 30% of testing (TS) data. The figure demonstrated that the AOADL-MXIC system

Table 2 Result analysis of AOADL-MXIC approach with distinct measures under 70% of TR data

Training phase (70%)					
Class labels	Accuracy	Precision	Recall	F-score	Jaccard index
COVID-19	97.43	97.06	95.10	96.07	92.44
Normal	98.10	96.88	97.44	97.16	94.48
Viral Pneumonia	99.14	98.04	99.43	98.73	97.49
Average	98.22	97.33	97.32	97.32	94.80

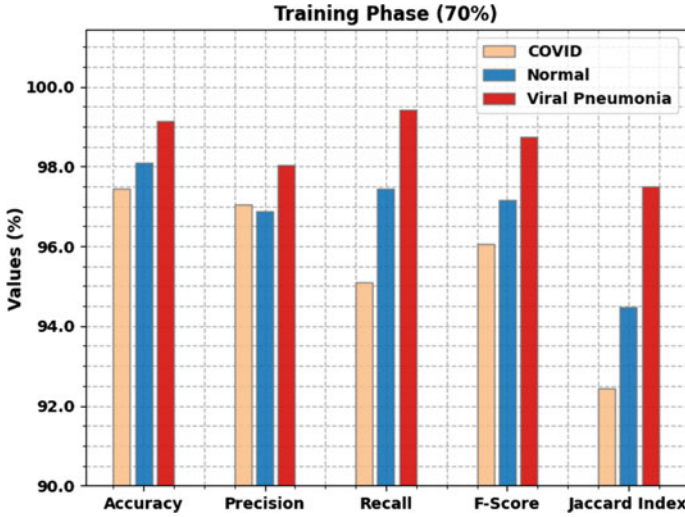


Fig. 4 Result analysis of AOADL-MXIC approach under 70% of TR data

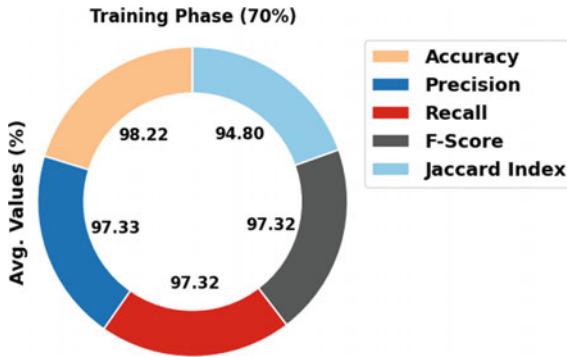


Fig. 5 Average analysis of AOADL-MXIC approach under 70% of TR data

has identified 148 samples under COVID-19 class, 146 samples under Normal class, and 148 samples under Pneumonia class.

Table 3 and Fig. 7 showcase an overall X-ray classification performance of the AOADL-MXIC algorithm under distinct classes on 30% of TS data. The experimental values demonstrated that the AOADL-MXIC technique has achieved maximal performance in each class. For example, the AOADL-MXIC approach has recognized X-ray images under COVID-19 class with $accu_y$ of 98.22%, $prec_n$ of 98.01%, $reca_l$ of 96.73%, F_{score} of 97.37%, and $Jaccard_{Index}$ of 94.87%. Furthermore, the AOADL-MXIC technique has recognized X-ray images under normal class with $accu_y$ of 98.89%, $prec_n$ of 98.65%, $reca_l$ of 97.99%, F_{score} of 98.32%, and $Jaccard_{Index}$ of 96.69%. In the meantime, the AOADL-MXIC method has recognized X-ray images

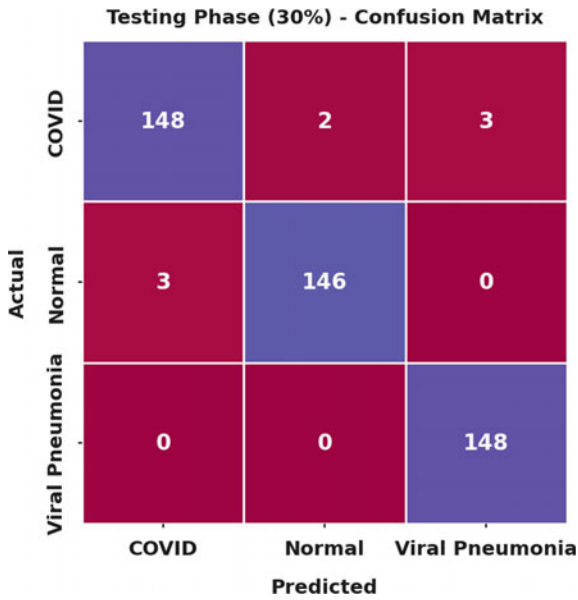


Fig. 6 Confusion matrix of AOADL-MXIC approach under 70% of TR data

under viral pneumonia class with $accu_y$ of 99.33%, $prec_n$ of 98.01%, $reca_l$ of 100%, F_{score} of 99%, and $Jaccard_{Index}$ of 98.01%.

An average X-ray image classification performance of the AOADL-MXIC approach under 30% of TS data is given in Fig. 8. The figure revealed that the AOADL-MXIC methodology has resulted in average $accu_y$ of 98.81%, $prec_n$ of 98.23%, $reca_l$ of 98.24%, F_{score} of 98.23%, and $Jaccard_{Index}$ of 96.52%.

The training accuracy (TA) and validation accuracy (VA) reached by the AOADL-MXIC approach on test dataset are established in Fig. 9. The experimental outcome exposed that the AOADL-MXIC method has gained maximal values of TA and VA. In specific, the VA performed that higher than TA.

Table 3 Result analysis of AOADL-MXIC approach with distinct measures under 30% of TS data

Testing phase (30%)					
Class labels	Accuracy	Precision	Recall	F-score	Jaccard index
COVID-19	98.22	98.01	96.73	97.37	94.87
Normal	98.89	98.65	97.99	98.32	96.69
Viral Pneumonia	99.33	98.01	100.00	99.00	98.01
Average	98.81	98.23	98.24	98.23	96.52

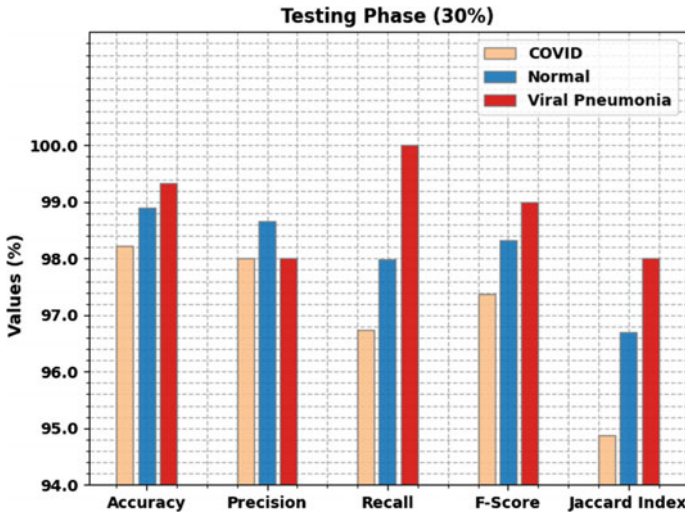


Fig. 7 Result analysis of AOADL-MXIC approach under 30% of TS data

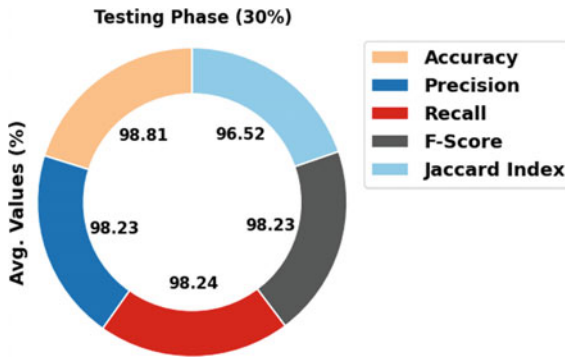


Fig. 8 Average analysis of AOADL-MXIC approach under 30% of TS data

The training loss (TL) and validation loss (VL) gained by the AOADL-MXIC system on test dataset are recognized in Fig. 10. The experimental outcome represented that the AOADL-MXIC methodology has been able minimal values of TL and VL. In specific, the VL appeared that lesser than TL.

At the final stage, a detailed comparative examination of the AOADL-MXIC approach with other DL algorithms takes place in Table 4 and Fig. 11 [21, 22]. The experimental values demonstrated that the VGG19 and SqueezeNet models have shown lower classification performance. At the same time, the TL_SqueezeNet and ResNet-1 models have demonstrated slightly enhanced classifier results. Moreover, the TL-ResNet-2 model has exhibited reasonable classification performance with acc_y of 98.44%, $prec_n$ of 97.97%, $reca_l$ of 97.94%, and F_{score} of 98.02%. But the



Fig. 9 TA and VA analysis of AOADL-MXIC algorithm

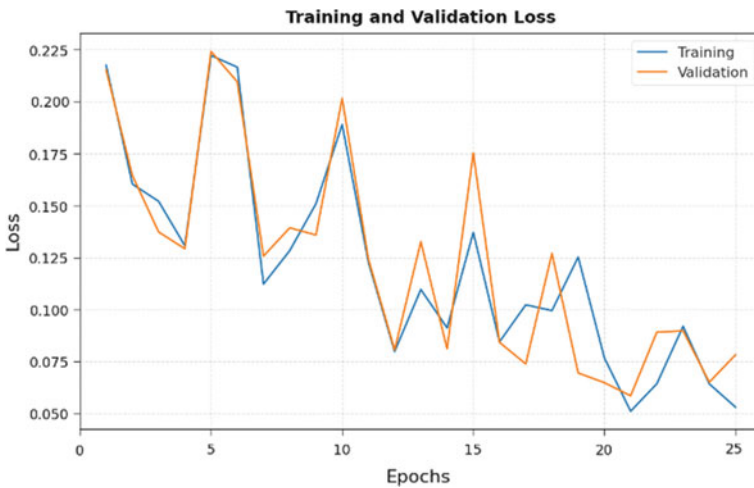


Fig. 10 TL and VL analysis of AOADL-MXIC algorithm

AOADL-MXIC model has attained superior results with $accu_y$ of 98.81%, $prec_n$ of 98.23%, $reca_f$ of 98.24%, and F_{score} of 98.23%. From the detailed results and discussion, it is confirmed that the AOADL-MXIC model has accomplished maximum performance on X-ray image classification process.

Table 4 Comparative analysis of AOADL-MXIC approach with existing methodologies

Methods	Accuracy	Precision	Recall	F-score
Squeeze net	92.96	91.73	95.10	93.46
TL_Squeeze net	96.44	96.08	96.84	96.59
VGG19	93.30	91.78	94.73	93.44
TL-ResNet-2	98.44	97.97	97.94	98.02
ResNet-1	97.23	96.74	97.43	96.79
AOADL-MXIC	98.81	98.23	98.24	98.23

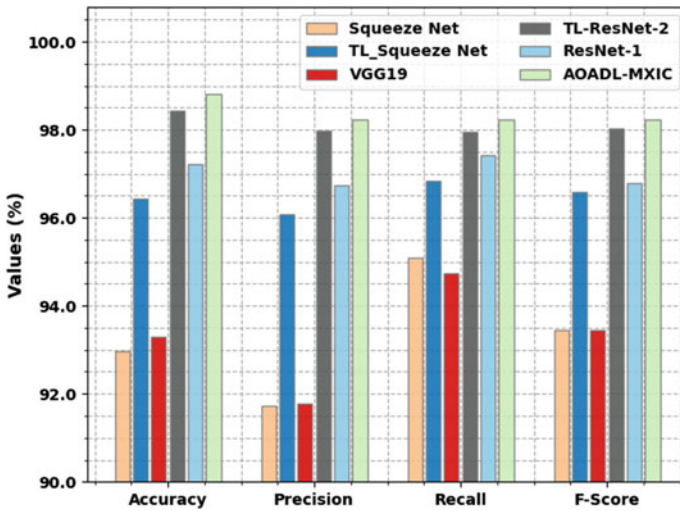


Fig. 11 Comparative analysis of AOADL-MXIC approach with existing methodologies

5 Conclusion

In this study, a novel AOADL-MXIC approach was introduced for the effectual detection and classification of medical X-ray images. The presented AOADL-MXIC approach primarily applied GF approach to pre-process the images. Next, the feature extraction process was executed by the use of AOA with CapsNet model. At last, the BiGRU approach was utilized for the classification of medical X-ray images. The utilization of the AOA helps for optimizing the hyperparameters related to the CapsNet system, the AOA is exploited. For validating the performance of the AOADL-MXIC model, a wide range of experiments is executed. The experimental result analysis of AOADL-MXIC approach on a set of medical images stated the promising performance over the other models. In future, the proposed technique was extended to the design of hybrid DL techniques for enhanced classification performance.

References

1. Mall PK, Singh PK, Yadav D (2019) GLCM based feature extraction and medical X-RAY image classification using machine learning techniques. In: 2019 IEEE conference on information and communication technology. IEEE, pp 1–6
2. Das A, Mohapatra SK, Mohanty MN (2022) Design of deep ensemble classifier with fuzzy decision method for biomedical image classification. *Appl Soft Comput* 115:108178
3. Chen J-Z (2021) Design of accurate classification of COVID-19 disease in X-Ray images using deep learning approach. *J ISMAC* 3(02):132–148
4. Sunghheetha A, Rajesh Sharma R (2021) Classification of remote sensing image scenes using double feature extraction hybrid deep learning approach. *J Inf Technol* 3(02):133–149
5. Mall PK, Singh PK (2022) BoostNet: a method to enhance the performance of deep learning model on musculoskeletal radiographs X-ray images. *Int J Syst Assur Eng Manag* 1–15
6. Chakraborty S, Paul S, Hasan KM (2022) A transfer learning-based approach with deep CNN for COVID-19-and pneumonia-affected chest X-ray image classification. *SN Comput Sci* 3(1):1–10
7. Sharma A, Singh K, Koundal D (2022) A novel fusion based convolutional neural network approach for classification of COVID-19 from chest X-ray images. *Biomed Sig Process Control* 103778
8. Masud M (2022) A light-weight convolutional Neural Network Architecture for classification of COVID-19 chest X-Ray images. *Multimedia Syst* 1–10
9. Kc K, Yin Z, Wu M, Wu Z (2021) Evaluation of deep learning-based approaches for COVID-19 classification based on chest X-ray images. *SIVIP* 15(5):959–966
10. Mahdy LN, Ezzat KA, Elmousalami HH, Ella HA, Hassaniien AE (2020) Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine. *MedRxiv*
11. Gour M, Jain S (2022) Uncertainty-aware convolutional neural network for COVID-19 X-ray images classification. *Comput Biol Med* 140:105047
12. Kong L, Cheng J (2022) Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion. *Biomed Signal Process Control* 77:103772
13. Bhattacharyya A, Bhaik D, Kumar S, Thakur P, Sharma R, Pachori RB (2022) A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images. *Biomed Signal Proc Control* 71:103182
14. Munusamy H, Muthukumar KJ, Gnanaprakasam S, Shanmugakani TR, Sekar A (2021) FractalCovNet architecture for COVID-19 Chest X-ray image classification and CT-scan image segmentation. *Biocybernetics Biomed Eng* 41(3):1025–1038
15. Chen B, Li J, Lu G, Yu H, Zhang D (2020) Label co-occurrence learning with graph convolutional networks for multi-label chest x-ray image classification. *IEEE J Biomed Health Inform* 24(8):2292–2302
16. Li HA, Fan J, Yu K, Qi X, Wen Z, Hua Q, Zhang M, Zheng Q (2020) Medical image coloring based on gabor filtering for internet of medical things. *IEEE Access* 8:104016–104025
17. Verma S, Jahangir S, Chug A, Singh RP, Singh AP, Singh D (2022) SE-CapsNet: automated evaluation of plant disease severity based on feature extraction through Squeeze and Excitation (SE) networks and Capsule networks. *Kuwait J Sci* 49(1)
18. Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021) The arithmetic optimization algorithm. *Comput Methods Appl Mech Eng* 376:113609
19. Luo X, Zhou W, Wang W, Zhu Y, Deng J (2017) Attention-based relation extraction with bidirectional gated recurrent unit and highway network in the analysis of geological data. *IEEE Access* 6:5705–5715
20. <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>

21. Ragab M, Alshehri S, Alhakamy NA, Alsaggaf W, Alhadrami HA, Alyami J (2022) Machine learning with quantum seagull optimization model for COVID-19 Chest X-Ray image classification. *J Healthc Eng* 2022
22. Khan SH, Sohail A, Khan A, Hassan M, Lee YS, Alam J, Basit A, Zubair S (2021) COVID-19 detection in chest X-ray images using deep boosted hybrid learning. *Comput Biol Med* 137:104816

Lightweight Encoder-Decoder Model for Semantic Segmentation of Hand Postures



Siddhant Kumar Das, Rishabh Lahkar, K. Antariksha, Ayanmani Das, Aryamaan Bora, and Amrita Ganguly

Abstract Hand gesture recognition (HGR) is becoming more important in human-computer interaction as people's expectations for interactive experiences rise. It has a wide range of applications in the automobile industry, consumer electronics, home automation, and other areas. Proposal of an alternative for gesture recognition without extra sensor requirement of depth images, by using semantic segmentation images of the hand for real-time hand gesture recognition is done. Segmentation is a challenging problem by itself, due to different restrictions such as illumination variation, a complicated background, and so on. The use of an encoder-decoder architecture for pixel wise semantic segmentation is done which is conducive for real-time application with a segmentation accuracy of 83.57%. A comparison is drawn between the performance of different existing categorization models as encoder or discriminator to attain a balance of efficiency and accuracy as well as observed the performance upon fine tuning on pretrained encoders on ImageNet to achieve the task of segmenting hand postures. The results of applying semantic segmentation for HGR have been demonstrated, highlighting the proposed model's computational efficacy of 87.8 GFLOPS.

Keywords Deep neural network · Convolutional neural network · Hand gesture recognition · Human computer interaction

1 Introduction

Hand Gesture Recognition has been a challenge researchers have been trying to solve for decades. Many traditional models were proposed for this challenge but development was quite slow in this regard due to unavailability of resources and impractical solutions. With the progress of computer technology, better computing with improved handling of large data was possible. This led to increase in use of convolutional neural networks, because of that deep neural networks also emerged

S. K. Das (✉) · R. Lahkar · K. Antariksha · A. Das · A. Bora · A. Ganguly
Assam Engineering College, AEC Road, Jalukbari, Guwahati, Assam 781013, India
e-mail: siddhantkd@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_41

579

which was lying dormant for a very long time. A major advantage of deep neural network models for gesture recognition was that it eliminated the use of peripherals, making hand gesture recognition possible only with images and videos. Hand Gesture Recognition models have one aim, which is to recognize the different postures and gestures made by the human hand. This is achieved by semantic segmentation of the input image by the model for the creation of the feature extraction map from which the model differentiates between the background and the object, and the model recognizes the object, i.e., the hand, and particularly the gesture of the hand. The main challenges of hand gesture recognition are skin color detection, complex background removal and variable lighting conditions. Researchers have tried to tackle these problems through proper training of their models using various datasets. In this paper, the objective to create a lighter, faster and accurate deep neural network model is achieved which is trained and optimized for real-time hand gesture recognition after various experimentations and observations. The model does semantic segmentation of an RYB-image which eliminates the need for extra peripherals like gloves or depth cameras in real-time. The experimentations included creation of different models using various kinds of convolutional, pooling and sampling blocks in different orders to get the best results from which the best model was chosen using various performance parameters. A comparative study of performance between the proposed model with existing models is described in this paper. The work has been documented in the upcoming sections.

2 Related Works

The problem of Hand Gesture Recognition (HGR) is as old as Human Computer Interaction (HCI) Systems themselves. It is one of the most challenging tasks related to HCI. Due to differences in illumination, background, shadows, differences in ethnicities, etc. the task becomes even more complicated. Hand gestures have a wide use case. They are used to convey extra information during conversations, like when pointing at an object, the index finger is used, and also, they help add structure and emphasis to the words being spoken, by blind, deaf, and dumb people as a means of communication.

From [1], Bhuyan et al showed the raw data acquisition of hand gestures as either sensor-based or vision-based. Vision-based gestures [2] have become more popular due to the limitations of sensor-based approaches, like requirement of gloves, which are expensive devices. The gestures have been obtained by the use of special cameras. It also has to be noted that the majority of problems related to hand gesture segmentation, like variations in illumination, background, occlusion, etc. can be solved by the use of depth cameras for capturing image RGB-D data. But, the feasibility and cost [3] of depth cameras make them a rather unpopular choice.

With the advancement of powerful graphics processing units [4], and recently Tensor Processing Units (TPUs) [5], deep neural networks have become the standard

for image and video related problems. Raina [4] was also the first paper that showed the enormous advantages of using GPUs over multicore CPUs.

AlexNet [6] was the first model which incorporated deep neural networks in this challenge. Furthermore, in [7], residual networks are introduced. These differ from plain networks. These ResNets, for short, have shortcut skip connections connecting different layers. The ResNets perform much better than the networks without skip connections. They are easily optimizable and have a far lower error rate. In [8], DeepLab network is introduced. The authors have used Atrous Spatial Pyramid Pooling (ASPP) to increase the receptive field of the network.

In [9], Dadashzadeh et al. proposed a hand gesture recognition model incorporating [7, 8]. Two streams constitute the network architecture, the first named the shape stream, while the second is named the appearance stream. The gesture is recognized by fusion of these two streams.

Next, the encoder-decoder module is introduced in [10]. The network is called UNet. The encoder module consists of a contracting path which captures context, while the decoder module is made up of a symmetric expanding path which enables exact localization. In [11], Attention Gates (AG) were introduced. These gates were easily integrated into the UNet architecture, creating Attention UNet. The advantage of using AGs were that it helped suppress irrelevant information while highlighting important parts of an image.

To tackle the problem of dynamic gesture recognition, MobileNetV2 is introduced in [12]. Depthwise Separable Convolution, Residual Connections with Linear Bottlenecks, and an Inverted Residual Structure, are the main features used in this model. These novelties make MobileNetV2 an effective option for use in scenarios where resources are limited.

For efficient semantic segmentation in real-time, LinkNet [13], was developed. In LinkNet, the spatial information from the encoder to the decoder is directly bypassed. This decreases the network's processing time while also increasing accuracy.

In [14], the author compared some well known loss functions and summarized their convergence rate in a model employed for the task of image segmentation.

3 Methodology

3.1 Overview

The task of segmenting region of interest (hand postures) from a given RGB input frame using deep neural network methods can be categorized into three extensive steps.

- Step 1: Collection and preprocessing of dataset.
- Step 2: Design and training of semantic segmentation deep learning model.
- Step 3: Proposed deep learning model statistical evaluation and output generation.

3.2 Dataset

The dataset used in this work is the “OUHANDS” dataset. The “OUHANDS” dataset is a publicly available collection of static hand posture images captured while a user is demonstrating gesture commands to a handheld device [15]. The authors used “The Intel RealSense F200” camera to collect the dataset samples. The dataset consists of 3150 hand samples (2150 training samples and 1000 test samples). The resolution for all the images is $640 * 480$ pixels. The OUHANDS dataset comprises of 10 unique hand postures demonstrated by 23 people. The dataset is captured in Human Computer Interaction (HCI)-like settings, such as handheld camera position, uncontrolled backgrounds, variation in illumination, hand-face obstruction with different hand shapes and sizes. This makes this dataset conducive for training and testing Human Computer Interaction (HCI) methods.

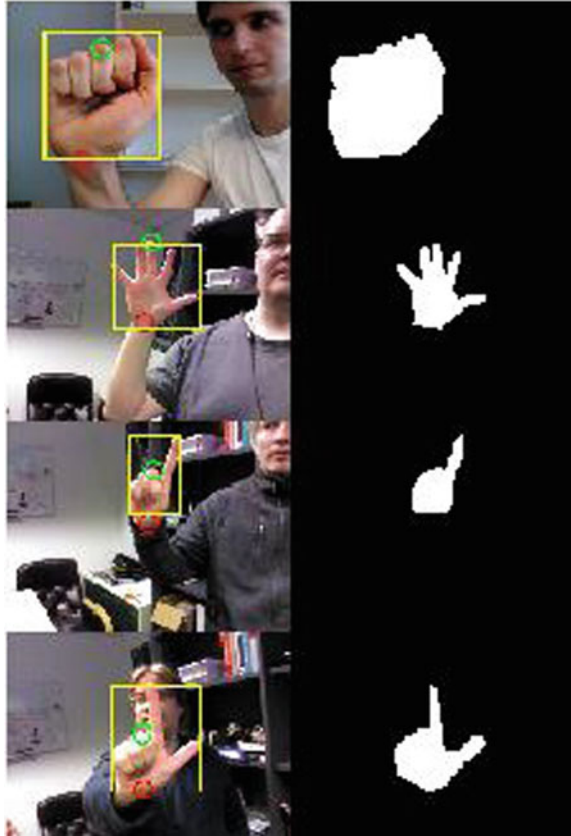
Data Preprocessing: The dataset samples were originally of the size $640 * 480$ pixels. This was resized into $320 * 320$ pixels for training the semantic segmentation model. The ground truth data which were the segmentation masks were binary thresholded before feeding into the models (Fig. 1).

3.3 Network Architecture

The network comprises of a U-shaped decoupled encoder-decoder structure similar to the architecture of U-NET [10], The network consists of the encoding or contracting path, the decoding or expanding path and there is a section of layers that link the two paths to minimize loss of spatial information during the decoding stage. The architecture of the proposed model for segmenting hand regions is represented in Fig. 2. The encoder is the narrowing contracting path of the architecture which encodes the input features from the given RGB frame. Lightweight MobileNetV2 [12] feature extraction model was employed to modify the encoder. The feature extractor was pretrained on ImageNet in order to benefit from regularization induced by transfer learning. The decoder is the expanding path of the architecture whose goal is to project the extracted lower dimensional features by the encoder into a higher resolution spatial dimension pixel space by multiple up-scaling operations. The pixel resolution of the final output map is $(320 * 320)$, identical to the original input RGB image. The decoder architecture was adopted from LinkNet [13] network architecture which utilizes the skip connections in an efficient manner conducive for real-time application. The skip connections in the proposed model architecture links the encoder and decoder at four points.

Encoder: To encode the semantic information from the given hand pose image, the model employs the MobileNetV2 [12] model for contextual feature extraction. The architecture for MobileNetV2 feature extraction model is shown in Fig. 3.

Fig. 1 OUHANDS dataset samples



For end-to-end segmentation tasks, the fully connected layers of the architecture are omitted. MobileNetV2 uses depth-separable convolution, residual links with linear bottlenecks, and an inverted residual structure.

Depth-separable convolution is used extensively in real-time tasks for two reasons

- (i) Compared to the classic convolution, fewer parameters need to be tuned, minimizing the likelihood of overfitting the model.
- (ii) Fewer calculations are performed, making them computationally efficient and more conducive for real-time applications.

Depthwise convolution constitutes convolution blocks called “inverted residual blocks” in MobileNetV2. There are three layers in each block. The first layer in each block is a 1×1 convolution layer with the ReLU6 activation function, which expands the number of feature channels. The second layer is a 3×3 depthwise convolution layer. The third 1×1 convolution layer compresses the network back to the original number of channels. Inverted residual blocks squeeze the layers where

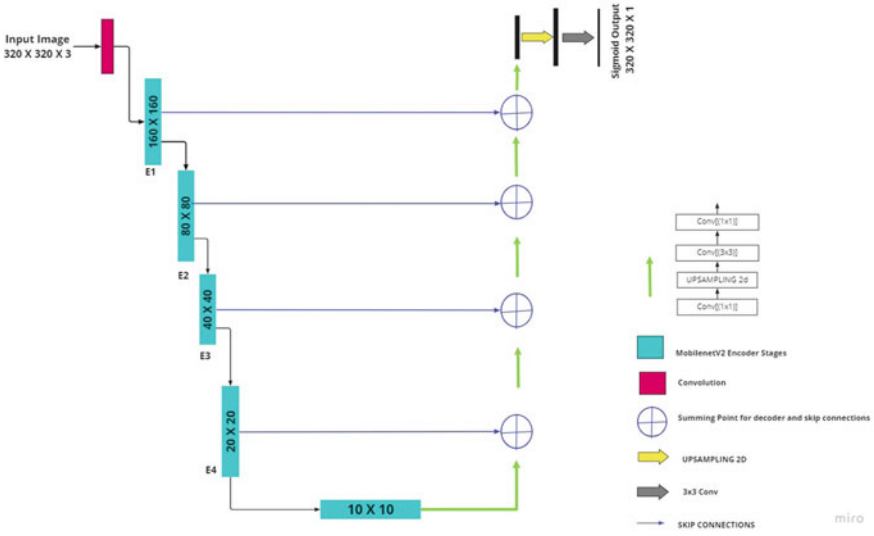


Fig. 2 Model architecture

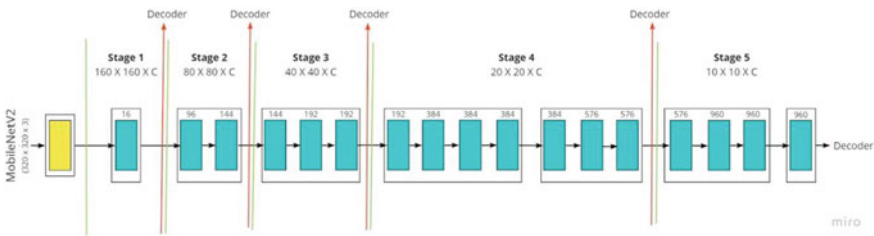


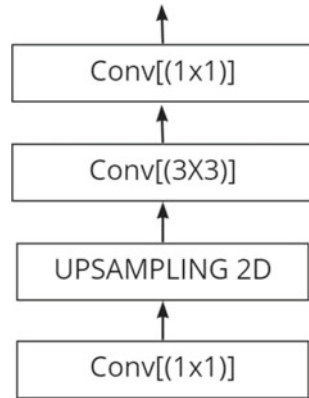
Fig. 3 MobileNetV2 architecture

skip connections are linked, degrading network performance. To overcome this, the authors introduced the linear bottleneck architecture, where the last convolution of a residual block has a linear output before adding it to the initial activations.

The input to the encoder is a RGB image with $(320 \times 320 \times 3)$ dimensions. Feature maps are extracted by the encoder blocks and are repeatedly downsampled during the process. The highest level feature map obtained during the encoding stage is of the dimension $(10 \times 10 \times 1280)$. This layer acts as the bridge between encoder and decoder.

Skip Connections: Skip connections link the higher level layer from the deeper part of the architecture to the lower level layer of the decoder for fine-grained feature re-usability during the up-scaling operations. The deeper layers of the network from the encoder have rich feature information. Skip connections make the model more robust by combining these rich feature maps to the decoder blocks for effective localization.

Fig. 4 Decoder block structure



Decoder: In the decoder, the expanding mechanism of the Link-NET architecture is employed. Link-NET is a fully convolutional network, structured in a U-shape similar to the U-NET. The Link-NET varies from the U-NET by the mechanism by which it links each encoder with its corresponding decoder block through the skip connections. Unlike U-NET, the decoder of the Link-NET does not concatenate the feature map extracted from the lower level encoder path and the higher level feature map from the deeper layers. The feature maps are added directly (through a convolutional layer) to reduce the computational parameters. The decoder produces segmentation maps at full resolution. The expanding structure constitutes four decoder blocks stacked together each linked to its corresponding encoder block through a summing point and skip connection. The decoder block structure is shown in Fig. 4. Each block has three layers. First layer is a (1×1) convolution layer with batch normalization and RELU activation. Second layer is an upsampling layer which upsamples the feature map using bilinear interpolation. It is then followed by a two convolution layers. The final decoder block of the expansion path is followed by (3×3) convolution layer with the sigmoid activation which outputs the required $(320 \times 320 \times 1)$ segmentation mask of region of interest (hand postures).

3.4 Training

The training data was used to train six different models. Each of the six models was subjected to identical training. Images having a resolution of 640×480 pixels were included in the training data which were then scaled to a dimension of 320×320 pixels and fed into the network to train the model. While experimenting with various batch sizes, it was experimentally determined that batch size of 16 provided the optimal learning pattern. “Adam” was used as the optimizer to train the deep learning model. According to Diederik [16], the method is “computationally efficient,

has little memory requirement, invariant to diagonal rescaling of gradients, and is well suited for problems that are large in terms of data/parameters”.

The set argument values for the optimizer used for training the models:

- Learning Rate: 0.001
- Exponential Decay Rate for the 1st moment estimates (1): 0.9
- Exponential Decay Rate for the 2nd moment estimates (2): 0.999
- Epsilon value: $1e - 4$.

All the models were trained for 500 epochs.

- **Loss Function:** After experimenting with many loss functions it was decided to use the dice coefficient loss [14] for the task of semantic segmentation.

$$D = \frac{2 \sum_i^N y_i x_i}{\sum_i^N y_i^2 + \sum_i^N x_i^2} \quad (1)$$

Here, y_i and x_i are a representation of pixel values of corresponding predicted mask and the ground truth mask, respectively. The values of y_i and x_i are either 0 or 1. The denominator in the dice loss equation represents the sum of total pixels from both the predicted mask and the ground truth mask. The numerator is a representation of successfully predicted pixels, since the sum only advances when y_i and x_i match (both of value 1). If two masks completely overlap or the predicted output is accurate, the Dice Coefficient gets its maximum value to 1. Otherwise, the Dice coefficient starts to decrease, getting to its minimum value to 0 if the two masks don't overlap at all or the predicted output is totally inaccurate.

- **Intersection Over Union metric (IoU):** To quantitatively evaluate the model performance, the IoU metric commonly known as the Jaccard Index is used. It is a commonly used metric for evaluating segmentation performance. It measures the percentage overlap between the ground truth mask and the predicted mask. The value of the IoU score is in the range of 1–0. A good prediction will result in a higher IoU score whereas a poor prediction will result in a lower IoU score. In mathematical terms, it is defined as the number of pixels that are common between the ground truth mask and predicted mask divided by the total number of pixels present in both of the masks,

$$IoU = \frac{(\text{Ground Truth Mask Pixels}) \cap (\text{Predicted Mask Pixels})}{(\text{Ground Truth Mask Pixels}) \cup (\text{Predicted Mask Pixels})} \quad (2)$$

Mean IoU is the average IoU score of all the individual image segmentation in the dataset. This is used to evaluate and compare the various models.

For deep learning implementations, Keras and Tensorflow frameworks were used. The system was configured with the NVIDIA P100 GPU (Fig. 5).

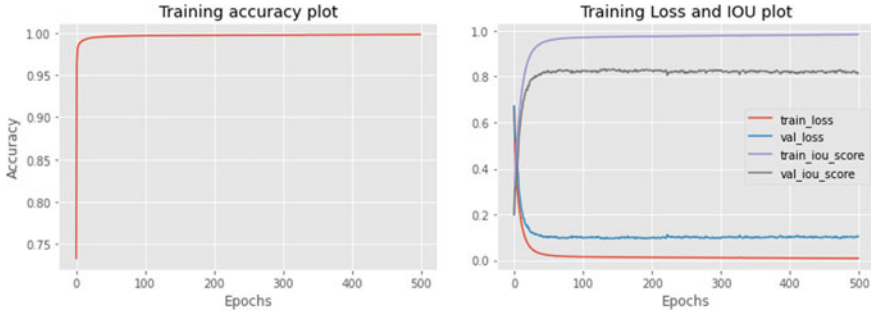


Fig. 5 Training plots

3.5 Results and Discussion

Training of the Networks The training of the proposed model took approximately 450 h, i.e., 54 s per epoch. Evaluation of the models during the training phase was based on the convergence of the Intersection Over Union score. Table 1 summarizes the convergence of the proposed model during the training. With a low IoU score at the first epoch, the IoU score increased rapidly reaching a score of 0.6819 after 10 epochs. The rate of convergence decreases significantly after 100 epochs. After 500 epochs of training, the model achieved an IoU score of 0.9839. The dice coefficient loss function decreases at a high rate initially but the rate slows down at higher epochs. After 300 epochs, the rate of convergence of the loss function reduces greatly and by the end of the 500 epochs, the model recorded a loss function value of 0.0081. The Training and Validation IoU score and Loss function plot is shown in fig 10 along with the training accuracy plot.

Table 1 Convergence of training IoU score and loss of the proposed methods

Epoch	Proposed MobileNetv2 LinkNet	
	Training IoU	Training loss
1	0.2045	0.6600
10	0.6819	0.1897
30	0.9200	0.0417
50	0.9566	0.0222
100	0.9710	0.0147
200	0.9759	0.0122
300	0.9787	0.0108
400	0.9813	0.0094
500	0.9839	0.0081

Results and Discussion: For quantitative comparison of the accuracy of the proposed model, Mean Intersection Over Union (mIoU), commonly known as the Jaccard Index is measured on the Ouhands Test dataset. In the beginning, to determine the efficacy of using a pretrained mobileNetV2 encoder instead of a non-pretrained mobileNetV2 encoder, the proposed model is trained with ImageNet pretrained weights and was compared with their non-pretrained versions in Table 1. It can be seen that the mIoU percentage of the pretrained model has a significant difference from the non-pretrained model. Thus, the ImageNet pretrained mobileNetV2 encoder was proceeded to be used in the proposed model (Fig. 6; Table 2).

For the demonstration of the performance of the proposed model the difference between the proposed model and other segmentation models are compared and analyzed. The original architecture of three state-of-the-art segmentation models, HGR-Net, U-Net and Google's Deeplabv3+ (with ResNet-50 feature extractor) have been implemented. In Table 3, it is observed that the proposed model is much superior compared to HGR-Net and U-Net. When compared with Deeplabv3+, it can be seen that the proposed model with significantly lesser parameters has similar mIoU percentage. A qualitative comparison of the segmentation result on the OUHANDS dataset are illustrated. It shows that the proposed segmentation model performs efficiently in different environments, i.e., in different backgrounds and different lighting conditions.

Further, the number of parameters and GFLOPs (floating-point operations per second) of the proposed model is compared with those of other state-of-the-art models in Table 4. HGR-Net has the lowest number of model parameters and GFLOPS, but at the same time, experimentation has shown it has the lowest segmentation accuracy compared to all other experimented models. The number of parameters of the proposed model is significantly less than that of UNet and Deeplabv3+. The GFLOPS of the proposed model accounts for only 16.9 and 3.64% of the GFLOPS of DeeplabV3+ and U-Net, respectively. Testing the model locally run on an AMD Ryzen 7 4800H CPU gave an average inference time of 0.10580351 s, a 0.1 s difference compared to the state-of-the-art DeeplabV3+ model with an inference time of 0.20978545 with almost the same segmentation performance. Thus, MobileNetv2 Linknet with lower computation cost and model size is a fast and efficient segmentation model for mobile devices.

Conclusion

Real-time performance of semantic segmentation in mobile devices is greatly restricted by the limited processing power of the hardware installed in such devices, used extensively in the field of hand gesture recognition systems. The existing work in this field involves a trade-off between segmentation accuracy and speed, the latter being more important for real-time application. Thus, there is a requirement for a lightweight model that can maintain a good balance between speed and accuracy for optimum performance. The model proposed in this paper adopts a MobileNetv2

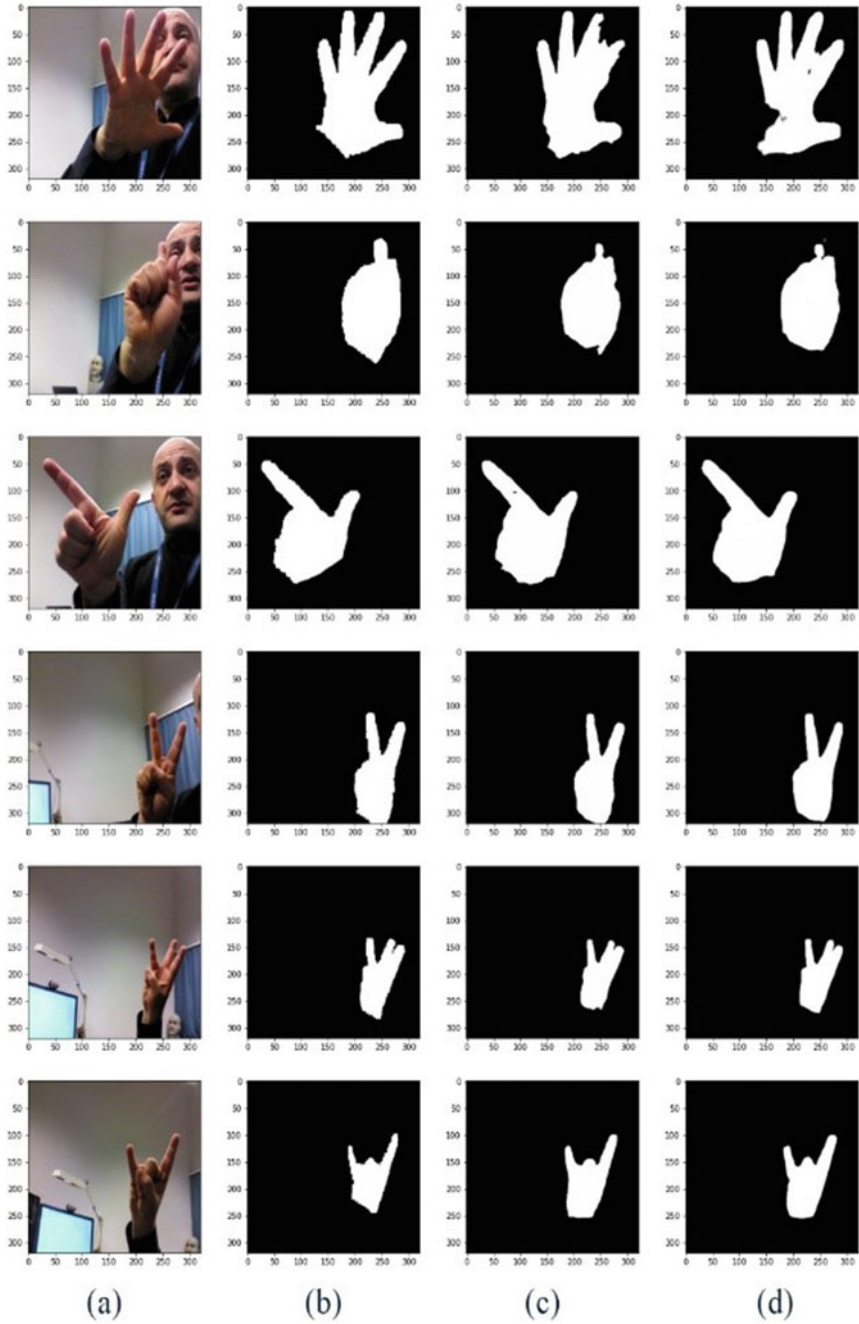


Fig. 6 Example segmentations on OUHANDS test set. **a** Input images. **b** Ground truth segmentation mask. **c** Mobilenetv2 Linknet. **d** Deeplabv3+

Table 2 Performance of models with or without ImageNet pretrained encoders

Model	mIOU (%)
Non-pretrained Mobilenetv2 Linknet	71.33
ImageNet pretrained Mobilenetv2 Linknet	83.57

Table 3 Performance comparison with other segmentation methods

Model	Total parameters	mIOU (%)
HGR-Net [9]	279,633	72.06
U-Net [10]	31,055,297	77.82
Deeplabv3+ [8]	11,852,353	83.87
Proposed Mobilenetv2 Linknet	4,146,617	83.57

Table 4 Parameters and GFLOPS comparison of the methods

Model	Total parameters	GFLOPS
HGR-Net [9]	279,633	30
U-Net [10]	31,055,297	2410
Deeplabv3+ [8]	11,852,353	519
Proposed Mobilenetv2 Linknet	4,146,617	87.8

backbone pretrained on the ImageNet dataset and the decoder architecture of LinkNet using lateral skip connections to reduce model parameters and computation cost. To verify the overall efficacy of the models, the model was compared with state-of-the-art model on the OUHANDS benchmark dataset and confirmed that the proposed model, MobileNetv2 LinkNet with 4.15 million parameters and 87.8 G floating-point operations, strikes the best balance between accuracy and speed, achieving a mean IoU of 83.57%.

Limitations: The main limitations of the study involve the ability to segment video data and the amount and variety of data available on different types of hand gestures. Temporal information needs to be considered for the model to perform efficiently on video data. The model is not able to properly segment two hand gestures at the same time. It was also observed that the skin color of the hand affects the segmentation performance of the model. The collection of more data will help enhance the performance of the model, thereby addressing these limitations.

Future Scope: The proposed model focuses on semantic segmentation of hand postures which is a pre-requisite for real-time recognition of hand gestures in complex environment using RGB cameras. The project can be extended to build a classification model with low computational complexity conducive for real-time application which can classify and recognize the hand gestures given by the user. Also future research can be carried out to reduce the model parameters and achieve a trade-off between accuracy and model size which will make it more feasible to employ in real-time application which can run on embedded devices.

References

1. Debajit S, Bhuyan MK (2021) Methods, databases and recent advancement of vision-based hand gesture recognition for hci systems: a review. *SN Comput Sci* 2(6):1–40
2. Liao C-J, Su S-F, Chen M-C (2015) Vision-based hand gesture recognition system for a dynamic and complicated environment. In: 2015 IEEE international conference on systems, man, and cybernetics, pp 2891–2895
3. Jani AB, Kotak NA, Roy AK (2018) Sensor based hand gesture recognition system for english alphabets used in sign language of deaf-mute people. *IEEE Sens* 1–4
4. Raina R, Madhavan A, Ng AY (2009) Large-scale deep unsupervised learning using graphics processors. In: Proceedings of the 26th annual international conference on machine learning, pp 873–880
5. Jouppi N, Young C, Patil N, Patterson D (2018) Motivation for and evaluation of the first tensor processing unit. *IEEE Micro* 38(3):10–19
6. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Proces Syst* 25
7. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
8. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans Pattern Anal Machine Intel* 40(4):834–848
9. Dadashzadeh A, Targhi AT, Tahmasbi M, Mirmehdi M (2019) Hgr-net: a fusion network for hand gesture segmentation and recognition. *IET Comput Vis* 13(8):700–707
10. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, pp 234–241. Springer
11. Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B et al (2018) Attention u-net: learning where to look for the pancreas. arXiv preprint [arXiv:1804.03999](https://arxiv.org/abs/1804.03999)
12. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
13. Chaurasia A, Culurciello E (2017) Linknet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE visual communications and image processing (VCIP), pp 1–4. IEEE
14. Jadon S (2020) A survey of loss functions for semantic segmentation. In: 2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB). IEEE
15. Matilainen M, Sangi P, Holappa J, Silvén O (2016) Ouhands database for hand detection and pose recognition. In: 2016 sixth international conference on image processing theory, tools and applications (IPTA), pp 1–5. IEEE
16. Diederik PK, Ba J (2014) Adam: a method for stochastic optimization

Flexi-Pay System for Public Transport Based on Passenger Comfort Using Global Positioning System



G. Mageshkumar , S. Suthagar , D. Maher Shalal, S. Muruganantham, and T. Manoji

Abstract Bus is a mode of transportation for people around the world. Bus fare is calculated based on the distance travelled by the passenger. The fare collected from both sitting and standing passenger is the same, but they do not enjoy the same comfort. This paper proposes a flexi-pay system, that is, a variable fare calculation system considering the distance travelled by the passenger whether he/she is seated or standing during the journey. The system uses ultrasonic sensor to find seating or standing passenger. The radio frequency identity is used to scan whenever the passenger gets a seat or leaves it. Global positioning system is used to calculate the distance travelled. Then, fare is calculated based on the distance travelled by the passenger considering seating and standing time into account. A Web application is designed to manage the system with admin and user accounts. The proposed work justifies the difference in comfort of the passengers in the existing system.

Keywords Bus fare · Global positioning system (GPS) · Public transport · Radio frequency identity (RFID) · Ultrasonic sensor

1 Introduction

The growth of the nation to a developed nation is achieved by a well-connected public transport system. Public transport gives access to jobs, education, medical care, markets, social and recreational activities, and transportation of agricultural products for sale in towns and cities, amongst other things for the people. Furthermore, public transportation is a mode of local transportation that allows a group of people to travel in a planned route. People continue to find public transportation to be a far more efficient and advantageous mode of transportation. Travel automation can improve user experience and encourage use of public transport [4].

G. Mageshkumar (✉) · S. Suthagar · D. Maher Shalal · S. Muruganantham · T. Manoji
Kongu Engineering College, Perundurai, Erode 638060, India
e-mail: magesh.me.est@gmail.com

1.1 Fare Calculation

The greatest number of passengers that may be transported past a single point on a particular route in a given amount of time is referred to as capacity in transit operations. Passengers per hour are the most prevalent unit of capacity measurement. The headway (or time headway) between buses passing a certain site is measured in minutes in transit operations. The frequency is the opposite of the headway, recording the number of buses per unit time passing through a specific place on a route, commonly measured in buses per hour [18].

$$C_{\text{route}} = f * C_{\text{bus}} \quad (1)$$

where, C_{route} is the capacity in transit, f is the frequency on the route (bus per hour) and C_{bus} is the maximum number of persons per bus. The capacity in transit is calculated using Eq. 1.

1.2 Concessions in Fare

Concessions on public transportation tickets are available for the following groups of people. They include current and former members of Parliament, the Legislative Assembly, and the Legislative Council, as well as accredited journalists, media persons, women passengers, transgender, differently abled persons with a disability of 40% or more, as well as escorts, cancer patients, freedom fighters drawing central or state pensions, and widows and legal heirs of freedom fighters drawing central or state pensions, aged Tamil scholars, and participants in language protests [18].

2 Literature Review

Shanmugapriya et al. had provided a crucial solution to the urban bus passenger's plight [17]. This smart bus system uses a display board to show passengers the number of available seats and space, which is denoted by red (full) and green (empty) LED lights. Bus riders can take their tickets by scanning their RFID tags, and a softcopy will be made and delivered to the user's phone. Lavanya et al. [10] offered real-time data on bus arrival times, crowd density, and bus transmission data. All of this informs passengers about the current occupancy and future buses, allowing them to make better decisions that aid fleet management. This data also aids bus operators in analyzing patterns in public transportation utilization along various routes, allowing them to provide more services and amenities. Dhanasekar et al. [7] proposed a method that intends to improve student safety whilst travelling to and from col-

lege regularly. The student's RFID tags are detected by an RFID reader positioned inside the bus. It delivers an immediate notice to parents through the Internet with the appropriate data from the college database server. Shaikh et al. [15] introduced a smart public transportation system prototype that is designed and implemented using IoT technology, GPS, and the ESP32 microcontroller. This system demonstrates that it can only use GPS-data to obtain real-time bus information, such as the bus's current location, speed, arrival time, and distance. Shaik Saidu Masthan [16], the papers have examined, and many techniques of benefit have been investigated. Using this computerized bus ticketing system, cash collection difficulties will be avoided. It is observed that the system was simple to operate and has delivered excellent results. Baskaran et al. [3] have suggested designing and developing a low-cost transportation management system based on RFID and GSM connectivity. The system is made up of many components that are wirelessly connected to GSM modems. The GSM network's SMS service is cost-effective and is used to send data between different modules. Ramya et al. [13] displays are mounted at every bus stop to inform passengers who are not using the service of the vehicles' whereabouts as they approach that stop. The technique is also effective and advantageous in the event of an error or an emergency. For example, if a bus has a mechanical problem, the operator at the bus terminal is notified, and the departure time between vehicles is shortened to save passengers' time. It is hoped that by implementing this approach, issues such as less utility of the bus fleet and lengthy lines at bus stations will be alleviated. As a result, the technology will assist both passengers and bus station management since it will give real-time information. Raymond and Imamichi [14] projected a system to resolve the trouble of resolving car courses from the sequences of discordant geospatial-material datasets. Nowadays, we witness the collection of instrument course datasets in the form of the sequences of GPS points. Braz et al. [5], the chance of real data on the radio signal receiver courses of taxis and fare boarding news for public transport squadrons of large municipalities has likely investigators and experts the opportunity to investigate new challenges concerning the reasoning of public transit service systems. They are further studied deeply for improving the transit system by Rahbar et al. in their work [12]. Simulation of smart buses in smart cities was performed by Kumar et al. [9] which is a new improvised electric bus. To improve the passenger travel experience, a study was performed by Cheng et al. [6] to find customized bus demand spots using smart cards. The graphical user interface for an industrial application is created using Python or C# or Java as designed by Barmase et al. [2].

According to the literature review conducted, it is found that there are methods for tracking the bus, advanced ticket booking systems, accident detection and alert systems, etc. As of now, there is no system available to monitor the passenger is sitting or standing during the travel and calculate the fare based on sitting and standing distances. The proposed work addresses this as the research gap.

3 Methodology

The proposed work of flexi-pay system is divided into two parts, such as the hardware, and the software. The software is designed based on a Web server programming [1, 19]. The idea is to find out the travel made by the passenger standing in the bus during the transit and reduce the convenience fee for that and to provide that as a cashback offer for the travel.

3.1 Hardware Design

The hardware part consists of a travel card (RFID), RFID reader [11], ultrasonic sensor [8], and an embedded device that is used for the ticketing purpose and is modified with a screen to view the occupancy of the seat. The travel card is an RFID that is used as an identity card for every passenger, and it is assumed that every passenger carries their travel card to use the bus in a similar fashion that fastag RFID to be bought by every car owner to pass through the toll gate in highways of India. The placement of the ultrasonic sensor is behind every seat, and this helps us to identify whether the particular seat is occupied. The RFID reader has been placed in front of every seat this helps the passenger on the bus to scan his RFID when he is seated in a seat. The ticket vending device in the conductor's hand helps the conductor to have a view of the passengers who correctly show their cards in front of the RFID reader. The design of ticket vending device is not discussed here as the focus of this work is mainly on calculating the standing and sitting distance of the passenger during the journey to compute the fare based on it. Presently, the seat occupied or vacant is updated in the database through serial port node package manager (NPM) module, and it is displayed in command window as shown in Fig. 2.

Figure 1 shows the circuit connection for the hardware, and it consists of the Arduino connections with an ultrasonic sensor and RFID reader (Fig. 2).

The prototype of the proposed work is shown in Fig. 3. When the seat is occupied by the passenger, it is sensed by the ultrasonic sensor, and if the RFID reader did not load with any values, then it will be notified by the conductor on his device, so then he can instruct them to place the card on the RFID. If the seat is unoccupied, the distance sensed by the ultrasonic sensor will indicate a high, and it will also be notified by the conductor's device.

3.2 Software Design

The scanner will be connected to the database servers, in which it updates the tables with the RFID unique numbers that are scanned. This table with the latest travel id will be used by the software to track the distance that the passenger has travelled in standing mode.

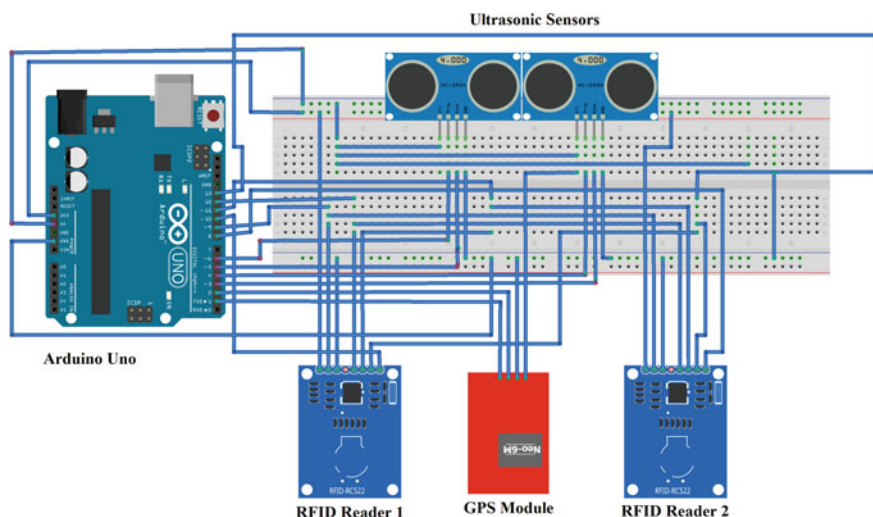


Fig. 1 Schematic diagram of the proposed work

```
C:\WINDOWS\system32\cmd.exe - node Index.js
Microsoft Windows [Version 10.0.22000.613]
(c) Microsoft Corporation. All rights reserved.

C:\Users\tkart>cd C:\Users\tkart\Desktop\8th sem\Bus

C:\Users\tkart\Desktop\8th sem\Bus>node Index.js
Success
```

Fig. 2 Node package manager shows the read status of RFID tag

The software includes administrator and user pages, and Fig. 4 shows the admin page for adding new buses into the database. The form collects the information regarding the bus basic information, the route, and the timing details. On the button click, a record with the bus details is created in the database. Figure 5 describes the admin page for adding stop details to the particular buses. This stores the latitude and longitude details and more details on the stop details and their respective stops.

Figure 6 shows the landing page of the Web application for the passenger. The top left part shows the logo, and the right side shows the travel id of the user, which is a link that will take us to the add family member page to link travel cards with

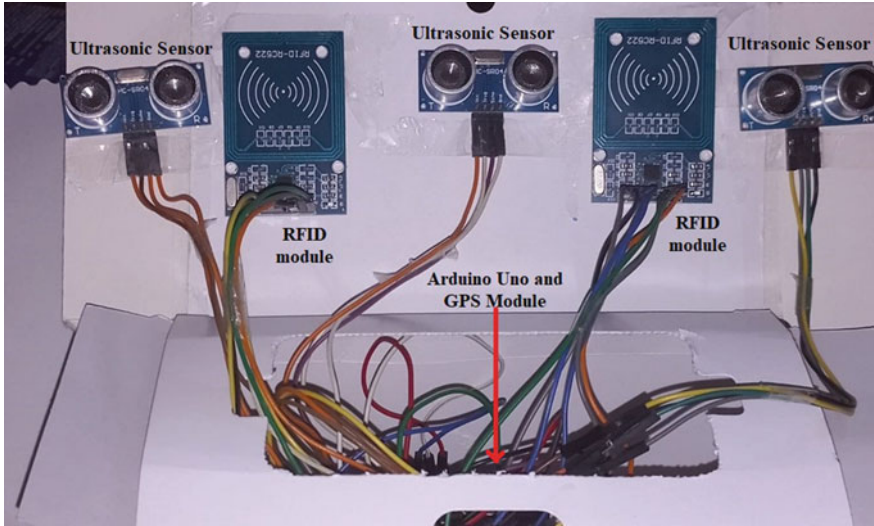


Fig. 3 Hardware prototype of the proposed work

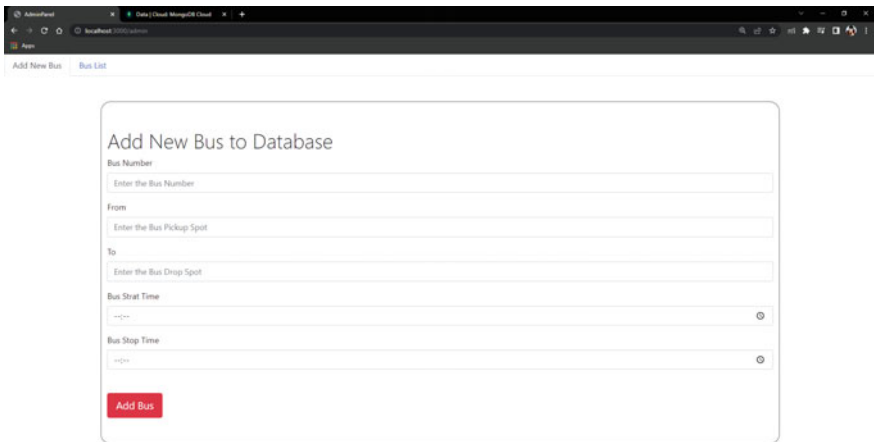


Fig. 4 Administrator user interface to add new buses in a route

the current user. Figure 7 shows that a passenger is adding co-passenger details in the Web page designed to serve the purpose of adding friends and family members during the travel. This helps in the case of the group travel in the bus and avoid individual person taking ticket for themselves.

The next step is to take a ticket it can be done using the Web app which help us to book a ticket and could also be taken using the travel card with the conductor's ticketing device, when booked in the Web app, choose the pick and drop spots from the two drop-down fields as shown in Fig. 8.

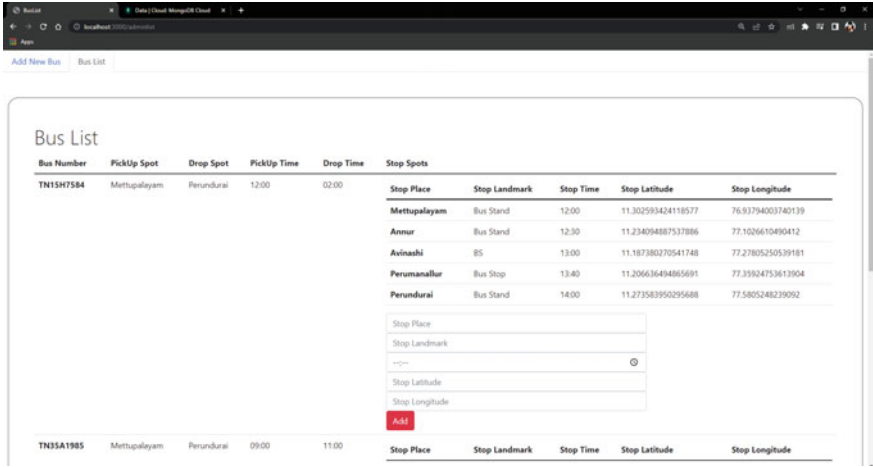


Fig. 5 Administrator user interface to add and edit the schedule of a bus route

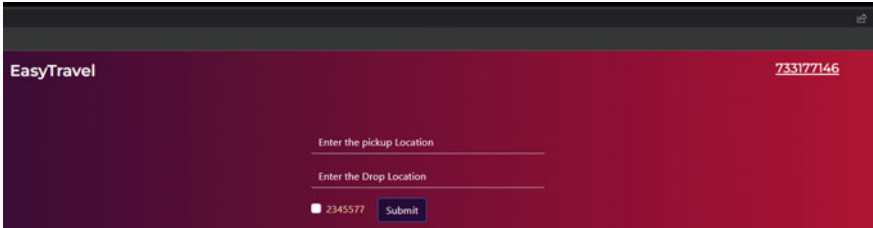


Fig. 6 Passenger Web user interface—home page

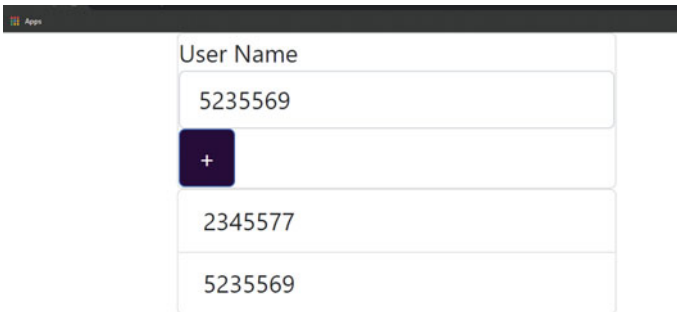


Fig. 7 Passenger can take multiple tickets by adding co-passengers

Figure 9 shows the search result for the route specified by the user. From the available buses, he could choose the bus for the travel. Once he chooses the bus and click book ticket, it would create an object in the server with the details like travel id, bus no, destination, pickup, first and second scan lat and long. This will also have a unique tracking number to track the travel of the passenger.

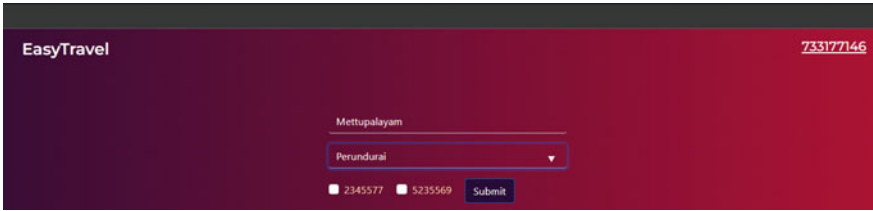


Fig. 8 Passenger user interface—select source and destination of the travel

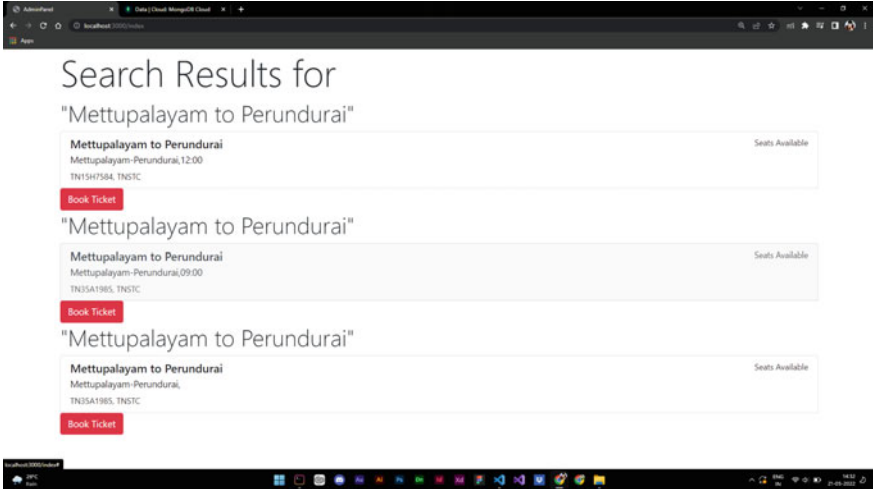


Fig. 9 Search result of the available buses for the selected source and destination

An object is created in the event of booking a ticket, and the card is added to the landing page with attributes like tracking id. The tracking id is unique, and it will help to track the particular travel in the database. A sample booking ticket for a group of three passengers travelling from Mettupalayam to Perundurai is shown in Fig. 10. Figure 11 shows the object that is created in the database for the sample ticket. Figure 12 shows the tracking page before the travel passenger is seated. There is a array in the table that contains the stops through which the bus travel. The card below that list displays the distance travelled between the first and the second scan. The blue box shows the current location from the global positioning system that is recorded to calculate the distance travelled and the fare for travel.

4 Operation Principle

The passenger enters the bus, he used take the ticket for the travel, which can to be taken from the conductor or the Web app. After that, the unique travel id with the

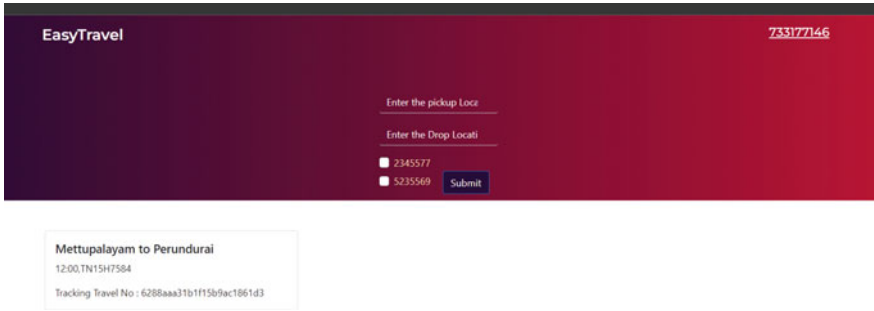


Fig. 10 Travel ticket has been booked for the group of passengers from Mettupalayam to Perundurai

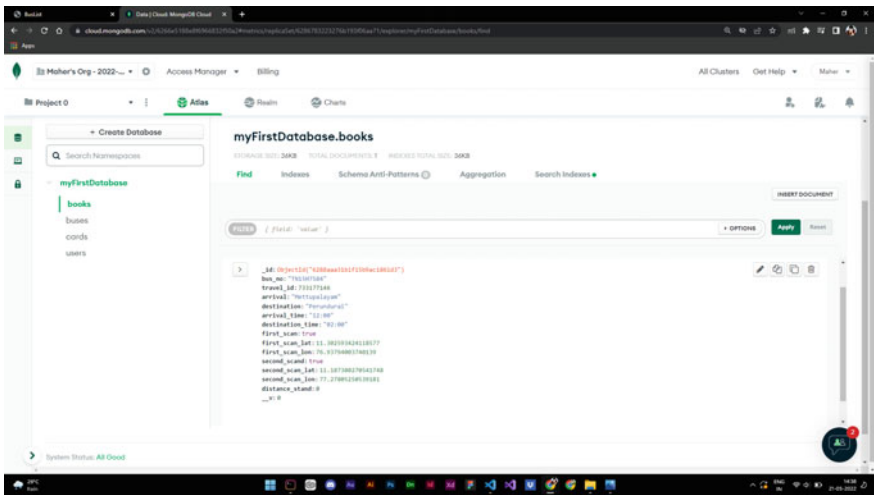


Fig. 11 Record of object created in the cloud database

passenger’s pickup and dropping spot will be noted. Meanwhile, the conductor will also have a display to monitor the seat vacancy of the bus. The vacancy is obtained with the help of ultrasonic sensors placed in the seats. It will sense the presence of the passenger and intimate the conductor with different colours like green, red, and black. Green denotes seat which is occupied and scanned, red denotes seat which is occupied and not scanned, and black denotes seat which is not occupied. If there is any seat available, the passenger can be seated, and if there is no seat available, the distance needs to be calculated for the time in which the passenger travelled in standing. When the passenger finds a seat after the travelling for a particular distance in standing, he should scan his card in the RFID reader in front of the passenger. The attributes that hold the latitude and longitude are updated on the instance of the second scan.

The coordinates are used to find the distance between the pickup spot and sitting spot, and it will be calculated with the Haversine formula shown in Eqs. 2, 3, and 4. In the equation, a represents square of half the chord length between two coordinates, $\Delta\varphi$ is latitude difference between two coordinates, $\Delta\lambda$ is the longitude difference between two coordinates, c is angular distance in terms of radians, R is radius of the earth, and d is the distance between two coordinates. Concerning the distance calculated the amount will be refunded to the passenger's wallet as a cashback which can be used by the passenger at a later period. The concept of wallet integration is inspired from the wallet system that is used for fastag.

$$a = \sin^2(\Delta\varphi/2) + \cos \varphi_1 \cdot \cos \varphi_2 \cdot \sin^2(\Delta\lambda/2) \quad (2)$$

$$c = 2 \cdot \text{atan2} \left(\sqrt{a}, \sqrt{(1-a)} \right) \quad (3)$$

$$d = R \cdot c \quad (4)$$

5 Result and Discussion

As we find above the travel between Mettupalayam and Perundurai, the passenger has travelled for about 44.12 km between the first scan (booking the ticket) and the second scan that is done when the passenger is seated. The distance between the first scan latitude, longitude, and second scan latitude, longitude with the help of the Haversine formula, which helps us to calculate the distance between two locations based on their latitude and longitude values.

The fare for the travel between Mettupalayam to Perundurai, according to Tamil Nadu bus fare calculation will be Rs. 66.00. Figure 13 shows that half the fare from the original ticket is provided as a cashback for the travel from Mettupalayam to Perundurai but got seated only from Avinashi. The total distance travelled by the passenger in standing mode is found to be 44.12 km, and the cashback is calculated as Rs. 10.71. It will be added in the travel wallet account which could be used in next travel by the passenger. The cashback amount is roughly assumed by the authors, and they do not ensure half the price of the fare is refunded. The proposed work is a proof of concept that it is possible to calculate the cashback amount and display it. The result of the travel from Mettupalayam to Perundurai is for different cases which is listed in Table 1. The various cases are categorized into single passenger and multiple passengers travelling both in standing and sitting mode, standing only, and sitting only modes.

Number of passengers is travelling in the bus is represented as P_n , and the distance travelled by the passenger in sitting mode and standing mode is denoted as D_{Sit} and D_{Stand} , respectively. The fare for the travel in sitting and standing mode is denoted as F_{Sit} and F_{Stand} , respectively.

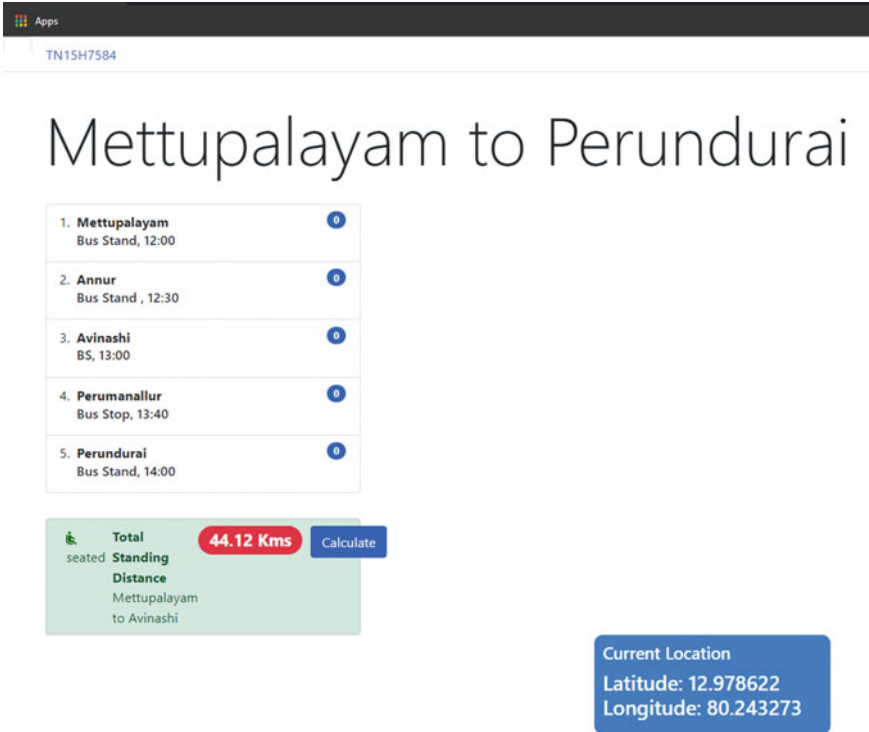


Fig. 12 Web page displays the location at which the scan is done by the passenger after taking seat

Fig. 13 Website displays the cashback to be given to the passenger

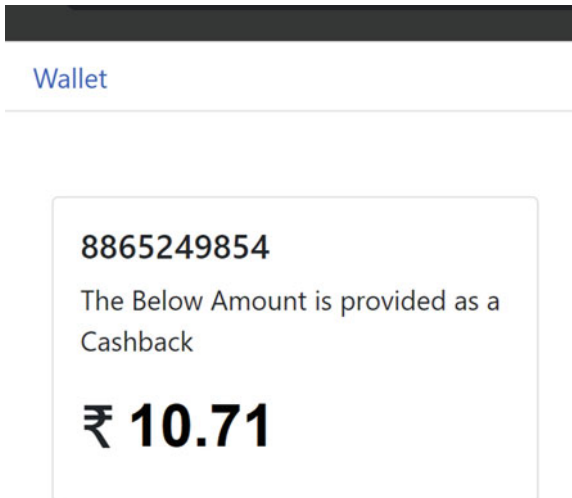


Table 1 Table captions should be placed above the tables

No. of passenger, (P_n)	($P/D_{Sit}/F_{Sit}$)	($P/D_{Stand}/F_{Stand}$)	Total fare	Cashback
1, (P_1)	$P_1/44.12/33.86$	$P_1/41.88/21.43$	55.29	10.71
1, (P_2)	$P_2/0.00/0.00$	$P_2/86.00/33.00$	33.00	33.00
2, (P_3, P_4)	$P_3/22.50/17.27$ $P_4/48.33/37.09$	$P_3/63.50/24.37$ $P_4/37.67/14.45$	93.18	38.82
2, (P_3, P_4)	$P_3/22.50/17.27$ $P_4/48.33/37.09$	$P_3/63.50/24.37$ $P_4/37.67/14.45$	93.18	38.82
2, (P_5, P_6)	$P_5/0.00/0.00$ $P_6/0.00/0.00$	$P_5/86.00/33.00$ $P_6/86.00/33.00$	66.00	66.00
2, (P_7, P_8)	$P_7/86.00/66.00$ $P_8/86.00/66.00$	$P_7/0.00/0.00$ $P_8/0.00/0.00$	132.00	0.00

6 Conclusion

Fare determination for the passengers travelling in the bus based on whether the passenger is seated or standing during the travel is calculated using ultrasonic sensors, GPS and RFID are reported here. The flexi-pay system may reduce the number of people waiting for the bus to arrive with empty seats and encourage the people to travel even if the bus seats are occupied already. It may also reduce the feel of discomfort during the travel as they have to pay less than the seated passengers. The results conclude the evidence that such system could be implemented with certain limitations in the performance and lot of challenges.

The proposed system is only a semi-automation system, to find the distance travelled by the passenger either in sitting or standing position. The performance of the system is currently very limited. The maximum number of users in the current system is limited to 10 users. The number of bus routes that are incorporated is 5, and the stops in-between the 5 places are also included in the database. The GPS tracking is updated every two seconds, but all of the tracking data are not stored as it will consume large memory. The GPS data is stored only when the passenger scans the RFID travel card. The system uses MongoDB to store the data of tickets, bus routes, types of buses, location of each bus stop, user data, etc. It is observed that, if implemented in real time, it will become necessary to use a big data server to handle the requests from all the buses.

There are a few challenges that one may face in real-time implementation of the proposed work. Firstly, all passengers will like only to travel in standing until they get tired. Secondly, a passenger may take a seat but does not scan the RFID travel card. Thirdly, a passenger leaving the seat without scanning the RFID, travel card. The above said challenges have to be monitored manually by the conductor of the public transport. Finally, there is threat for data security in the current system which can be improved using security controls like Web application firewall, subdividing the network into microsegmentation, and implementing anti-distributed denial of service (anti-DDoS) algorithms to avoid data breach.

References

1. Adam T (2019) Automated bus ticket reservation system for Ethiopian bus transport system. *IOSR J Comput Eng (IOSR-JCE)* 21(3):22–27
2. Barmase GV, Khopade GV, Thawkar SP, Bawankule SP, Gajbhiye ND, Sakhare S (2021) GUI based industrial monitoring and control system. *J Inf Technol Digital World* 3(2):108–117
3. Baskaran D, Pattumuthu M, Priyadharshini B, Akram PS, Sripriya S (2016) RFID based smart bus using embedded system. *Int J Eng Res Technol COCODANTR* 4(11)
4. Biswas D, Rahaman M, Anup AK et al (2020) Development an automation transport system for DIU with IoT based tracking. Technical report. Daffodil International University
5. Braz T, Maciel M, Mestre DG, Andrade N, Pires CE, Queiroz AR, Santos VB (2018) Estimating inefficiency in bus trip choices from a user perspective with schedule, positioning, and ticketing data. *IEEE Trans Intell Transp Syst* 19(11):3630–3641
6. Cheng Y, Huang A, Qi G, Zhang B (2019) Mining customized bus demand spots based on smart card data: a case study of the Beijing public transit system. *IEEE Access* 7:181626–181647
7. Dhanasekar N, Valavan C, Soundarya S (2019) IoT based intelligent bus monitoring system. *Int J Eng Res Technol (IJERT)* 7(11):1–4
8. Harini K, Saithri A, Shruthi M (2021) Smart digital bus ticketing system. In: *Advances in automation, signal processing, instrumentation, and control*. Springer, pp 853–858
9. Kumar A, Srikanth P, Nayyar A, Sharma G, Krishnamurthi R, Alazab M (2020) A novel simulated-annealing based electric bus system design, simulation, and analysis for Dehradun smart city. *IEEE Access* 8:89395–89424. <https://doi.org/10.1109/ACCESS.2020.2990190>
10. Lavanya R, Rani K, Gayathri R, Binu D (2017) A smart information system for public transportation using IoT. *Int J Recent Trends Eng Res* 3(04):222–230
11. Oudah A (2016) RFID-based automatic bus ticketing: features and trends. *IOP Conf Ser Mater Sci Eng* 114:012146. <https://doi.org/10.1088/1757-899x/114/1/012146>
12. Rahbar M, Hickman M, Mesbah M, Tavassoli A (2018) Calibrating a Bayesian transit assignment model using smart card data. *IEEE Trans Intell Transp Syst* 20(4):1574–1583
13. Ramya R, Prakruthi N, Sahil B, Reshma N, Asha KN (2019) Smart bus ticketing using barcode scan-one time barcode tickets on android application for travelling through metropolitan buses. *Int J Eng Sci Invent Res Dev (IJESIRD)* 05(10):16–20
14. Raymond R, Imamichi T (2016) Bus trajectory identification by map-matching. In: *2016 23rd international conference on pattern recognition (ICPR)*. IEEE, pp 1618–1623
15. Shaikh DA, Ansari AB, Khan R, Khan AS (2018) Internet of things (IoT) based real time bus monitoring and passenger information system. *Int J Adv Eng Res Dev (IJAERD)* 5:1042–1047. <http://www.ijaerd.com/index.php/IJAERD/article/view/2530>
16. Shaik Saidu Masthan DJ (2020) Design of smart bus ticketing system using IoT and RFID. *Int J Control Autom* 13(02):1267–1272. <http://sersc.org/journals/index.php/IJCA/article/view/27013>
17. Shanmugapriya R, Sowmya A, Sowmya S, Sriram SR (2020) Smart bus monitoring and ticketing system using IoT. *Int Res J Eng Technol (IRJET)* 07:6167–6170
18. Transport Department of Tamil Nadu Government (2021) Policy note—2021 to 2022, demand no. 48. Govt. Central Press, Chennai. <https://www.tn.gov.in/documents/dept/33>
19. Zhuo-Yi C, Hwei L, Miao-Miao W (2012) Design of portable checking and ticketing systems based on GPRS. *Procedia Eng* 29:4002–4006

An Improvised Algorithm for a Dynamic Key Generation Model



D. V. Guru Saran and Kurunandan Jain

Abstract Data may be kept private, secure, and authentic via cryptography. Computational resources and communication channel performance impede the development of an efficient key generation model in IoT devices for encryption and decryption applications. Most IoT networks encode information with session keys, which are less secure than dynamic key encryption used in communication channels. Furthermore, because of the resource limitations of most IoT networks, it is impractical to use the current dynamic key generation approaches. To solve these problems, a dynamic key generator model has been designed that continuously generates unique keys in the range of 1000–10,000,000. We also discuss our proposed model's security and performance analysis to validate the feasibility of its operation in such a resource-constrained IoT environment.

Keywords IoT · Dynamic key · Cryptography · Security

1 Introduction

It is only a matter of time until the Internet of Things (IoT) becomes an increasingly crucial part of our everyday lives, thanks to the arrival of 5G technology. A new security danger is created, or malicious attackers can get into the information and equipment that are supposed to be secured [1]. Certificate-based encryption and public key distribution are adequate to assure data confidentiality, integrity, and validity at OSI network layer levels. These services are unsuitable for IoT devices due to their memory, computational power, energy consumption, and equipment space limitations [2]. In addition to standard network security weaknesses, the Internet of Things offers new security challenges due to its unique features and the fact that it

D. V. G. Saran (✉) · K. Jain
Center for Cyber Security Systems and Networks, Amrita School of Engineering, Amritapuri
Campus, Vallikavu, Kerala, India
e-mail: amenp2csn20010@am.students.amrita.edu

K. Jain
e-mail: kurunandanj@am.amrita.edu

includes several sensor nodes. Data transiting through the Internet of Things (IoT) network is protected by lightweight cryptographic algorithms [3].

Data confidentiality, integrity, and authenticity may all be ensured by using cryptography. The information must be kept private to prevent it from being accessed or misused by others. Authenticity ensures that data cannot be disputed, whereas data integrity ensures that it is trustworthy and reliable information being carried across the network [4].

Cryptographic keys are used for encryption and decryption to ensure security. These keys encrypt the data by converting plaintext to ciphertext. A cryptographic key can be created in one of two ways. Symmetric and asymmetric keys are two sorts of keys. In symmetric key encryption, the very same secret is used to encode and decode the message. In asymmetric keys, one encrypts the message with one key and decodes it with another key [1, 5]. While dynamic keys are used once for each message transmitted, session keys are used to encrypt and decrypt the information in a session to ensure session security. The system's security is enhanced because dynamic keys are utilized for every data transmission [6].

In the IoT network context, malicious users can target the data or devices that generate the data. Active and passive attacks on an IoT network/system are two types of attacks. Passive attacks include the attacker listening to and analyzing traffic passed between objects without altering it. This enables the attacker to ascertain sensitive data like credentials or keys shared, which can be abused to perform cyberattacks. During cyberattacks, the attacker alters, deletes, or replaces transmitted data with malicious messages [7]. It can also impersonate a legitimate node, engage in the key formation procedure, or replay valid messages to gain the system's trust to steal sensitive data such as cryptographic keys.

In previous research [8], we created and implemented a method for generating discrete, transitory keys known as dynamic keys, which are used to determine the delivery time of numerous packets in data transmission or the delivery time of a single packet. Using symmetric key encryption, the suggested dynamic key generation mechanism is developed to work in IoT environmental system. In this investigation, we assume that every individual message transmission between communicative parties uses the dynamic key created by our technique. In comparison with conventional dynamic key generation approaches, we provide a robust system for cryptanalytic assaults while using fewer resources. Three steps comprise the key generation algorithm of the cyclic dynamic key generator (CDKG). It utilizes an input pair of secret and seed to create dynamic keys without a key distribution center (KDC) intervention.

This paper redesigns the algorithm to boost efficiency while preserving strong resilience to cryptanalytic attacks. Furthermore, we replace the SHA256 hash with the Blake2b hash because BLAKE2 is much more efficient in software over most platforms than SHA256. When hashing 64 bytes, for example, it is often twice as efficient as SHA256 on ARM devices [9]. BLAKE2 includes potentially valuable security features such as protection against length extension, a common technique of keying, "personalization tags" to ensure domain separation, etc. In addition, we implemented configuration stage and random stage, which will be discussed in Sect. 3.

The following are the main contributions for this paper:

1. Redesign and enhance the model's performance for dynamic key generation intended for use in an IoT ecosystem.
2. Analyze the efficiency and safety of the suggested method for protecting IoT devices.

2 Literature Survey

Most key management solutions prefer to encrypt and decode messages with session keys. Using session keys to encrypt data is vulnerable to eavesdropping or compromising the data. Using a session key for more than one session may compromise the data, so dynamic keys are used because they are one-time use keys employed per message instead of per session. Dynamic keys are preferred over session keys due to their security resilience [6], whereas session keys are preferred for computation and storage purposes.

In this research, Thanh Nha Dang [10] suggested an improved AES scheme that creates dynamic keys. A 16-byte data packet was delivered in a numerically indexed pattern. Based on the pattern, the key in an Internet of Things (IoT) system was dynamically updated with encrypted data.

At the PHY layer, the researchers researched, assessed, and uncovered flaws in OFDM-based encryption methods [11]. It is been discovered that frequency response encryption reduces the effects of channel fading and improves bit rate error performance. A new method for altering encryption blocks for input OFDM frames and a dynamic secret key mechanism has been presented to strengthen the security of OFDM-based encryption systems.

Uchiteleva's method dynamically updated the secret key on both endpoints of a network connection by using pseudo-random (PR) [12] sequences produced at the physical layer of wireless transmitters throughout the data transmission. In an Industrial Internet of Things (IIoT) setting, this method is adaptable and suited for large network nodes with limited computational power.

Utilizing a Diophantine variant of the nonlinear equation [13], Thirumalai's suggested solution repurposes the RSA technique for storage efficiency. In addition, this technique worked exceedingly well due to its sole use of an RSA public key. MEMK does not involve a multiplicative inverse function or extended Euclid's method. To acquire a test result for MEMK PKC key generation, encryption, and decryption phases, they ultimately increased the N-bit modulo bits from 1 to 10 K.

Moosavi [14] provided two ECG-based cryptographic key generation algorithms. They first suggested using the Fibonacci linear feedback shift register (LFSR) pseudo-random number generator to generate a sequential set of APIs. Second, they demonstrated a key generation mechanism that is both highly secure and low cost. Researchers discovered that traditional IPI-based solutions take 12.3% and 41.2% more time when comparing key generation execution times.

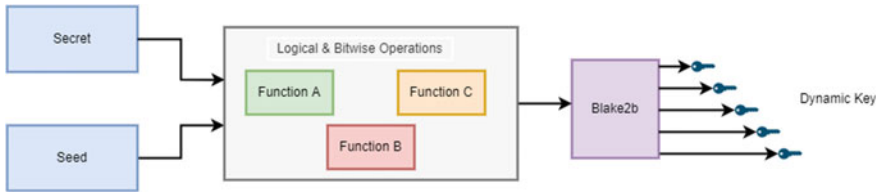


Fig. 1 General overview of key generation methodology

This research proposed the generalized triangle-based security algorithm [15] as an energy source, resource data encryption algorithm with an acceptable encryption generation mechanism (G-TBSA). The proposed G-TBSA is implemented in constrained Wi-Fi wireless sensor networks (WSNs). The key generation process is at the algorithm's core because it requires low resources to produce keys, reducing algorithm complexity and improving energy efficiency.

3 Proposed System

3.1 General Overview of the Key Generation System

Figure 1 depicts a broad overview of the key generation approach. The three functions use logical and bitwise operations to produce random keys supplied as an input to the Blake2b cryptographic hash function. Using the Blake2b hashing technique, a 256-bit hash is generated. The result generates the dynamic key from the specified secret and salt. There are four stages of operation in the key generation scheme: the initialization stage, the configuration stage, the randomness stage, and the cyclic stage.

3.2 Initialization Stage

Figure 2 depicts the broad operational perspective of the initialization stage. Two pre-shared alphanumeric values are required as inputs for this stage: a secret and a seed. Only Function A, the configuration stage, and BLAKE are active at this level. The configuration stage is contained inside Function A. After completing the configuration stage, it executes logical and bitwise operations to generate a new secret. The freshly generated secret is sent to Function C or Function B along with the Blake2b hashing method. Blake2b combines the pre-shared salt and freshly produced secret to construct a 256-bit dynamic key. The key generation algorithm then moves on to the cyclic and random stages.

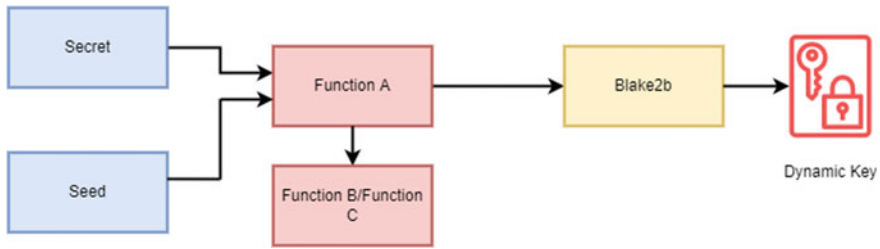


Fig. 2 Initialization stage

3.3 Configuration Stage

This stage is implemented in Function A. Instead of sending the input directly to logical and bitwise operations, it performs mathematical operations on the secret and seed pair, as shown in Fig. 3. Two pre-shared alphanumeric values are inputs in the configuration stage, i.e., a secret and seed, combined and stored in the input. The combined input is divided into six equal groups by padding with zero. Apply mod 6 to these six equal groups and make adjustments such that S1, S2, S3, S4, S5, and S6 are obtained. After the configuration stage, the divided input performs logical and bitwise operations in Function A, yielding a new secret as an output. The newly created secret is sent into Function C/Function B and the Blake hash algorithm.

3.4 Randomness Stage

The general workings of the randomness stage are indicated in Fig. 4. This stage is implemented in Function C. The stage requires a secret generated from Function A and a key generated from the previous result, i.e., a dynamic key. Function C performs logical and bitwise operations on the secret generated from Function A and the key generated from the previous result to generate a new secret. The new secret is supplied to function B as an input. If there is no previous result stored in memory, the new secret generated from Function A is directly passed to Function B. By implementing this stage, for the same secret and key, we generate different sets of random keys.

3.5 Cyclic Stage

Figure 5 depicts the main functioning overview of the cyclic stage. At this level, all modules are operational (Function A, Function C, Function B, and BLAKE). Function A aims to create new secrets from Function B's output pairs. Function C's

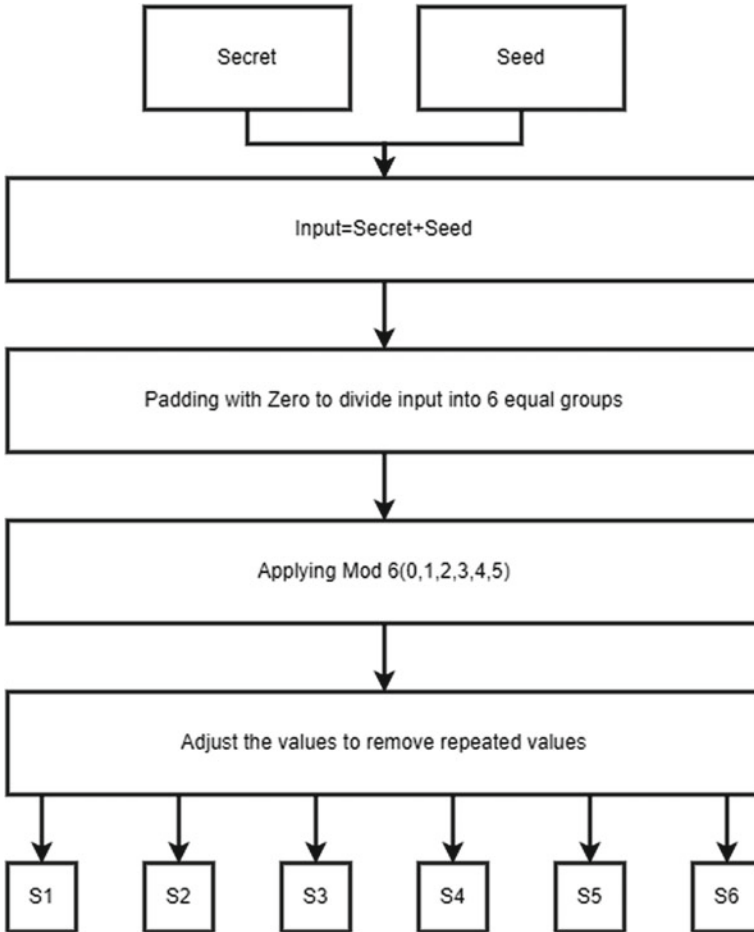


Fig. 3 Configuration stage

purpose is to give Function B a secret. From the outputs of Function A and Function C, Function B tries to create fresh secrets and seeds for Function A. The secrets created by Functions A and B are continuously fed into the 256-bit dynamic key generating Blake hash function.

4 Key Generation Methodology

This section provides a detailed summary of each function. Table 1 gives each function’s terms and notation used in this section.

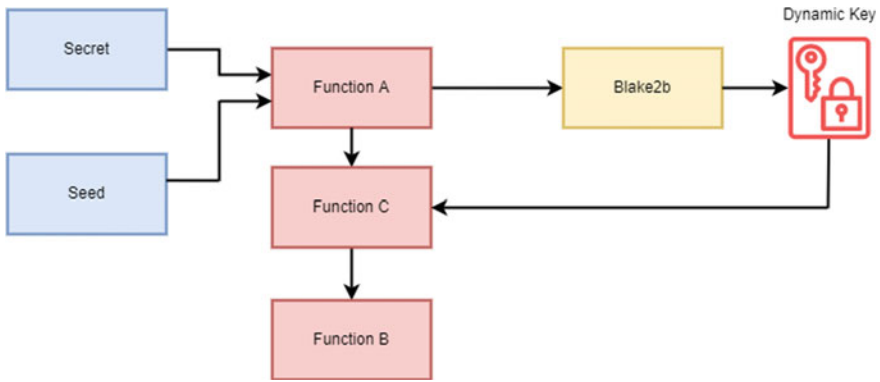


Fig. 4 Randomness stage

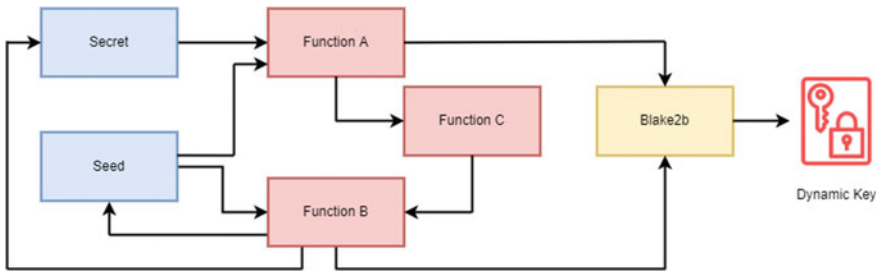


Fig. 5 Cyclic stage

Table 1 Terms and notations

Symbol	Notation
XOR	Bitwise exclusive
LCS	Left circular shift
RCS	Right circular shift
LRS	Logical right shift
LLS	Logical left shift

4.1 Function A

Figure 6 depicts the inner workings of Function A in detail. The secret and seed are first transferred to the configuration stage. The detailed working of the configuration stage is indicated in Fig. 3. After the configuration stage, the input divides the secret into six parts: S1, S2, S3, S4, S5, and S6; S5 and S6 undergo the left and right circular shift operations, respectively. The circularly shifted S5 and S6 perform an Exclusive OR operation with S3 and S4, resulting in S7 and S8. Similarly, S3 and S4 perform the left and right circular shift operations, respectively. The circularly shifted S3 and

S4 perform an Exclusive OR operation with S1 and S2, yielding S9 and S10. S7, S8, S9, and S10 are subjected to a logical right shift operation, with the resulting outputs combined together to generate a new secret.

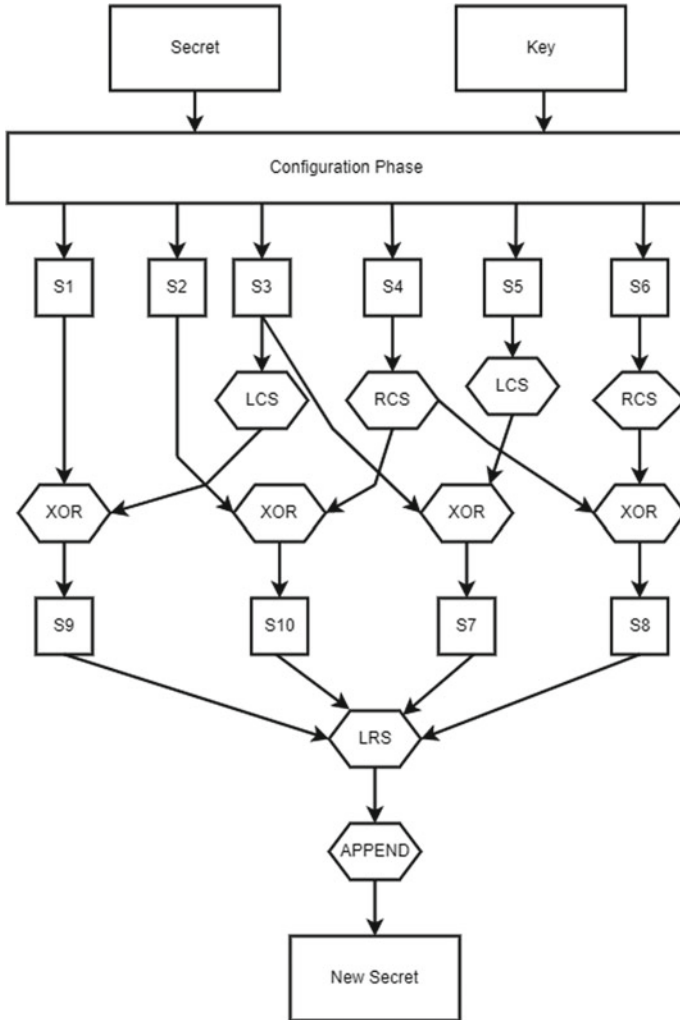


Fig. 6 Function A

4.2 Function C

Figure 7 illustrates the inner workings of Function A in great detail. The stage requires a secret produced by Function A and a key produced by the preceding result. This function C is valid if it returns the previous value. If there is no prior result, Function A immediately passes the newly produced secret to Function B. The method divides the secret and the key into four parts: S'1, S'2, S'3, and S'4. K'1 and K'3 suffer circular shifts to the left and right, respectively. Circularly shifted K'1 and K'3 are subjected to an Exclusive OR operation, with S'1 and S'3 producing S'6 and S'5. S'2 and S'4 suffer circular shifts to the left and right, respectively. S'8 and S'7 are created when S'2 and S'4 are circularly moved. The outputs of the logical right shift operation performed on S'5, S'6, S'7, and S'8 are concatenated to produce a new secret.

4.3 Function B

It is shown in Fig. 8 how Function B works from a high-level perspective. S"1 to S"4 are the four components of the secret that the algorithm divides into and the seed into two parts (K1 and K2). Shifts in the direction of left and right circular motion occur for K1 and K2. As a result of this process, K4 is generated. A left circularly shifted K1 conducts an Exclusive OR operation with S"3 to create K3 as a consequence of the Exclusive OR operation with S"1. K3 and the result of S"2's left circular shift is supplied as input to an Exclusive OR operation, which generates the output. S"6. Using Exclusive OR, the result of the right circular shift operation on S"1 and K4 is used to produce the desired output. S"5. Like S"3, the result of Exclusive ORing S"2 and a right circularly shifted S"3 is produced by S"3. S"7. An Exclusive OR operation on S'1 with a left circular shift is carried out S"4 is responsible for the final product. S"8. After performing a logical left shift operation on each S5, S6, and S7, the outputs are appended together to form a new secret. Logical right shift is applied to K3 and K4, and the outputs are appended together to form a new seed.

5 Security Methodology

This section goes through the numerous attack vectors used against the proposed approach key system. These tests demonstrate that the proposed approach key system is resistant to different attacks that an attacker might attempt to reveal the keys.

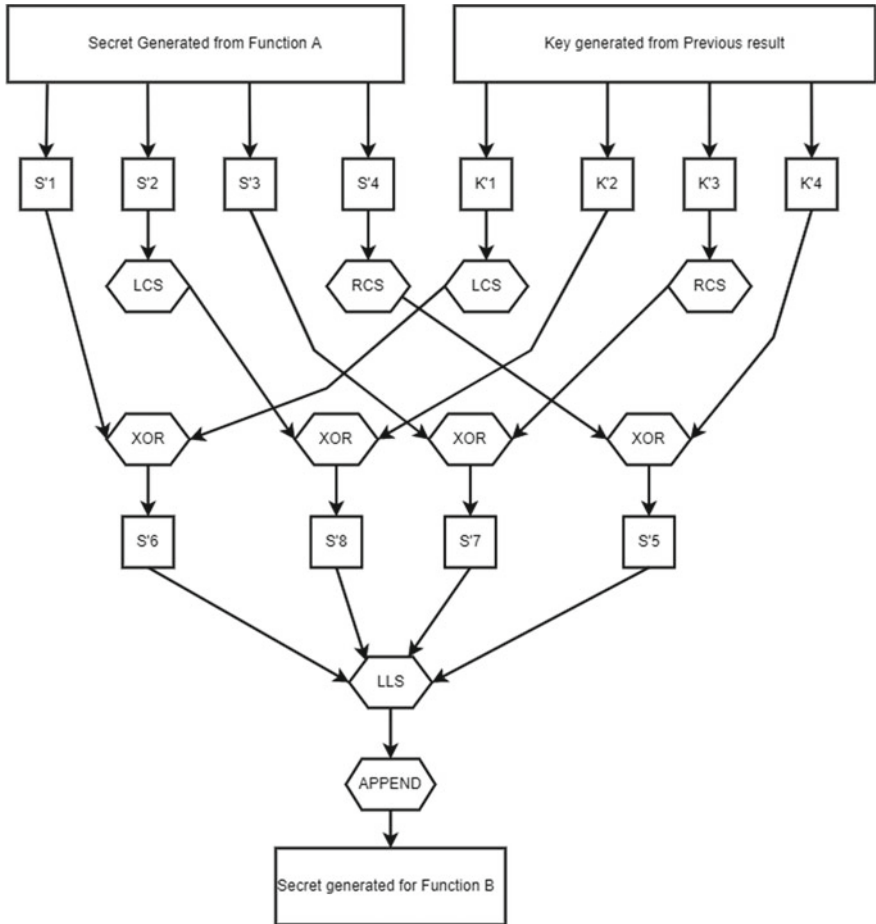


Fig. 7 Function C

5.1 Chosen Plaintext Attack

A cryptanalysis attack can result in illegitimate access to sensitive data due to the compromised symmetric keys. The adversary sends plaintexts of his or her choosing to the encryption oracle to get comparable cipher messages. The adversary attempts to deduce the keys used by the oracle based on the plaintext–ciphertext pairs retrieved. Figure 9 depicts the general attack scenario. Due to the exposure of symmetric keys, this attack may result in unauthorized access to sensitive data.

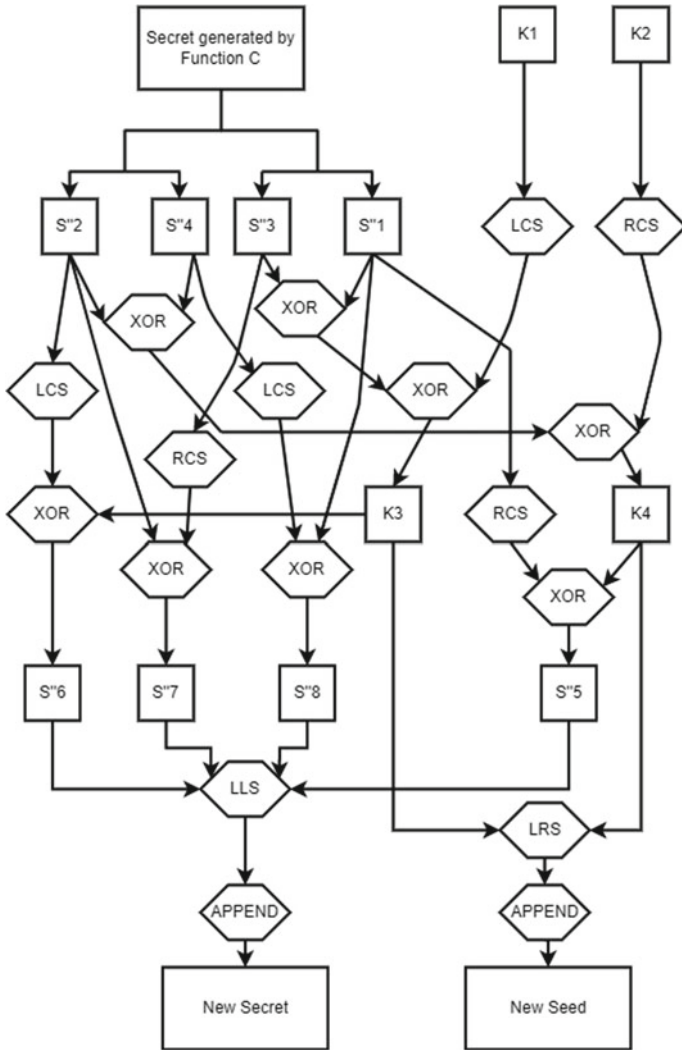


Fig. 8 Function B



Fig. 9 Chosen plaintext attack

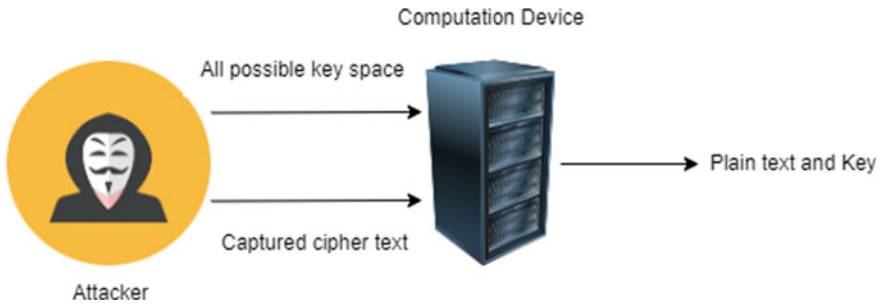


Fig. 10 Brute force attack

5.2 Brute Force Attack

The term “exhaustive key search methodology” refers to this cryptanalysis method. In this attack, the decryption of ciphertext using all possible key spaces determines the key used to encrypt the data. The strength of the algorithm’s cryptographic keys may be evaluated using this attack. The key to one message on the transmission medium can be used to decode all other messages if the attacker is successful. As seen in Fig. 10, the general assault scenario. An attacker cannot use the same key to decode many messages in the same transmission medium because of the dynamic key. There must be a way for an attacker to brute force an initial seed or secret, in addition to being able to authenticate message decryption using the dynamic key generation method.

5.3 Preimage Attack

The attack puts the hashing technique to the test against preimage resistance. This attack vector has two variants: the first and second preimage attacks. A first preimage attack allows the attacker to identify a message with a hash length of less than 2^N . Resistance to the first preimage assault indicates that it is impossible to identify any second input, i.e., given message M . It is impossible to create a second preimage $M' \neq M$ such that $H(M) = H(M')$.

5.4 Replay Attack

This is a version of the Man-In-The-Middle (MITM) attack, in which attackers capture network communication packets and retransmit tampered but genuine packets to impersonate the legitimate user. This form of attack is frequent when

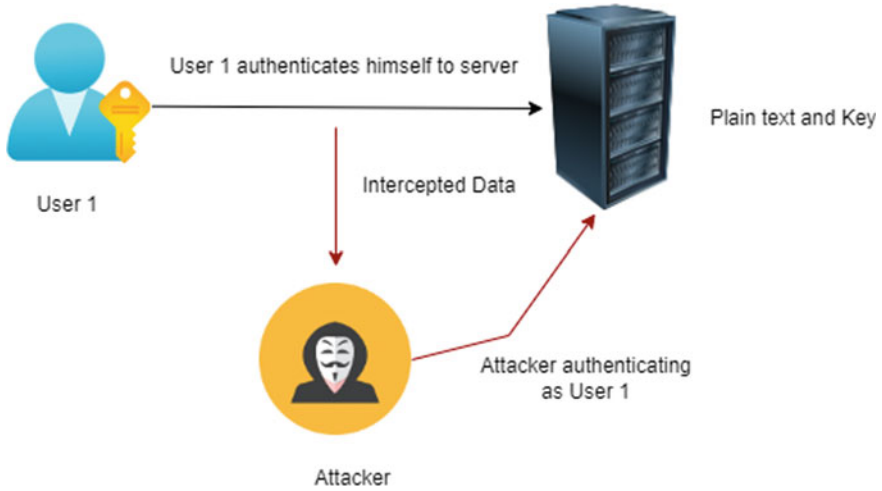


Fig. 11 Replay attack

a legitimate system uses reusable authentication keys. This attack may allow an adversary to gain access to system resources. Attack scenarios are depicted in Fig. 11.

5.5 Session Hijacking

After acquiring the stolen session key or token, an adversary operating as a gateway can get unauthorized access to sensitive information and applications in devices/services/networks. This attack can be carried out by either stealing or guessing an authentic session key to acquire unauthorized access to the system. Figure 12 depicts the general attack scenario. This attack methodology can aid in determining whether a suggested scheme for session monitoring is possible.

6 Performance Analysis

The suggested dynamic key scheme’s performance measures are outlined in this section. Randomness, storage, and computation requirements assess if the algorithm can work in an IoT environment.

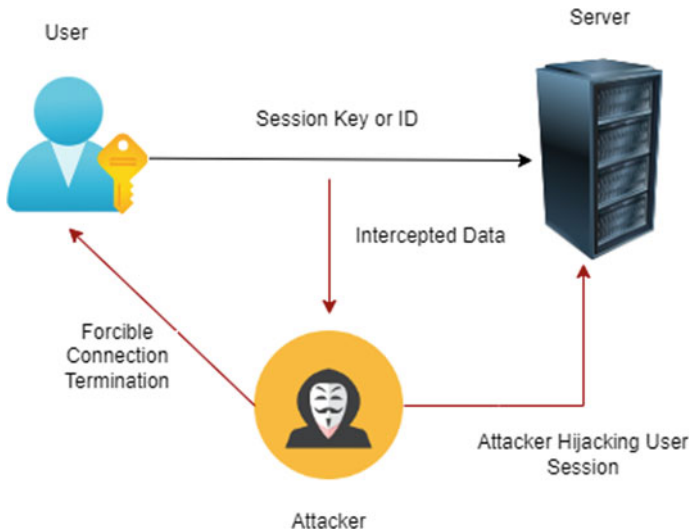


Fig. 12 Session hijacking attack

6.1 Test Configuration

The key generation scheme was tested on the Raspberry Pi 4 Model B, a tiny single-board device commonly used in IoT networks [16, 17]. Table 2 lists the hardware specifications for this single-board computer.

Before transmitting data across the network, this single-board computer may process raw data obtained from different sensor feeds and execute cryptographic procedures. The suggested approach is intended to produce dynamic keys spanning from 1000 to 10,000,000, and various performance features in compute, and storage requirements are tested. The prototype was built using Python 3, an interpreted, high-level, general-purpose programming language. All Python libraries required for the prototype’s proper operation are supported by the Raspberry Pi OS, which runs on the hardware.

Table 2 Raspberry Pi 4 model B hardware specifications

Version	Type
Instruction set	ARMv8 (64/32-bit)
SoC	Broadcom BCM2711
FPU	Neon-FP-ARMv8
CPU	4 X Cortex-A72 @ 1.5 GHz
GPU	Broadcom VideoCore VI 500 MHz
Memory (SDRAM)	1 GB

6.2 *Randomness Analysis*

Randomness analysis assures that the output created by an algorithm is pattern less. This analysis helps to confirm that the produced keys are unique. The National Institute of Standards and Technology (NIST) proposed several to verify the randomness of generated dynamic keys [18, 19]. These tests use conventional normal and chi-square (χ^2) distributions as reference distributions. If the pattern under analysis is not random, the resulting test statistic will be in the extremes of the reference distribution. Table 3 provides an overview of the test completed.

7 **Results and Discussions**

This section describes the proposed dynamic key generation scheme's security features and the results of several testing.

7.1 *Security Features*

1. **Mutual Authentication:** Mutual authentication happens when both sides of a communications connection, rather than simply one, authenticate one others identity. Most IoT devices require a connection to a remote server to work correctly. They may also be required to connect to other IoT devices. IoT devices must communicate via an untrustworthy network (the Internet). Mutual authentication decreases the possibility of attackers compromising their connections by guaranteeing that the data they receive is trustworthy and from a legitimate source.
2. **Session keys security:** In each communication session, dynamic keys are limited to a fixed amount of messages ranging from one to hundreds. On the other hand, session keys are utilized for the duration of the communication session. The maximum number of messages that can be encrypted with a single dynamic key exceeds the maximum number of messages that may be exchanged in a communication session.
3. **Perfect forward secrecy:** Forward secrecy (FS) is a cryptographic primitive that changes the keys used to encrypt and decrypt data automatically and frequently. Even if the most recent key is compromised, just a small amount of necessary data is released due to this ongoing process. Even if the server's private key is compromised, absolute forward secrecy helps protect session keys. Since dynamic keys could only be used once in messages in a single communication session, they add an extra layer of protection to secrecy.
4. **User Anonymity:** Intruders continuously look for security weaknesses and sensitive data such as passwords and user IDs. Intruders can use these credentials

Table 3 NIST test overview

Name of the test	Description
Frequency test	There should be an approximately equal amount of ones and zeros in a series, which is what the test measures
Frequency test within a block	Ones in an M-bit block are being tested to check if their frequency equals $M/2$
Run test	The number of one- and zero-runs of varying durations is tested to see if the random distribution using the runs test. With this test, you can see if the rate of change between 0 and 1 s is too rapid or too sluggish
Test for the longest run of ones in a block	Tests if the longest run of ones in the tested sequence is consistent with the longest run anticipated in a random distribution
Binary matrix rank test	Tests if fixed length subsequences of the original sequence are correlated linearly
Discrete Fourier transform test	Assuming randomness is the assumption, this approach checks the analyzed sequence for periodic characteristics (i.e., recurrent sequences that are near together). We are looking for a noticeable difference between 95 and 5% in peak counts, surpassing that requirement
Non-overlapping template Matching test	With this test, we are trying to determine which generators create an abnormally high number of non-periodic (aperiodic) patterns we have defined
Overlapping template matching test	A pre-specified objective string's frequency is examined using the overlapping template matching test. To distinguish this from the non-overlapping matching test, the window only advances one bit before the search is resumed when the pattern is found. This is the sole difference between the two
Maurer's universal statistical test	With this test, we are looking to see whether there is any way to reduce sequence size without sacrificing quality dramatically. It is possible to reduce the size of a non-random sequence considerably
Linear complexity test	Determines if the sequence is sufficiently complicated to be considered random. Longer LFSRs can distinguish random sequences. An LFSR that is too short shows non-randomness
Serial test	We will examine if the 2 m m-bit overlapping patterns appear as frequently as would be anticipated from a random sequence in this test

(continued)

Table 3 (continued)

Name of the test	Description
Approximate entropy test	The topic of this test is the frequency of all conceivable overlapping m-bit patterns in the whole sequence. With this test, we are looking to see how often blocks of lengths m and m + 1 overlap with the predicted frequency of random blocks
Cumulative sums test	In order to identify if the examined sequence contains too many or too few component sequences, a test is conducted. The overall cost is about the same as going on a walk at random. For some non-random sequences, the random process's deviations from zero will be considerable
Random excursion test	This test aims to examine if the frequency of a certain state's occurrence in a cycle differs from the random sequence. There are eight tests in this assessment, one for each of the states: -4, -3, -2, -1 and + 1, + 2, + 3, + 4. There are eight tests in this evaluation (and conclusions)
Random excursion variant test	This experiment aims to see if the unpredictability of the number of trips to different states differs from what is predicted. Each state -9, -8,..., -1 and + 1, + 2,..., + 9 is represented by one of the eighteen tests included in this study (and the conclusions drawn from them)

to carry out MITM and impersonation attacks. We can keep specific network identities hidden when exchanging messages using dynamic keys.

5. Key dynamics: When you utilize a fixed key for an extended period, the likelihood of breaking it increases. This method generates a new key each time it starts, recognizing the dynamics of the key and significantly increasing its security.
6. Full randomness test pass: The redesigned model passed all the series of randomness tests given by NIST

7.2 Security Test

1. Cryptanalysis and Brute force attacks: The algorithm generates a one-time master key that is used to secure individual messages rather than entire sessions. It functions similarly to a one-time pad. The attacker never finds the keys using the ciphertext attack because analyzing encrypted messages with many unique keys is difficult. The breach of a single key reveals only one message; it is impervious to cryptanalysis assaults. A critical space of $1.1579209e + 77$ is required to brute force a 256-bit key. In theory, it would take 3.671052 years for a supercomputer with a speed of 100 petaflops to deplete all 256-bit key space. Rendering

this type of attack against the redesigned key generation mechanism is infeasible. However, this situation is only applicable to fetching a single dynamic key. Because the key space is the same size, the amount of computing required to brute force the input seed and secret pair and validate the decryption of individual information increases nearly 10 million times to decipher a complete communication session correctly.

2. **Replay and Session Hijack Attacks:** Replay attacks are common on network devices that use reused authentication keys. By employing the provided dynamic approach, the attacker is prevented from using the same access code to get access to network resources. To prevent replay attacks, a single dynamic key can be used to generate an access code that can only be verified once. If an attacker obtains the session key, he or she can take over a user session by intercepting the user's session key and impersonating this user to maintain the connection with the system. This attack strategy is not conceivable in our redesigned model because session communications are encrypted using dynamic keys.

8 Hardware Consumption

The algorithm is efficient regarding entropy and storage because the prototype and keys take up only 7 KB of code and key storage space, as given in Table 4. Ten million keys were produced from an alphanumeric seed and a secret during the prototype stage. Figure 13 depicts the rise in time as the number of keys, increases, and we also compared the redesigned algorithm to the previous algorithm. The process takes 4688 s to produce ten million keys. While creating ten million dynamic keys, the technique consistently used less than 20 MB of RAM.

Table 5 gives the results of the NIST randomness testing. The test results indicate that the method generates random output sequences, which correspond to produced keys being unique.

Table 5 gives the minimum P-Value of all possible states examined for random excursion and excursion variant testing. The predefined threshold value, i.e., P-Value ≥ 0.01 for all x states indicates that the output is random.

Table 4 Storage key space

Keys	1 K	10 K	100 K	1 M	10 M
Storage	63.5 KB	634.8 KB	6.2 MB	62 MB	619.9 MB

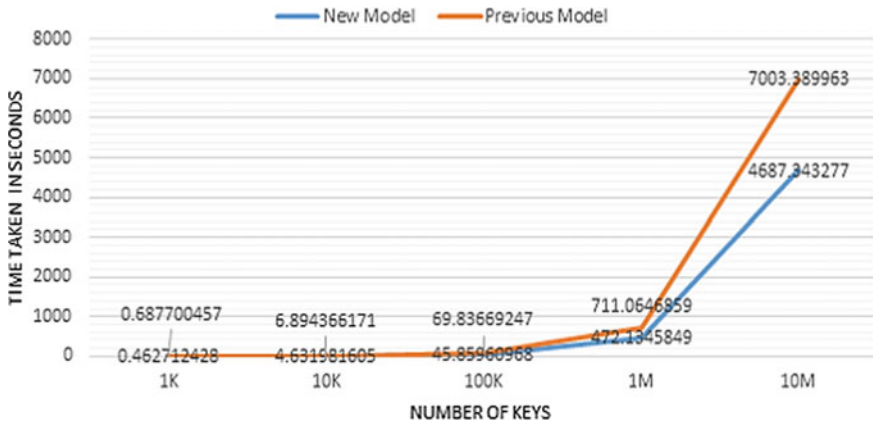


Fig. 13 Comparing performance analysis of previous and redesigned model

9 Conclusion

In conclusion, we describe a cyclic dynamic key generation technique capable of producing many unique keys. Most cryptanalytic assaults, such as dictionary attacks, ciphertext-only attacks, brute force, preimage, and replay attacks, are resistant to the method (and, by extension, the keys). The redesigned algorithm has better performance and full randomness compared to the previously designed model. Since this key generation technique is intended to be deployed in an IoT ecosystem, the prototype developed for the algorithm operates efficiently in resource-constrained environments.

Table 5 Randomness test result

S. No	Test type	Previous model result (P-Value)	Modified model with Blake2b (P-Value)
1	Frequency test	0.62202	0.5815
2	Frequency test within a block	0.0261	0.47955
3	Run test	0.6593	0.96586
4	Test for the longest run of ones in a block	0.6539	0.68468
5	Binary matrix rank test	0.0852	0.45961
6	Discrete Fourier transform test	0.1661	0.49789
7	Non-overlapping template matching test	0.0292	0.99986
8	Overlapping template matching test	NAN	0.06119
9	Maurer's universal statistical test	NAN	0.10592
10	Linear complexity test	0.03486	0.08011
11	Serial test	0.0523	0.49896
12	Approximate entropy test	1	1
13	Cumulative sums test (Forward)	NaN	0.5876
14	Cumulative sums test (Reverse)	NAN	0.79917
15	Random excursions test	0.0508	0.2109
16	Random excursions variant test	0.1573	0.58388

^a NAN-Not a number. Comparing the NIST randomness testing results of previous and redesigned model.

References

1. Kimani K, Oduol V, Langat K (2019) Cyber security challenges for IoT-based smart grid networks. *Int J Crit Infrastruct Prot* 25:36–49
2. Samaila M, Neto M, Fernandes D, Freire M, Inácio P (2018) Challenges of securing Internet of Things devices: a survey. *Secur Priv* 1:e20
3. Abomhara M, Kjøien G (2014) Security and privacy in the Internet of Things: Current status and open issues. In: 2014 international conference on privacy and security in mobile systems (PRISMS). pp 1–8
4. Kumari S (2017) A research paper on cryptography encryption and compression techniques. *Int J Eng Comput Sci* 6:20915–20919
5. Mushtaq M, Jamel S, Disina A, Pindar Z, Shakir N, Deris M (2017) A survey on the cryptographic encryption algorithms. *Int J Adv Comput Sci Appl* 8:333–344
6. Ngo H, Wu X, Le P, Wilson C, Srinivasan B (2010) Dynamic key cryptography and applications. *Int J Netw Secur* 10:161–174

7. Singh K, Tomar S, Jain U, Hussain M (2020) Simplified and secure session key sharing for Internet of Things (IoT) networks. In: International conference on internet of things and connected technologies. pp 319–332
8. Pothumarti R, Jain K, Krishnan P (2021) A lightweight authentication scheme for 5G mobile communications: a dynamic key approach. *J Ambient Intell Humanized Comput* 1–19
9. Aumasson J, Neves S, Wilcox-O’Hearn Z, Winnerlein C (2013) BLAKE2: simpler, smaller, fast as MD5. In: International conference on applied cryptography and network security. pp 119–135
10. Dang T, Vo H (2019) Advanced AES algorithm using dynamic key in the internet of things system. In: 2019 IEEE 4th international conference on computer and communication systems (ICCCS). pp 682–686
11. Melki R, Noura H, Mansour M, Chehab A (2018) An efficient OFDM-based encryption scheme using a dynamic key approach. *IEEE Internet Things J* 6:361–378
12. Uchiteleva E, Hussein A, Shami A (2020) Lightweight dynamic group rekeying for low-power wireless networks in iiot. *IEEE Internet Things J* 7:4972–4986
13. Thirumalai C, Kar H (2017) Memory efficient multi key (MEMK) generation scheme for secure transportation of sensitive data over cloud and IoT devices. In: 2017 innovations in power and advanced computing technologies (i-PACT), pp 1–6
14. Moosavi S, Nigussie E, Virtanen S, Isoaho J (2017) Cryptographic key generation using ECG signal. In: 2017 14th IEEE annual consumer communications networking conference (CCNC). pp. 1024–1031
15. Ahmed S, Islam M, Nath T, Ferdosi B, Hasan A (2019) G-TBSA: a generalized lightweight security algorithm for IoT. In: 2019 4th international conference on electrical information and communication technology (EICT). pp 1–6
16. Akshay S, Vishnukumar B, Mohan V, Anand M (2018) Energy and performance analysis of raspberry pi with modern computing devices. *Int J Eng Technol* 7:777–779
17. Zhao C, Jegatheesan J, Loon S (2015) Exploring iot application using raspberry pi. *Int J Comput Netw Appl* 2:27–34
18. Rukhin A, Soto J, Nechvatal J, Smid M, Barker E (2001) A statistical test suite for random and pseudorandom number generators for cryptographic applications. (Booz-allen, 2001)
19. Bassham III L, Rukhin A, Soto J, Nechvatal J, Smid M, Barker E, Leigh S, Levenson M, Vangel M, Banks D et al (2010) Sp 800–22 rev. 1a. a statistical test suite for random and pseudorandom number generators for cryptographic applications. (National Institute of Standards Technology, 2010)

A New Authentication Protocol for Hardware-Based Authentication Systems in an IoT Environment



Rohit Pilla and Kurunandan Jain

Abstract Many companies have recently introduced smart IoT and wearable devices. Most of these devices have sensors that interact with them as well as features that enable them to connect to the Internet. However, if proper security and safety are not considered, the extent of damage will be greatly increased. This research work proposes a novel authentication scheme for two devices that can mutually authenticate and use dynamic keys for encryption and decryption based on the AES-GCM authenticated encryption algorithm. The security framework of the proposed scheme demonstrates that it is resistant to a variety of common attacks and enhances security in an IoT environment.

Keyword AES-GCM · Cryptography · Dynamic keys generation

1 Introduction

IoT has attracted a significant interest from both industry and academia. Recently, industries have made significant progress in developing new applications. Each domain has its own set of standards, goals, challenges, and security requirements. IoT ensures the system's confidentiality, integrity, and availability for various network components and parts.

Authenticated encryption techniques are a subset of symmetric key cryptographic algorithms that ensure the confidentiality and integrity of data. The protection of data from being disclosed without permission is referred to as confidentiality, whereas authenticity refers to the integrity of data and the authentication of its source. Some applications require the handling of extra data, such as packet headers, which require

R. Pilla (✉) · K. Jain

Cybersecurity Systems and Networks, Amrita Vishwa Vidyapeetam, Kollam, Kerala 690525, India
e-mail: amenp2csn20024@am.students.amrita.edu

K. Jain

e-mail: kurunandanj@am.amrita.edu

authentication but not encryption. This work will focus on authenticated encryption with associated data (AEAD), a broad classification for such schemes.

Authenticated encryption with associated data (AEAD) [1] refers to cryptographic primitives that may be used to solve a range of problems. The AEAD methods encrypt a part of the message and leave the rest unencrypted, but the whole packet gets authenticated. AEAD encryption is a symmetric encryption method for data transferred in packets that enables secrecy and data authentication via a single software interface. AEAD techniques are simpler and more efficient than separate encryption and authentication algorithms and use fewer resources. It also avoids key flaws in combining encryption and authentication, which have led to several real-world assaults against protocols and programs such as SSL/TLS. Key management is the process of managing the lifespan of keys that involves the production, exchange, storage, replacement, and deletion of cryptographic keys. Key administrative procedures are structured as either centralized or decentralized. The key distribution center (KDC) is shared among the bunch regulators, which is often formed of a progressive structure. For scattered conventions, hubs operate together to guarantee crucial administration activities such as key creation, circulation, reestablishment, and denial. This study presents a machine-to-machine authentication procedure that is end-to-end protected by AES-GCM encryption.

Since IoT devices are frequently low-cost and simple, with limited CPU, memory, and energy resources, implementing authentication methods for them is a difficult and time-consuming process. As a result, any security protocol or application designed for IoT devices should be as simple as possible in terms of processing complexity and energy consumption [2, 3]. Furthermore, since this single factor is easily intercepted and obtained by adversaries, most of the one-factor authentication solutions presented in the literature fail to prevent device authentication attacks [4].

The main research contributions made in this paper are as follows:

- A suitable authentication protocol is developed for device-to-device mutual authentication.
- Dynamic key generation algorithm is implemented in both the devices [5].
- A performance analysis has been carried out on the proposed scheme, which use lesser bytes compared to other protocols and security methodology resilient to various common attacks.

2 Related Work

In recent years, a great variety of two-factor authentication approaches have been proposed. However, most of these approaches [2, 6, 7] are concentrated on the user. These methods employ passwords and smart cards/devices as two-factor security. Since smart cards are not impervious to tampering, these systems are commonly

exposed to different physical assaults. On the other hand, a few interesting PUF-based authentication techniques have recently been suggested for IoT systems [8–11]. However, most of them are based on computationally inefficient public-key systems.

Anshel [12] suggested a lightweight key agreement protocol based on AE. The protocol performs well on low-cost devices by outperforming the previous public-key-based alternatives. E-multiplication is a new operation introduced by AE. An adversary cannot deduce the input from a given output since it is a one-way function. The difficulty of E-multiplication rises in lockstep with a different level of security. It enables the AAGL protocol to achieve considerable gains in efficiency. The mathematical foundations of AE are not the same as those of classic public-key cryptosystems. The braid group and its features are used in the AAGL protocol to create an algebraic erase concept.

On the other hand, in [13], the authors presented an RFID-based authentication architecture for distributed IoT applications, which they believe will be helpful in future innovative city contexts. Compared to the previous techniques, the suggested protocol (which is not dependent on PLS) is both lightweight and efficient. It also provides RFID tags to enable secrecy, anonymity, as well as safe location. Similarly, [14] presents a lightweight authentication system for IoT-enabled devices in distributed cloud computing environment. The authors offer an architecture with a smartcard-based authentication protocol, which allow registered users to access all private information stored on private cloud servers safely. In addition, the authors in [1] address the three-layered data flow architecture for fog computing and propose many unique architectures inside the fog computing paradigm, including energy lattices, MediFog, UXFog, linked parking systems, and FoAgro.

For authentication, the authors of [15] use discrete algorithm called zero-knowledge proof of knowledge (ZKPK). Physical security is provided by a combination of a PUF and user password. However, this suggested technique does not require any stored secrets, and it is less effective for IoT devices since it requires a user password for authentication each time [16]. Furthermore, because of the discrete logarithm ZKPK, their proposed approach is more computationally complex.

The authors of [17] created a foldable architecture for AES encryption/decryption by using pre-computed keys and researched the minute features of Spartan-II FPGAs to efficiently implement the S-box on block RAMs (BRAMs). The MixColumn/InvMixColumns idea is implemented using shared logic resources, consuming roughly 222 slices with a throughput of 166 Mbps.

3 Proposed Scheme

The proposed scheme mainly consists of two devices, where one device acts as a verifier, while the other tries to authenticate and vice versa, as shown in the Fig. 1.

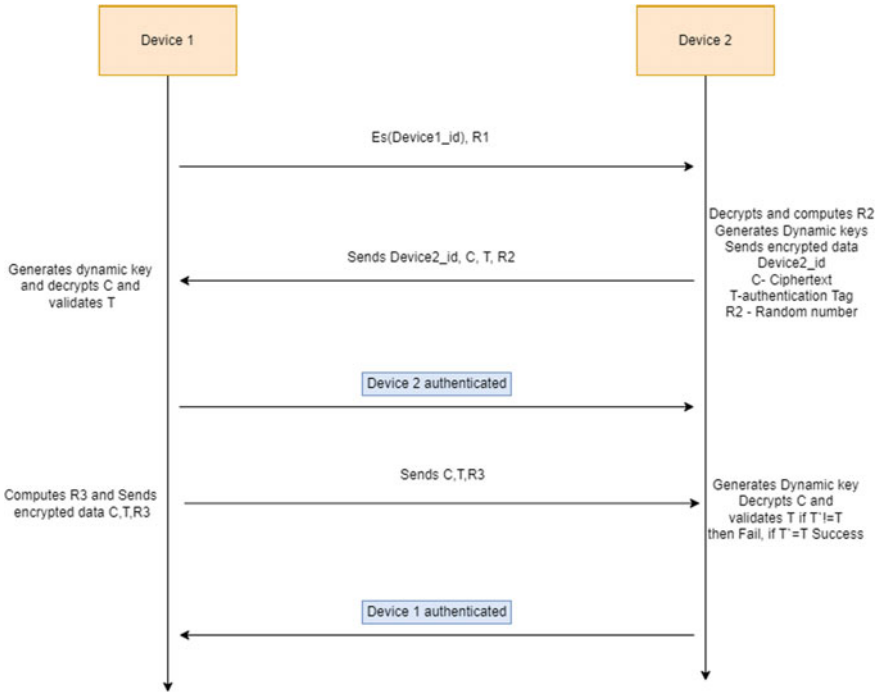


Fig. 1 Proposed scheme

Prerequisites:

1. Ciphertext—128-bit
2. Key size—256-bit
3. IV—96-bit
4. Plain text—128-bit
5. Random numbers—64-bit

Message 1: Device 1 → 2:

Initially, the first device starts transmitting its Device-1ID and random cryptographic nonce r_1 by encrypting them using its 256-bit pre-secret key with AES-GCM. In this case, this can be a secret key assigned to each device by the manufacturer, assuming it as device 2.

- Step 1: Encrypts Device-1ID using secret key $E = E_s(id)$
- Step 2: Computes a pseudo-random number r_1
- Step 3: Transmits E and r_1

Message 2: Device 2 → 1:

Device 2 receives this data and tries to decrypt it using AES-GCM with the secret key used by device 1. Since r_1 is a random nonce generated by device one. After

receiving Device-1ID and $r1$, device 2 checks whether this pair (key, nonce) is repeated in the previous messages and continues if no match is detected. If the match is detected, the process gets terminated. Device 2 starts generating the random nonce $r2$, uses a dynamic key generator algorithm, and generates key K . Initially, a hash is calculated using a block cipher of 128-bit zeros using the K key produces H . The block cipher B_o is generated when the random IV is 96-bit size, and it is appended to the padding string $0^{31}||1$, and 32-bit incrementing function is applied to block cipher. Using block cipher encryption, we calculated ciphertext C_b , and R is a time stamp bound to C_b . Q is calculated by the GHASH function, which takes AAD “A,” and random numbers $r1$, $r2$, and the size of A and C are 64-bit variables. The tag is calculated by MSB, which is the most significant bit, and the output is encrypted with the previously generated block in step 2. The result is then reduced to the given tag length to create the authentication tag. The output consists of the ciphertext and tag. Transmits ciphertext, tag, and device2_ID.

Step 1: Define hash subkey $H = E(K, 0^{128})$

Step 2: Defining a block size B_o

$\text{len}(\text{IV} = 96 \text{ ten } B_o = \text{IV} || 0^{31} || 1$

Step 3: Generating counter block

$C_b = E(K, \text{inc}_{32}(B_o))$

$P = E(K, C_b, R)$

Step 4: $Q = \text{GHASH}_H(A || \text{len}(A) || \text{len}(C) || C || 0^{128} || r1 || r2)$

Step 5: Auth tag $T = \text{MSB}(E(K, B_o, Q))$

Step 6: Send $C, T, R2$

Device 2 Authentication:

Device 1 computes key K by using a dynamic key generation algorithm before validating the IV, ciphertext, and authentication tag, their sizes will be examined in step 1. If the length matches the specified parameters, it either continues or fails. Step 2 uses key K to execute block cipher encryption on the “zero” block and creates H values. When the size of the IV matches, step 3 builds a pre-block counter. It generates a block in the same way as encryption does. Step 4 constructs the decryption block using a 32-bit incrementing function on the counter block and retrieves the plain text by decrypting the block using key K . Steps 5 and 6 combine the AAD and ciphertext strings to generate 64-bit representation strings. To generate the GHASH result, they are concatenated with random integers $r2$, $r1$, and “zero” bits. MSB receives the GHASH output. The output is then reduced to the tag length specified in order to generate the authentication tag. When both tags match, the plaintext is revealed.

Step 1: If the length of ciphertext, AAD, and IV are not supported, then fail. If $\text{len}(T) \neq t$, then fail

Step 2: $H = H^{-1}(K, 0^{128})$

Step 3: Defining a block size B_o^1

$\text{len}(\text{IV}) = 96 \text{ then } B_o^1 = \text{IV} || 0^{31} || 1$

Step 4: Generating counter block

$P = E(K, \text{inc}_{32}(B_o^1), C, R)$

Step 5: $Q^1 = GHash_H (A \parallel \text{len}(A) \parallel \text{len}(C) \parallel C \parallel 0^{128} \parallel r1 \parallel r2)$

Step 6: $T^1 = \text{MSB}(E(B_o^1, Q^1))$

Step 7: Check $T = T^1$ if they both match return P or return fail.

Furthermore, device 1 verifies the received data and tag. The tag is invalid if the attacker modifies the data packet while in transit since the timestamp is concatenated with the encrypted data. We can achieve confidentiality and data integrity and device-to-device authentication.

Message 3: Device 1 \rightarrow 2:

Similarly, for authenticated encryption function, device 1 computes the pseudo-random number R3. Device 1 performs similar steps, but key K, IV, R, A, and C these parameters vary or might be different in this case because a new key is used for each input message, and IV is incremented by 1 for each message to avoid reuse of IV attacks and to the ciphertext timestamp R and is bound so that when an attacker tries to modify the ciphertext, the tag will not get validated due to modification. After encrypting the data, device 1 sends C, T, and R3 to device 2 for verification.

Step 1: Define hash subkey $H = H_k (0^{128})$

Step 2: Defining a block size B_1

$\text{len}(\text{IV}) = 96$ then $B_1 = \text{IV} \parallel 0^{31} \parallel 1$

Step 3: Generating counter block

$C_b = \text{GCTR}_k (\text{inc}_{32} (B_1), P, R)$

Step 4: $Q = GHash_H (A \parallel \text{len}(A) \parallel \text{len}(C) \parallel C \parallel 0^{128} \parallel r2 \parallel r3)$

Step 5: Auth tag $T = \text{MSB}(GCTR_k (B_1, Q))$

Step 6: send C, T, R3

Device 1 Authentication:

Device 2 receives C, T, and R3 by using the AES-GCM decryption algorithm. Device 2 tries to validate the authentication tag. Device 2 performs a dynamic key generator algorithm to retrieve key K. The following steps represent the decryption and verification process. If the $T = T^1$, device 1 is authenticated.

Step 1: If the length of ciphertext, AAD, IV are not supported, then fail. If $\text{len}(T) \neq t$, then fail

Step 2: $H = H_k^1 (0^{128})$

Step 3: Defining a block size B_1^1

$\text{len}(\text{IV}) = 96$ then $B_1^1 = \text{IV} \parallel 0^{31} \parallel 1$

Step 4: Generating counter block

$P = \text{GCTR}_k (\text{inc}_{32} (B_1^1), C, R)$

Step 5: $Q^1 = GHash_H (A \parallel \text{len}(A) \parallel \text{len}(C) \parallel C \parallel 0^{128} \parallel r2 \parallel r3)$

Step 6: $T^1 = \text{MSB}(GCTR_k (B_1^1, Q^1))$

Step 7: Check $T = T^1$ if they both match return P or return fail.

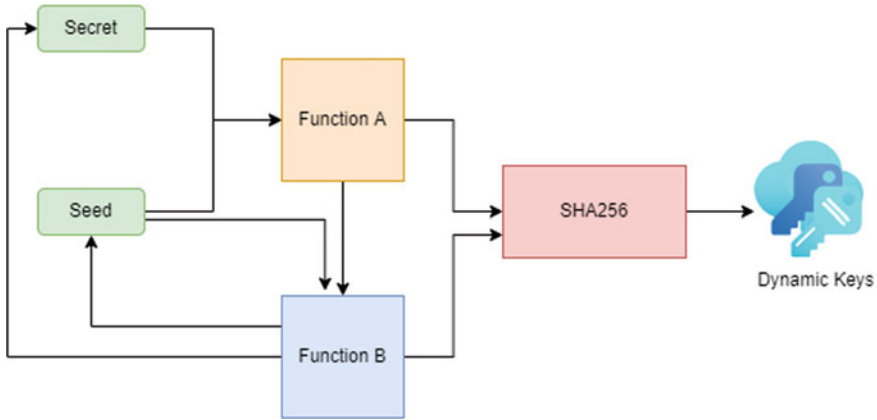


Fig. 2 Dynamic key generation algorithm [5]

3.1 Overview of Dynamic Key Generation Phase

The dynamic key generator phase is really adapted from the authors [5]. The algorithm works with two inputs, such as secret and seed which are stored in a non-volatile memory location. There are two functions, as shown in Fig. 2. They are built-in with some logical and bitwise operations. This algorithm is the cyclic process, and lastly, the SHA-256 function generates 256-bit dynamic keys as output. This algorithm is implemented in both devices. By doing this, we reduce the burden of having a back-end server to compute and key exchange.

4 Security Methodology

4.1 Man-In-The-Middle Attacks

In a man-in-the-middle attack, an attacker will intercept the communication between the verifier and the authenticating device. They will try to record and watch the messages exchanged between both devices in the authentication phase, and they will try to come up with a rogue device that is a real device and tries to authenticate with the verifier by re-transmitting the messages. However, here in our proposed scheme, at the encryption phase, we bound the random IV and time stamp together with ciphertext because when an attacker tries to modify the ciphertext in between at the verifier side that authentication tag is not validated since the packet has been modified in transit. We can assert that the suggested system is robust to this form of attack.

4.2 *Replay Attacks*

After every successful session, the two IoT devices will change their random IV and key. There will be a new value generated uniquely and different from the previously used sessions. During authentication, when the attacker tries to intercept and eavesdrop on the entire authentication session and tries to replay by using the same key and data in the next session, they will be failed because, in this protocol, we are using a dynamic key generator which generates a unique for every session, and the session key value is bound to ciphertext; however, the data cannot be modified. If it is modified, it will not get authenticated.

4.3 *Side-Channel Attacks*

Side-channel security issues are well-known for shared key encryption. In the protocol, we employ AES-GCM 256-bit for authenticated encryption and authenticated decryption. Suppose that an attacker attempts to get the information by observing and analyzing and utilizing side-channel assaults; as we mentioned, we are employing a dynamic key generation strategy. Here, all keys are produced by executing additional logical and bitwise operations. Due to this complexity, it is difficult to get any key, and with the uniqueness, of the keys, we do not reuse the same key everywhere in the session. Hence, it is not feasible to recover sensitive information.

4.4 *DDoS Attacks*

DDoS is a sort of cyber assault where the attackers attempt to flood the server with multiple requests from many devices to the server. By doing this, the server is overburdened with requests. Our protocol does not allocate any resources before authentication, hence this attack is impossible.

4.5 *Brute Force Attacks*

Brute force attack is typically used to break passwords and encryption keys. The suggested approach employs AES-GCM 256-bit for encryption and decryption; therefore, a 256-bit key is being utilized at both ends to encrypt and decode the data. The amount of options to break the encryption key is approximately 2^{256} , which takes a lot of time to crack. The longer the key, the tougher for an attacker to break.

Version	Tyoe
Instruction set	ARMvVv8-A (64/32-bit)
SoC	Broadcom BCM2837B0
FPU	VFPv4 + NEON
CPU	4x Cortex-A53 1.4 GHz
GPU	Broadcom VideoCore IV @ 250 MHz
Memory (SDRAM)	1GB

Fig. 3 Raspberry Pi 3 model B configuration

5 Performance Analysis

This section evaluates the proposed technique and compares them with different specified papers based on AES-GCM 256-bit encryption and decryption time execution and communication cost to complete authentication.

5.1 Test Configuration

The following is the test configuration as shown in Fig. 3, where the entire code is developed in Python 3.6 using Raspberry Pi 3 since it is mostly used in IoT and has all supported libraries installed.

5.2 Encryption and Decryption Time Execution

It is generally known that AES-GCM enables authenticated encryption, which protects data integrity and secrecy, and the proposed technique has employed a 256-bit version of AES-GCM, where the results are compared with other similar studies. As you can see in the Figs. 4 and 5, the proposed system consumes less resources for encryption and decryption. It should also be noted that as message size increases, so does the time required to complete encryption and message authentication. This is normal as the number of bytes increases, so does the time required to process them.

5.3 Communication Cost

The communication cost of the proposed scheme uses three messages to perform mutual authentication between two devices. The proposed model is compared with related work of authors [5, 18–20], as shown in Fig. 6 you can see that the proposed model uses lesser bytes.

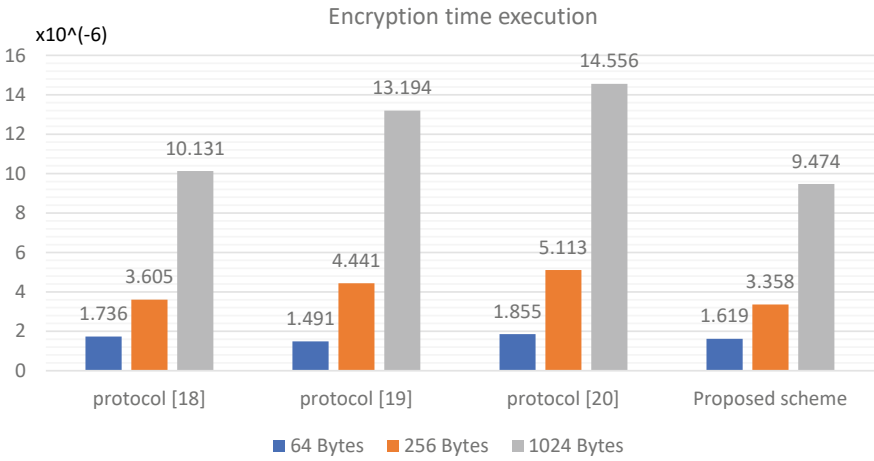


Fig. 4 Encryption time execution

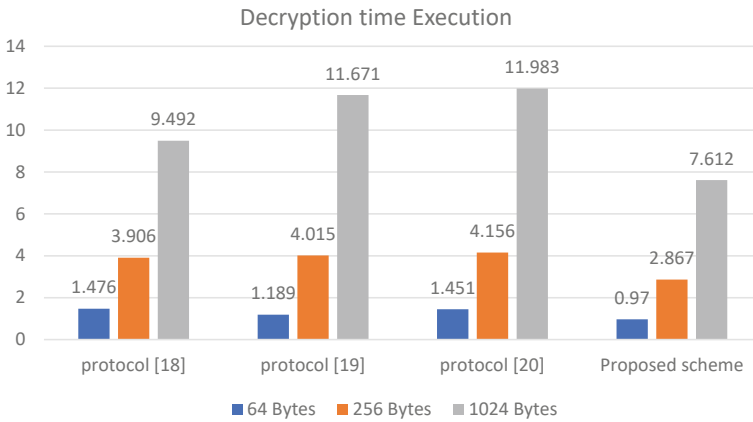


Fig. 5 Decryption time execution

Scheme	No. of messages	No. of Bytes
Protocol in [18]	3	1216
Protocol in [19]	4	1856
Protocol in [20]	4	2752
Protocol in [21]	4	2080
Proposed scheme	3	832

Fig. 6 Communication cost

6 Conclusion

This paper has successfully proposed a novel mutual authentication scheme suitable for resource-constrained environments like IoT. This study has reduced the use of back-end servers for issuing keys and verifying devices in a wireless environment by deploying a dynamic key generation protocol that allows each device to generate its own dynamic keys for authentication and session key establishment. By using the AES-GCM protocol, the proposed study has ensured data integrity and confidentiality. The proposed scheme's security methodology has successfully demonstrated that it is resistant to various attacks. The performance analysis domain consists of encryption and decryption time of the proposed protocol and communication cost of the scheme, which is less when compared with other protocols. With this, the proposed study ensures that the scheme is suitable for various IoT devices due to its less resources and robust authentication scheme.

References

1. Kunal S, Saha A, Amin R (2019) An overview of cloud-fog computing: Architectures, applications with security challenges. In: Security and privacy, vol. 2(4), p e72
2. Han L et al. (2016) An efficient and secure two-factor authentication scheme using elliptic curve cryptosystems. In: Peer-to-peer networking and application, vol. 11(12)
3. Aman MN, Chua KC, Sikdar B (2017) Secure data provenance for the internet of things. In: Proceedings of the 3rd ACM international workshop on IoT privacy, trust, and security. pp 11–14
4. Aman MN, Chua KC, Sikdar B (2017) Physically secure mutual authentication for IoT. In: 2017 IEEE conference on dependable and secure computing. IEEE, pp 310–317
5. Pothumarti R, Jain K, Krishnan P (2021) A lightweight authentication scheme for 5G mobile communications: a dynamic key approach. J Ambient Intell Humanized Comput 1–19
6. Amin R et al (2017) A two-factor RSA-based robust authentication system for multiserver environments. In: Security and communication networks 2017
7. Xie Q et al (2017) Provably secure dynamic ID-based anonymous two-factor authenticated key exchange protocol with extended security model. In: IEEE transactions on information forensics and security, vol. 12.6, pp 1382–1392
8. Kim Y-P, Yoo S, Yoo C (2015) DAoT: Dynamic and energy-aware authentication for smart home appliances in Internet of Things. In: 2015 IEEE international conference on consumer electronics (ICCE). IEEE, pp 196–197
9. Petrov V et al (2014) Towards the era of wireless keys: How the IoT can change authentication paradigm. In: 2014 IEEE world forum on internet of things (WF-IoT). IEEE, pp 51–56
10. Porambage P et al (2014) Two-phase authentication protocol for wireless sensor networks in distributed IoT applications. In: 2014 IEEE wireless communications and networking conference (WCNC). IEEE, pp 2728–2733
11. VL Shivraj et al (2015) One time password authentication scheme based on elliptic curves for Internet of Things (IoT). In: 2015 5th national symposium on information technology: towards new smart world (NSITNSW). IEEE, pp 1–6
12. Anshel I et al (2007) Key agreement, the algebraic Eraser™, and lightweight cryptography. Contemp Math 418:1–34
13. Gope P et al (2018) Lightweight and privacy-preserving RFID authentication scheme for distributed IoT infrastructure with secure localization services for smart city environment. Future Gener Comput Syst 83:629–637

14. Amin R et al (2018) A light weight authentication protocol for IoT-enabled devices in distributed cloud computing environment. *Futur Gener Comput Syst* 78:1005–1019
15. Aman MN, Chua KC, Sikdar B (2017) Mutual authentication in IoT systems using physical unclonable functions. *IEEE Internet Things J* 4(5):1327–1340
16. Frikken KB, Blanton M, Atallah MJ (2009) Robust authentication using physically unclonable functions. In: *International conference on information security*. Springer, Berlin, pp 262–277
17. Chodowiec P, Gaj K (2003) Very compact FPGA implementation of the AES algorithm. In: *International workshop on cryptographic hardware and embedded systems*. Springer, Berlin, pp 319–333
18. Das K, Wazid M, Kumar N, Khan M, Choo K, Park Y (2018) Design of secure and lightweight authentication protocol for wearable devices environment. *IEEE J Biomed Health Inform* 22(4):1310–1322
19. Li X et al (2018) A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments. *J Netw Comput Appl* 103:194–204
20. Das A et al (2016) Provably secure user authentication and key agreement scheme for wireless sensor networks. *Secur Commun Netw* 9(16):3670–3687

Hybrid Cell Selection Mechanism for V2X Handover



Faiza Al Harthi and Abderezak Touzene

Abstract The integration and implementation of vehicle-to-everything (V2X) in the transportation sector can potentially improve vehicle, passenger and traffic safety, and efficiency. This benefit, which is one of the objectives of intelligent transport systems, faces challenges and difficulties parallel to those that hinder the realization of the benefits promoted by 5G mobile networks and beyond. High-speed user equipment (UE) in the form of vehicles must be able to seamlessly connect to various entities in the network while avoiding common handover (HO) issues such as frequent and unnecessary HOs and radio link failures resulting in the degradation of quality of service (QoS). In this paper, we propose a hybrid cell selection mechanism based on signal quality, coverage, and bandwidth to reduce issues and optimize operations concerning HO that will boost V2X network performance.

Keywords Handover · Signal strength · Coverage · Bandwidth · Hybrid algorithm

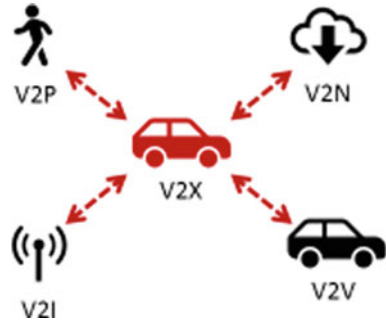
1 Introduction

A seamless user experience is the end goal of 5G networks, and for network communications to be smooth and uninterrupted, devices must be near a base station (BS) to increase signal strength and reduce interference. In the event of an unavailable BS, device-to-device (D2D) communication has been identified as the best solution for signal availability and BS proximity problems [1]. D2D is a promising component in 5G networks that enhances energy efficiency, throughput, latency, and spectrum utilization in cellular networks. This is clearly evident when it comes to the deployment of D2D use cases such as V2X where with the high velocity of UE, reliable and less frequent change connections are crucial to save people's lives [2].

F. Al Harthi (✉) · A. Touzene
Department of Computer Science, Sultan Qaboos University, Muscat, Sultanate of Oman
e-mail: S135943@student.squ.edu.om

A. Touzene
e-mail: touzene@squ.edu.om

Fig. 1 V2X communication



V2X communications, as shown in Fig. 1, are one of the evolving technologies that enable vehicles to connect to each other (V2V), to the road environment including pedestrians (V2P) and infrastructure (V2I), and to cellular networks (V2N). The Federal Communications Commission (FCC), in 1999, allocated a spectrum in the 5.9 GHz band for the purpose of intelligent transport services (ITSs) that resulted in extensive research for the implementation of V2X communications. In 2010, an IEEE radio standard for V2X based on 802.11p technology called dedicated short-range communication (DSRC) was released. DSRC allows communication through PC5, which is also known as Sidelink. Another type of connection is through cellular network (C-V2X). Both technologies operate in the 5.9 GHz band, although C-V2X can operate in the operator's licensed band (Uu interface). However, both technologies fail to provide the required performance when QoS in V2X applications is rigorous [3]. V2X use case requirements include speed, which can reach a maximum of 500 km/h without limitations, communication range, or latency, which shall be 100 ms between UEs in V2V/V2P/V2I deployment. These requirements characterize V2X as a high mobility use case [4].

Highly mobile UEs in a heterogeneous D2D network rely on device centric rather than a network centric approach to connectivity. It enables the UE to maintain a direct connection with other UEs with or without the presence of BS. The devices have the capability for network setup and control [5]. This is especially useful since highly mobile devices need to switch from one network to another when coverage to maintain communication fades. In order to avoid connection termination, HO mechanisms allow the active connection to be transferred from weak network coverage to another stronger cell. HO in wireless networks can be achieved between various access technologies, operators, and cells. HO is one of the improved mechanisms that plays a vital role during communication in heterogeneous networks (HetNets). HetNets are defined by 3GPP LTE Advance standard as the paradigm to handle varied wireless coverage zones, cell sizes, and capacities. This type of network is made up of different types of small base station cells (SBS) and low-power nodes such as microcells, picocells, femtocells, and relays [6]. However, this model poses a major challenge where densification of SBS increases cell edges and inter-cellular interaction. It is estimated that small cells (SCs) can make the handover happen every 11.6 s [7]. This performance is challenged, especially in V2X use cases where UEs

are fast moving and get in and out of coverage areas. This can negatively impact V2X performance as frequent connection switching can occur. Furthermore, a solution is required that would make this type of network handle situations well through automated means. The HO challenges can be solved by having the vehicle choose the ideal cell before the handover decision. This can significantly reduce HO issues and improve the network QoS [8].

2 Related Work

Various studies have provided the basis for this research. It was mentioned in [9] that in smart solutions for handling different technology characteristics, stringent QoS requirements and advanced V2X use cases can be addressed by context awareness. In [10], the mechanism implements context awareness using the previously stored user contexts and behavioral information to predict user mobility and skip unnecessary handover process. Although, this mechanism improves the overall HO process, and it requires a type of computing platform and processing technique that caters for V2X deployment in femtocells. In [11], the study focused on V2X communications solely on macrocell and femtocell infrastructure. They proposed a BS selection scheme using Markov decision policy for V2X communication. The mechanism uses received signal strength (RSS) prediction to select a BS.

Device-to-device (D2D), which was incorporated into the 3GPP Release 12 standards as proximity services (ProSe) [12], was created with the purpose of supporting the Internet of things (IoT) in 5G networks. Article [13] studies ProSe and V2X connections using Sidelink, which streamlines D2D interactions across cellular networks. The demodulation reference signal (DMRS) uses power-normalized-correlation (PNC) to detect active peers during mobility. The distributed system monitors out of coverage (mode 2 and 4) and network-assisted communications (mode 1 and 3). In the first context, discovery periods are an Aloha-like protocol, and in the second, a polling-like protocol leveraging media access control (MAC) layer or physical collision model. Only Sidelink is evaluated for network selection. In [14], the GPS system is investigated and compared with several location algorithms. A multi-stage localization approach is proposed that makes use of neighboring UEs in order to locate a target UE in harsh channel conditions. The main challenge of this channel is that UE has poor BS localization. Using neighbor discovery, the two nearest nearby UEs are determined. Distance calculations provide two angle of arrival (AoA) alternatives for target UE. Oriented beamforming is used to differentiate the accurate AoA estimate by forming small arcs around the two AoA possibilities. Using the predicted AoA and distance to adjacent UEs, the user's position is computed. The Poisson line Cox point process for vehicle locations and planar point process for users locations is used to model vehicle location [15]. In this context, the author also determines the effective rate provided to the typical user to address the impact of vehicle broadcast on the cellular network. In this case, it is associated with the closest BS, resulting on more HO. Due to the semi-static scheduling mechanism, the random

access procedure used by UE to initiate a data transfer has a low resource efficiency. With the large access requirements of the 5G network, this problem would be exacerbated. Intelligent access channel allocation must be used to efficiently decrease resource consumption while ensuring access latency and reliability, which needs the BS to precisely predict the number of UEs conducting random access. The author [16] developed long short-term memory (LSTM) to process the time series capacity of the data, making it suited for tackling the active UE number estimate problem.

In [17], it was pointed out that implementing vertical HO allows the use of different access technologies, multiple network interfaces, multiple QoS parameters, and multiple connections that can be changed. A common approach for cell selection in mobile communication is based on the signal interference noise ratio (SNIR) or received signal strength interference (RSSI) [18]. However, implementing the HO process using the same set of user parameters in HetNets could degrade mobility and performance [19]. This is quite a challenge in V2X use cases where most UEs are fast moving and require different QoS parameters (user preferred) dictated by application requirements. Since 5G is expected to handle V2X applications, HO mechanisms that were usually developed for centralized implementations will not be suitable for the distributed and hybrid HO requirements of V2X networks [20]. To effectively implement HO between cells and handle diverse requirements in HetNets and V2X, it is crucial to achieve optimal context interpretation [10].

Cell selection in HetNets must consider a number of parameters according to network context and use cases. The selected parameters need further to be optimized [18]. This is valid in HetNets infrastructure where different SBS deployed with a variation in terms of power transmission, coverage, and number of connections are accepted. Many recent studies (see Table 1) have focused on using more than one parameter for cell selection. The cell selection methods in [10, 18, 21, 23, 25, 28] result in improvement in HO performance, however, these studies did not consider the V2X use case. Applying this method in a different topology can lead to more frequent HO. In [22], RSSI, load, and speed are used as cell selection parameters, although only RSSI is considered for handover decisions. In [26], network weight is used to assess network performance and user preferences to choose the optimal cell selection, and the hierarchy analysis method is used to assign the appropriate weight to each selection metric. Load balancing is neglected in this study. The same limitation is applicable for [27]. Overloaded cells can affect network performance, resulting in more link failures.

Overall, in the ideal design of a network, supplementing less setup cost while providing good signal strength coverage is considered an ideal deployment approach for 5G BSs [21]. Enough resources at the target cell must be reserved to ensure that the connection will not be terminated during handover [8]. Based on that, this paper considers three parameters: link quality (RSSI), coverage, and bandwidth for ensuring efficient seamless connectivity for handover. The handover decision is based on an optimization algorithm. This approach is considered to control handover issues such as frequent and unnecessary HOs and link failures.

Table 1 Link parameters

Paper	Parameters																				
	Link Quality						Speed	Direction/User trajectory/mobility	Bandwidth/Channel Capacity	Coverage/Distance	Location	Power transmission	Channel Model	Power Consumption	Displacement angle	Packet Loss Rate	Transmission Delay Time	User Preference	Traffic Load	Dwell Time	
	RSRP	RSSI/RSS	SINR	Propagation	Signal Intensity	Jitter															
[10]	Yellow						Red					Yellow	Grey							Brown	
[11]		Blue					Red	Orange	Green												
[21]									Blue	Grey											
[18]		Blue							Blue				Green								
[22]		Blue					Red		Green												
[23]								Orange													
[24]								Orange						Brown							
[25]									Green			Yellow									
[26]									Green			Yellow			Dark Blue	Yellow					
[27]					Pink				Green				Dark Green		Dark Blue	Yellow	Grey				
[28]			Green						Green												
[29]		Blue																			Blue

3 Hybrid Cell Selection Method

V2X communications can be enhanced by selecting the best BSs based on QoS metrics and/or UE preferences. It was pointed out in [30] that prior to network selection, information about BSs should be gathered based on context and user preferences. The selected BS information is processed and stored in the UE buffer for HO use, which would lessen the need for centralized BS processing and best serve the UE preference according to current context, application requirements, or presence of

vertical networks. A periodic update of the stored selected BS information is based on operational levels of the required parameters [31]. Figure 2 illustrates the overall research flow by pre-processing and segmenting the handover mechanism to provide seamless connectivity based on the ideal selection of BS using signal quality, cell coverage, and bandwidth as the parameters and hybrid algorithm. Phase II and Phase III calculation are restricted to the output of Phase I.

These phases can be explained as follows:

Radio signal strength can be affected by many obstructions while propagating from transmitter to receiver. This can cause signal attenuation and affect the network performance. Even though radio frequency is developed with the ability to penetrate surfaces, the signal strength is degraded while moving across reinforced concrete in walls or changing in the weather. Moreover, the interference that happens between two waves is considered a killer for signal communication. To ensure the availability of the network, the vehicle must choose the best signal quality among the signals.

Coverage measures how broad an area around a wireless transmitter has sufficient signal strength for wireless devices. The findings are used by vehicles to associate with the best coverage BS that allows them to travel for long distances without the need for changing the connection. In many studies, the long distance is measured based on the velocity of the vehicle. A macrocell is used for the association of a high velocity vehicle. Middle and low vehicle speeds are associated with small cells, and in V2X, these are referred to as femtocells. However, in this study, connections are not segregated based on the level of speed. For instance, the term ‘long distance’ here means the coverage area.

The number of connected UEs is increased in HetNets. SCs in HetNets have different power transmission and link capacity. At certain times, the SC may reach its maximum capacity, which results in link failure. Therefore, bandwidth estimation is needed before the selection of a BS for handover. If the number of users is more in the BS, the bandwidth will be reduced. To calculate the number of connections, the researcher must monitor the BS interface which is carrying different types of packets. For instance, the source vehicle can understand the average bandwidth of the BS which is available for transferring the connection.

The entire system model shown in Fig. 3 is converted into a connected graph. The aforementioned parameters, link quality, coverage, and bandwidth are used for the selection of the context-based BS. The author has introduced an optimization algorithm to select the best BS among the other BSs which helps to reduce the

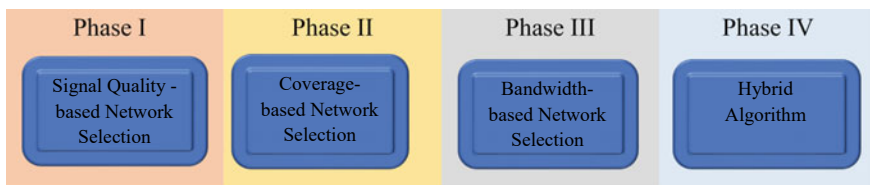


Fig. 2 Overall research workflow

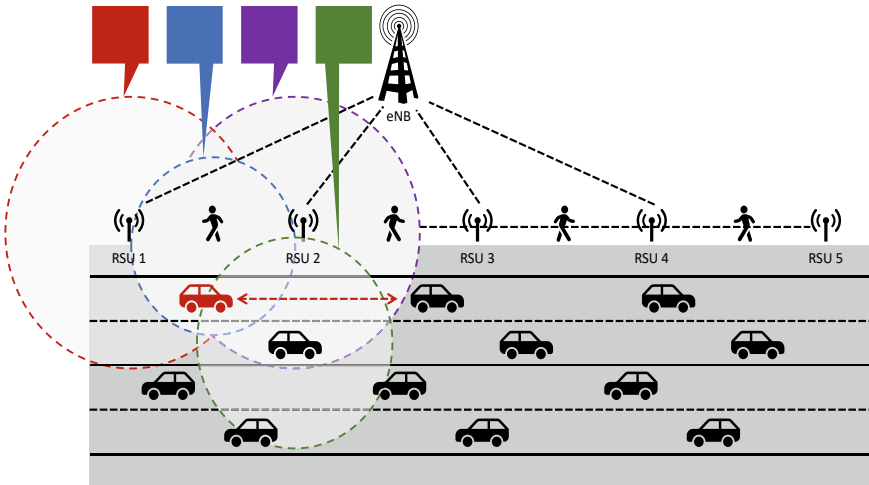


Fig. 3 BS selection model

number of HOs and latency and increase signal quality to V2X. Moreover, this algorithm provides an acyclic graph to eliminate loops between BSs.

3.1 System Model

The system model, as shown in Fig. 3, consists of a number of vehicles, pedestrians, road side units (RSUs), and a macrocell that represents V2X communications. It can be used in any small cells such as Wi-Fi, WiMax, and WLANs which are using different transmission power and coverage areas. There are four overlapping cells (C1, C2, C3, and C4) within the area of the red car (source vehicle) that may require different parameters. The source vehicle moves forward at a random speed. The macrocell (gNB) has a link connection with RSUs through U_u interfaces. A link connection can exist between the RSUs and vehicles of different use cases (V2V, V2I, V2P, and V2N) using PC5 Sidelink interface.

3.2 Signal Quality-Based Network Selection

Figure 4 depicts a simplified handover situation in which the source vehicle is linked to RSU1 and is driving toward the coverage zone of RSU2, pedestrian (P), and adjacent vehicle (AV) which are considered base stations. The source vehicle continuously monitors the signal strength of these BSs (RSU2, P, and AV) with the purpose of storing the BSs with the best signal quality (illustrated in signal strength algorithm).

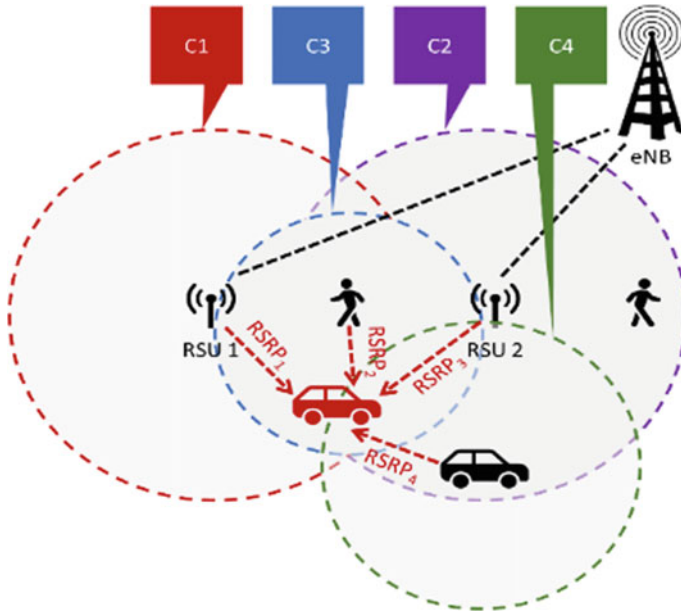


Fig. 4 Signal quality-based tower selection

If the detected signal strength of the V2P in the intersection region exceeds that of the RSU1, and the V2P can establish the connection to provide the resources required by the source vehicle, a cell selection of V2P will be considered as the best for the handover process based on the signal quality referred to in Table 2.

The surrounding environment and the maximum radius of the cell, both of which have a significant influence on the received signal, are two of the most crucial elements that determine wireless coverage. The received signal characteristics are primarily influenced by three factors: multipath fading, shadow fading, and path loss propagation. Although signal to interference and noise ratio (SINR) is a signal quality metric, it is not specified in 3GPP specifications and so is not provided to the network. The carrier received signal strength indicator RSSI indicates the average overall received power measured where $RSRP_n$ and $RSRP_p$ are the RSRPs of the neighboring and

Table 2 Signal quality [32]

RF signal condition	RSRP (dBm)	RSRQ (dBm)	SINR (dBm)
Excellent	$> = -80$	$> = -10$	$> = 20$
Good	-80 to -90	-10 to -15	13 to 20
Medium	-90 to -100	-15 to -20	0 to 13
Weak	$< = -100$	< -20	$< = 0$

servicing BSs, respectively. The serving cell’s offset is A3 offsets, and the cell individual offset between adjacent and servicing cells is the cell individual offset between the neighboring and servicing cells (CION, s). In this situation, however, the RSRP of the servicing cell—a femtocell in this case—drops below the threshold or the RSRP of neighboring cells only when the vehicle is inside the area of coverage [25].

RSRP1 is a servicing cell that is connected to the source vehicle, as indicated in Fig. 3. Due to its movement, the source vehicle discovered three more signals from different BS: RSRP2, RSRP3, and RSRP4. With the formula below (1), the car will choose the best signal quality among these signals. All received signals containing BS information are recorded in the vehicle’s buffer, which will be utilized in catastrophe scenarios.

$$RSRPN - RSRPs > A3offsets - CION, s \text{ or } RSRPN \geq RSRPs \tag{1}$$

where

- RSRPs Signal from source (serving) BS
- RSRPN Signal from neighboring BS
- A3 offsets Offset of the serving cell
- CION, s Cell individual offset between the neighboring and servicing cells.

The received signal strength indicator (RSSI) is a metric that indicates the overall received wide-band power in all symbols, including all thermal and interference noise. RSSI is not reported by vehicle to BS since it can be estimated from reference signal received quality (RSRQ) and RSRP reported by UE.

$$RSSI = \text{Serving cell power} + \text{interference power} + \text{noise}.$$

Reference signal received power (RSRP) is a variation of RSSI.

$$RSRP(\text{dBm}) = RSSI(\text{dBm}) - 10 \times \log(12 \times N).$$

$$\text{Reference signal received quality(RSRQ)} = RSRP / (RSSI/N).$$

where

- N = Number of physical resource blocks (PRBs).
- $RSSI$ = Noise + serving cell power + interference power during reference signal (RS) symbol.
- $RSRQ$ depends on serving cell power and the number of Tx antennas.

Signal Quality Algorithm

- Step 1: Keep decoding BSs (cells) DMRS (PSBCH, PSCCH, PSSCH) by measuring transmission power.
- Step 2: Vehicle must collect all sensing information of reserved resources and RSSI/RSRP measurements.

Step 3: The vehicle must exclude its connection from the common resource block (CBR) and make its available resource set.

Step 4: Select the Tx based with reliable BS based on the best signal strength.

Step 5: Sort the collected BS information according to signal quality from highest to lowest (BS ranking based on the current parameter).

Step 6: Select the top 50% BSs and store the information in the UE's buffer.

Step 7: Update the BS information in the buffer after T-time by collecting new BS information (Step 2) then proceeding to Step 4 to ensure the reliability of buffer information.

Step 8: If the number of BSs in the buffer is 0, reselection is triggered; go to Step 2.

3.3 Coverage-Based Network Selection

The source vehicle has to select best coverage BS based on large coverage area, which is calculated based on the height of the antenna and transmitter power. Furthermore, as shown in Fig. 3, the source car is getting signals from four different RSUs (RSU1, RSU2), pedestrian, and adjacent car. Out of four coverage intersections, RSU2 is providing more coverage to the source. The source vehicle can travel up to (D1) distance at (T1) time without need of handover. This maintains the connection for a long period. Once it reaches the radius of the coverage, it has to initiate the HO discovery process.

Coverage Area Calculation:

The antenna's height ' h ' is located at point ' S ' on the surface. ' O ' is the center of the antenna, and ' Re ' is the radius of the place. Let ' P ' be the point on the place at a distance of ' d ' from ' S '. Beyond the ' S ', the device cannot receive the signal from the transmitter ' T '. Point TP is tangent of the place, hence the height of the tower $SP = PT = d$. $PT = d$ as shown in Fig. 5.

ΔOPT is right triangle. As per Pythagoras theorem, $OT^2 = OP^2 + PT^2$

Therefore, $(Re + h)^2 = Re^2 + d^2$

$Re^2 + 2 Re h + h^2 = Re^2 + d^2$

Therefore, $d^2 = 2 Re h + h^2$

Hence, ' h ' is a small value compared to the radius of the earth, so h^2 can be removed.

$d^2 = 2 Re h$.

Therefore, $d = \sqrt{2hRe}$

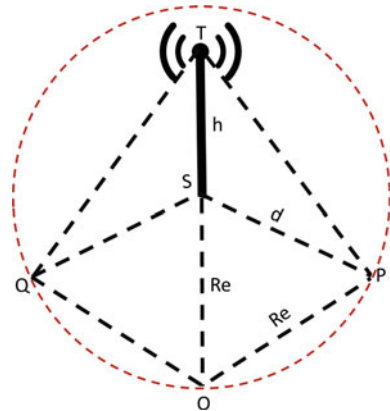
Coverage area is calculated as $A = \pi d^2$

Time Calculation

Time = Distance/Speed

$S = \{BS_1, BS_2, BS_3, \dots, BS_n\}$

Fig. 5 Coverage area calculation



$$\forall d, t \in \text{BS}$$

where BS is base station, $D1$ —maximum distance and $T1$ —minimum time

If the source vehicle is traveling toward the destination, it has to change the BS frequently. To avoid the number of HOs, the source vehicle has to calculate and prepare the neighbor BSs distance ($D1$) and reaching time ($T1$) to prepare the HO process. The local decision is made in the vehicle based on the stored value in the queue of the vehicle’s buffer. Based on the highest ($D1$) value and minimum ($T1$) value, the source vehicle has to choose maximum coverage BS for changing the HO using BS ($D1, T1$).

Coverage Algorithm

- Step 1: Set of available BSs will be based on the selected BSs from the link quality algorithm.
- Step 2: Calculate coverage area of BSs based on the ‘ $D1$ ’ and ‘ $T1$ ’.
- Step 3: $D1$ and $T1$ values are stored in the buffer for selecting BSs.
- Step 4: Sort the collected BS information according to $D1$ and $T1$ values from highest to lowest (BS ranking based on the current parameter).
- Step 5: Select the top 50% BSs and store the information in the UE’s buffer.
- Step 6: Update the BS information in the buffer after T-time by collecting new BS information (Step 2) to ensure the reliability of buffer information.
- Step 7: If the number of BSs in the buffer is 0, reselection is triggered, go to Step 2.

3.4 Bandwidth-Based Network Selection

The researcher recommends counting the number of vehicle connections utilized on PC5 and Uu interfaces to determine the number of user connections at a certain cell

or BS. This technique may provide assured bandwidth to each vehicle while avoiding congested network connections.

We make the following assumptions:

- Measure the number of users connected in the BS interface from time to time.
- To arrange BSs in the queue based on the available users and bandwidth.
- The best BS has to be selected from the above queue for HO.

Bandwidth calculation based on V2X interfaces is shown in Fig. 6.

To calculate the bandwidth need required, the following formula can be used:

$$(\text{Application throughput}) \times (\text{Number of concurrent users}) =$$

Aggregate application throughput

$BS = \{\text{Number of users, time}\}$.

$S = \{BS_1, BS_2, BS_3, \dots, m\}$.

$Bk(t)$ provided bandwidth of BS, $k \in BS$ at time t ,

Set of vehicles $U = \{1, 2, 3, \dots, n\}$.

$Bi(t)$ is the bandwidth needed for user, $i \in U$ at time t ,

Ai is a selected BS by the user i .

Bandwidth Algorithm

Step 1: Set of available BSs will be based on the selected BSs from link quality algorithm.

Step 2: Determine the number of users currently connected in each BSs.

Step 3: Select best BSs with minimum number of users.

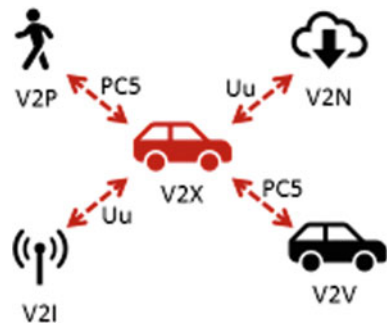
Step 4: Sort the collected BS information from lowest to highest users (BS ranking based on the current parameter).

Step 5: Select the top 50% BSs and store the information in the UE's buffer.

Step 6: Update the BS information in the buffer after T-time by collecting new BS information (Step 2) to ensure the reliability of buffer information.

Step 7: If the number of BSs in the buffer is 0, reselection is triggered; go to Step 2.

Fig. 6 Bandwidth calculation based on V2X interfaces



3.5 Hybrid Cell Selection Algorithm

Each vehicle has an option to select a BS based on the signal quality, coverage, and number of users strategies mentioned earlier. The elimination of BSs that do not pass the required specifications has already lessened the possibility of HO issues. An optimization technique is introduced based on the above three parameters to identify BS for HO.

Step 1: A vehicle must call the best result of signal quality, coverage, and bandwidth algorithms.

Step 2: Predict a vehicle's movement by searching through historical records stored in a table.

Condition 1: If a vehicle record is found, select the BSs that the vehicle will probably encounter in the trajectory. Proceed to Step 3.

Condition 2: If no historical record is found, proceed to Step 4.

Step 3: Any BS identified in Step 2, Condition 1 can be immediately used by the vehicle for HO, eliminating the need to HO to other BSs.

Condition 1: If there is a change in the selected value (signal quality, coverage, or bandwidth) in the currently used BS, other BSs can be evaluated for the next HO.

Condition 2: If the vehicle has existing connection history in the BSs, HO preparation and authentication can be ignored.

Condition 3: If no BSs are available in the current trajectory, proceed to Step 1.

Step 4: Compare $\text{signal} > S_t$, $\text{coverage} > C_t$ and $\text{bandwidth} > B_t$ for each BS.

Step 5: Apply an optimization algorithm.

Step 6: Arrange in buffer.

Step 7: Select optimized BS for HO then save this information in a history table.

Step 8: If $\text{BS} < = S_t$ or C_t or B_t .

Step 9: Go to Step 5 to select the next best option or else go to Step 1.

where

C_t Coverage threshold value,

S_t Signal threshold value and,

B_t Bandwidth threshold value.

4 Conclusion

Handover in HetNets has gained a lot of importance, especially with the emergence of 5G and 6G networks. This process presents a more challenging scenario in V2X due to the high mobility nature of vehicles. Although 5G has extended its coverage, by creating D2D communications, this can lead to a greater number of the HOs during the UE mobility that greatly affects network performance. This paper proposes a new approach to reduce issues affecting HO operations by the selection of best BSs using methods that rely on signal quality, coverage area calculations, and bandwidth

parameters. Furthermore, the author has proposed a new optimized algorithm using the above-mentioned parameters to select the best cell for optimal handover decision making. In future work, the proposed mechanism will be extensively evaluated using a simulation tool.

References

1. Gandotra P, Jha RK (2016) Device-to-device communication in cellular networks: a survey. *J Netw Comput Appl* 71:99–117
2. Fodor G, Dahlman E, Mildh G, Parkvall S, Reider N, Miklós G, Turányi Z (2012) Design aspects of network assisted device-to-device communications. *IEEE Commun Mag* 50:170–177
3. Naik G, Choudhury B, Park J-M (2019) IEEE 802.11bd & 5G NR V2X: evolution of radio access technologies for V2X communications. *IEEE Access* 7:70169–70184
4. Thanh, LTT, Moh S (2021) Comprehensive survey of radio resource allocation schemes for 5G V2X communications. *IEEE Access* 9:123117–123133
5. Garcia MHC et al (2021) A Tutorial on 5G NR V2X communications. *IEEE Commun Surv Tutor* 23:1972–2026
6. Lopez-Perez D, Guvenç I, Chu X (2012) Mobility management challenges in 3GPP heterogeneous networks. *IEEE Commun Mag* 50:70–78
7. Zhao D, Yan Z, Wang M, Zhang P, Song B (2021) Is 5G handover secure and private? *A Sur IEEE Internet Things J* 8:12855–12879
8. Huang L, Lu L, Hua W (2020) A survey on next-cell prediction in cellular networks: schemes and applications. *IEEE Access* 8:201468–201485
9. Noor-A-Rahim M et al (2022) 6G for vehicle-to-everything (V2X) communications: enabling technologies, challenges, and opportunities. *Proc IEEE Inst Electr Electron Eng* 110:712–734
10. Guidolin F et al (2016) Context-aware handover policies in HetNets. *IEEE Trans Wirel Commun* 15:1895–1906
11. Hua Q et al (2019) A novel base-station selection strategy for cellular vehicle-to-everything (C-V2X) communications. *Appl Sci (Basel)*, 9:556
12. 3GPP TS 24.333 version 12.5.0 Release 12, Universal mobile telecommunications system (UMTS); LTE; Proximity-services (ProSe) Management objects (MO), https://www.etsi.org/deliver/etsi_ts/124300_124399/124333/12.05.00_60/ts_124333v120500p.pdf
13. Nasraoui L, Ikki S (2021) Neighbor discovery for ProSe and V2X communications. *IEEE Internet Things J* 8:7241–7251
14. Sellami A, Nasraoui L, Najjar L (2021) Neighbor-assisted localization for massive MIMO 5G systems. In: 2021 18th International multi-conference on systems, signals & devices (SSD). IEEE
15. Analysis of vehicular safety messaging in cellular networks. <https://arxiv.org/abs/2004.04327>
16. Ding G, Yuan J, Bao J, Yu G (2020) LSTM-based active user number estimation and prediction for cellular systems. *IEEE Wirel Commun Lett* 9:1258–1262
17. Elsayed MM, Hosny KM, Fouda MM et al (2022) Vehicles communication handover in 5G: a survey. *ICT Express*
18. Embus DA, Castillo AJ, Vivas FY, Caicedo OM, Ordóñez A (2020) NetSel-RF: a model for network selection based on multi-criteria and supervised learning. *Appl Sci (Basel)* 10:4382
19. Liao Q, Penna F, Stanczak S, Ren Z, Fertl P (2013) Context-aware handover optimization for relay-aided vehicular terminals. In: 2013 IEEE 14th workshop on signal processing advances in wireless communications (SPAWC). IEEE
20. Xenakis D, Passas N, Merakos L, Verikoukis C (2014) Mobility management for femtocells in LTE-advanced: key aspects and survey of handover decision algorithms. *IEEE Commun Surv Tutor* 16:64–91

21. Wang C-H, Lee C-J, Wu X (2020) A coverage-based location approach and performance evaluation for the deployment of 5G base stations. *IEEE Access* 8:123320–123333
22. Alablani IA, Arafah MA (2021) An adaptive cell selection scheme for 5G heterogeneous ultra-dense networks. *IEEE Access* 9:64224–64240
23. Zhang L, Chen X, Ma Y (2021) A similarity-based method for base station selection in 5G networks. In: 2021 IEEE international conference on parallel and distributed processing with applications, big data and cloud computing, sustainable computing and communications, social computing and networking (ISPA/BDCLOUD/SocialCom/SustainCom). IEEE
24. Ndashimye E, Sarkar NI, Ray SK (2020) A network selection method for handover in vehicle-to-infrastructure communications in multi-tier networks. *Wirel Netw* 26(1):387–401
25. Turkmen A, Ansari S, Klaine PV, Zhang L, Imran MAC (2021) IMPRESS: Indoor mobility prediction framework for Pre-Emptive indoor-outdoor handover for mmWave networks. *IEEE Open J Commun Soc* 2:2714–2724
26. Jiang D, Huo L, Lv Z, Song H, Qin W (2018) A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking. *IEEE Trans Intell Transp Syst* 19:3305–3319
27. Si Q, Cheng Z, Lin Y, Huang L, Tang Y (2020) Network selection in heterogeneous vehicular network: a one-to-many matching approach. In: 2020 IEEE 91st vehicular technology conference (VTC2020-Spring). IEEE
28. Biswas S, Vijayakumar P (2021) AP selection in cell-free massive MIMO system using machine learning algorithm. In: 2021 sixth international conference on wireless communications, signal processing and networking (WiSPNET). IEEE
29. Alablani IA, Arafah MA (2021) Applying a dwell time-based 5G V2x cell selection strategy in the city of los angeles, California. *IEEE Access* 9:153909–153925
30. Duraipandian (2019) Enhanced network selection and handover schema for heterogeneous wireless networks. 01:160–171, Dec 2019
31. Suma V, Haoxiang W (2021) Optimal key handover management for enhancing security in mobile network. 2:181–187, Dec 2020
32. Acronyms RSRP, RSSI, RSRQ, SINR when measuring signal strength, <https://www.signalbooster.com/blogs/news/acronyms-rsrp-rssi-rsrq-sinr>

Utilization of Augmented Reality Technology in the Campus Environment



Ford Lumban Gaol, Mufti Ikhsan Kamil, Aria Muhammad Iswardhana, Steven Gusda Firnandes, Fazri Fahrezi, Tokuro Matsuo, and Fonny Hutagalung

Abstract Technological advancements are accelerating in the modern day; one such advancement is augmented reality (AR) technology. Naturally, these advancements have an effect on various sectors of life, one of which is education. The education industry will be the most impacted if campuses implement and deploy augmented reality technology. One of these is through the use of augmented reality technology as an application-based learning medium for students in order to motivate them to learn and make it easier for them to improve their academic performance on campus. The issue at hand is the absence of utilization of augmentation reality technology as a learning medium on campus, which, of course, does not capitalize on technical advancements to improve student interest in learning. This study employs a research and development model. The method is divided into four phases: (1) planning; (2) data collecting; (3) analysis, discussion, and development; and (4) documentation. The findings of this study reveal that students enthusiastically support the usage of augmented reality technology as a learning medium in order to boost accomplishment

F. L. Gaol (✉)

Binus Graduate Program—Doctor of Computer Science, Jakarta, Indonesia

e-mail: fgaol@binus.edu

M. I. Kamil · A. M. Iswardhana · S. G. Firnandes · F. Fahrezi

Alam Sutera, School of Information System, Bina Nusantara University, Jakarta, Indonesia

e-mail: mufti.kamil@binus.ac.id

A. M. Iswardhana

e-mail: aria.iswardhana@binus.ac.id

S. G. Firnandes

e-mail: steven.firnandes@binus.ac.id

F. Fahrezi

e-mail: fazri.fahrezi@binus.ac.id

T. Matsuo

Advanced Institute of Industrial Technology, Tokyo, Japan

e-mail: matsuo@aait.ac.jp

F. Hutagalung

University of Malaya, Kuala Lumpur, Malaysia

e-mail: fonny@um.edu.my

and introduce new experiences through its use. However, it is required to have a system capable of serving as a technological medium, so that the construction of a system capable of serving as a medium for augmented reality may be easily accessed by students, who hope that this system would enable them to better their academic performance.

Keywords Technology · Education · Augmented reality · Application · Environment

1 Introduction

Augmented reality technology is a technical advancement that transports objects from the virtual world to the actual world, allowing users to interact with them. [1] Augmented reality (AR) is a term that refers to a technology that blends the real and virtual worlds, making them interactive and in the form of three-dimensional animation. [2] Augmented reality (AR) is a visual representation of a physical object obtained directly or indirectly through the addition of data that can subsequently be shown. Virtually, virtual objects are used to present information that humans find offensive. [3] The utilization of VR technology in an educational setting constrains AR not in terms of technology itself, but in how this technology is used and how students learn.

Technology enables the establishment of a necessary learning environment conducive to the most effective implementation of the learning process

Kiryakova Gabriela [4] apart from being used in domains such as health, the military, and manufacturing, reality has also been incorporated into widely used gadgets, like as mobile phones [5]. In a college context, augmented reality technology is advantageous for giving real-time and unique information. Augmented reality is being developed in a variety of fields, including entertainment, medical training, design, robotics, manufacturing, and education. Specifically, e-learning is one of the domains of learning where augmented reality applications are most widely used [6]. In education, particularly on college campuses, augmented reality technology can broaden students' perspectives and encourage them to be creative and imaginative, making the learning process more efficient and less stodgy.

Suprihatiningrum [7] learning is a sequence of activities that include information and the environment in order to facilitate the process of learning for pupils. In essence, students use learning to grow their own potential [8]. Additionally, the utilization of augmented reality technology can be employed as campus media to assist prospective new students in promoting information about university facilities and lectures on campus. [9] Learning is a collaborative activity including learners, teachers, and instructional tools. With pupils experiencing cognitive stage development, augmented reality technology may be favored for learning, particularly when teaching abstract topics [10]. Sural [11] We require additional research and essential

learning materials generated using augmented reality technology and then integrated into the learning process.

By adopting augmented reality technology, information on brochures that are frequently used to deliver information to kids in their environment can be added in the form of virtual 3D information. As a result, by utilizing augmented reality, critical elements not included in the brochure can be completed. The advancement of augmented reality technology enables academics to create and assess augmented reality-based learning experiences [12]. The results indicate that pupils can examine brochures in greater detail using augmented reality technology.

The main goal with this technology is to attract the interest and attractiveness of students in using AR technology. This technology is used as a tool other than as a medium for learning, as well as a medium to promote the campus by utilizing augmented reality technology. In addition, this technology is also capable of being a communication medium, making it easier for marketing to be used and helping student learning. Students also can see the environment or information from inside the booklet will also feel more alive and real with the application of supporting animations such as architectural buildings art such as paintings, trees, and others. So that the teaching and learning process can do well, students should be invited to take advantage of all the senses [13].

1.1 AR for Criminal Purpose

Most likely, there will be no potential that aims to commit crimes using augmented reality technology, on the contrary, augmented reality technology can also be used by the police as a technology medium in uncovering crimes.

Although there will likely be no potential for crime by utilizing this technology, there is a small part that can potentially lead to criminal acts, namely by those who are very skilled in utilizing and managing this technology, and making a technology based on augmented reality that can perform hacking, phishing, and even digital sexual harassment.

1.2 AR for Entertainment Purpose

Augmented reality, or AR, provides a new level to entertainment media by allowing viewers to participate actively in the presentation, rather than as passive spectators who can just watch and see. In a show that merges the actual and virtual worlds, augmented reality adds interactivity and user involvement. Additionally, the implications inherent in this technology might serve as a catalyst for amusement in a variety of contexts. The implications of this research are that it may have a beneficial effect on the community, particularly students and campus parties, by demonstrating that the use of augmented reality technology in this era is critical and significantly

facilitates various activities such as learning, promotional media, and communication media, among others. Additionally, the findings of this study can contribute to the field of knowledge since they explain a fast evolving technology and educate people about how the usage of augmented reality technology can have a significant impact on the evolution of human life. This research indicates that research has a significant influence due to the usage of augmented reality technology, which benefits students by allowing for the printing of brochures contained in wall magazines that can be scanned using the AR application to generate a 3D model of the relevant brochure.

2 Research Technique

The technique we use as data collection is an online survey using a Google Form. This survey will be given to respondents randomly consisting of friends who are on campus, at home, and in the surrounding environment. Afterward, we conducted a survey targeting as many as 20 respondents. In the survey, there were several questions regarding the opinions of respondents regarding the use of AR technology in the campus environment in the field of education [14].

From the results of an online survey conducted by respondents, they argue that the use of AR technology in the campus environment is still underutilized, even though the technology is very useful if used properly by the campus if used properly [15].

The following as shown in Fig. 1 is a flowchart that describes the techniques in this research:

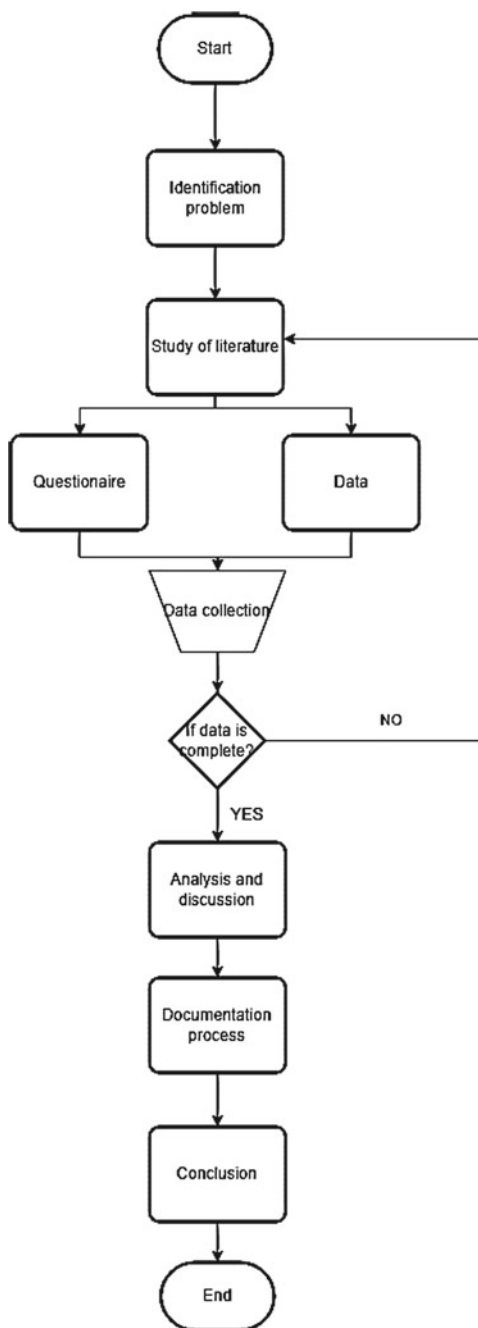
Additionally, we discovered that 100% of respondents agreed to deploy augmented reality technology on campus to make it easier for students to study. This is accomplished through the utilization of research and development methodologies to build augmented reality technologies on campus to aid in student learning. We are presenting this study report to learn about and utilize augmented reality technology on a college campus.

- How does the private campus promote learning through the use of augmented reality?
- What issues can arise as a result of implementing augmented reality?

3 Research Method

The purpose of this study is to determine the application and development of augmented reality (AR) technology in the creation of more attractive and efficient learning medium. This research employs both development and research methodologies. This strategy can result in the creation of a product based on study findings. Method research and development is defined as research procedures that result in the production of a product (which can be a model, module, or something else), and the product has an effect [10, 16].

Fig. 1 Research methods



This approach of research and development can be utilized by researchers to discover a model and determine its use. The concrete product is a learning media product that students can use to supplement their courses using augmented reality technology. Methods of research and development, abbreviated as R&D, have the same meaning as methods of development research. [7, 17] Educational research and development (R&D) is the process by which educational goods are developed and validated.

Educational development research is a method that instructional product producers employ to create educational products. It is vital to incorporate appropriate hypothesis testing strategies in this type of quantitative study [11, 18]. The research and development model is divided into four phases, which are as follows:

- A. **Planning Process:** Prior to performing research, preparations are done for the requirements. Preparations were made, which included a review of the literature on augmented reality, markerless augmented reality, and problem formulation.
- B. **Data Collection Method:** This process collects data via questionnaires distributed to campus users, ranging from students to professors. The data obtained includes information on the benefits of the augmented reality application and information that was not included in the application of augmented reality or in the data processing for analysis.
- C. **The analysis, discussion, and development process** involve conducting an analysis of the application, designing and designing, programming and coding, implementing and testing, and evaluating the application that will be used/implemented.
- D. **Documentation Process:** Following the completion of the three preceding stages, documentation is conducted to record the research findings.

A limited amount of product usability study was undertaken on campus to gauge student reactions and monitor student reactions to the usage of augmented reality technology as a learning medium. This is made feasible through the distinction between what is known and what should be known, which identifies what is missing and undiscovered in technique, theory, and the research literature in general. By identifying and analyzing these gaps, we can gain a better understanding of the study being conducted [19].

The steps in analyzing the data in this study are as follows:

4 Descriptive Analysis

- Looking for the average or the mean of variable X and variable Y, with the following formula:

$$M_y/y = y/N^{13}$$

$$\frac{M_x}{x} = \frac{X}{N}$$

$$\frac{My}{y} = \frac{y}{N^{13}}$$

- Looking for variance/deviation score (SD), with the formula:

$$S_x^2 = \frac{x^2}{d.k} \quad S_y^2 = \frac{y^2}{d.k}$$

$$S_x = S_x^2 \quad S_y = \sqrt{Y_x^2}^{14}$$

- Looking for the highest score of X and Y

5 Hypothesis Testing Analysis

This analysis is used to test the truth of the proposed hypothesis, while the way of analysis is through processing which will look for the relationship between X variable data and Y variable.

The steps are as follows:

$$r_{xy} = \frac{xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}^{15}$$

Description:

- r_{xy} Correlation coefficient between x and y .
- xy The product of x and y .
- x Student perception variable about the use of AR.
- y Student behavior variable.
- Σ Sigma (amount).

6 Analysis of Significance Test

This analysis is to make further interpretations by comparing the r value of the product moment correlation coefficient (r_{xy}) with the r table value (rt) with a significance level of 1% or 5% as follows:

1. If the r_{xy} value is greater than rt 1% or 5%, the results obtained are significant.
2. If the r_{xy} value is less than rt 1% or 5%, the result obtained is non-significant.

7 Research Result

The requirements needed to develop an augmented reality technology system are in the form of adequate hardware and software and are also able to create an augmented reality system properly. These needs include as follows:

Hardware requirements:

- (a) Processor: Intel-i5 10600 k
- (b) Graphics Card: NVIDIA Geforce GTX 1660
- (c) Memory: 16 GB

Hardware is very useful in developing a system that will run because it is able to manage inputted data and produce the expected output. In this case, the expected output is the application of augmented reality technology as a learning medium in the campus environment.

Software requirements:

- (a) Game Engine Unity IDE version 5.2.2
- (b) vuforia-unity-6-0-17.unitypackage
- (c) Microsoft Windows 10

Besides hardware, there is also software that plays a role in detecting hardware so that the system runs well in developing an augmented reality system as a learning medium in the campus environment.

From the results of research that has been carried out to find out how to use augmented reality technology in the campus environment in supporting learning and what problems will occur when using augmented reality technology. There are differences in terms of comparison of the results of this study with the previous studies. The previous research revealed that augmented reality technology still cannot be implemented properly on campus as a learning medium and many students object to changing the existing system into the latest system using augmented reality technology. However, in the results of this study, many students support the change from the old system to the newest system using augmented reality technology, and many campuses have used this technology as a medium of learning in the classroom. The following are the results of the research that has been carried out:

From the Table 1, these results show evidence that many students support the use of augmented reality technology in the campus environment as a learning medium. 12 strongly agree, 14 agree, 9 neutral, 4 disagree, and 2 strongly disagree.

From the Table 2 which contains questions regarding the use of augmented reality in the campus environment, many agree if it is applied in the campus environment as a learning medium for various reasons and also the respondents think about the impact that will be obtained if augmented reality is applied on campus.

We can learn from Table 3 that the problems that raised with the implementation of augmented reality relate with cost. This can be explained that due to the advance technology, the augmented reality will incur the high cost. Beside cost, there are issues that become problems such as system problem and errors, features that do not

Table 1 Assessment from students in the application of augmented reality technology

Respondent	Category
12	Strongly agree
14	Agree
9	Neutral
4	Disagree
2	Strongly disagree

Table 2 Augmented reality to support learning

Question	Answer
Do you think augmented reality can support learning? and please your reason?	Yes, because at this time, a lot of technology is developing, so we have to use it
	Of course, because augmented reality is trending at the moment, and it is used for learning
	Sure, because many other campuses have implemented it
	Yes, augmented reality is able to support learning because of its supportive features
	Its function is to make things real into things that are possible to support learning
What will be the impact if the implementation of augmented reality is carried out?	Improved learning achieved
	Learning just got easier
	Benefit if used properly
	Students can improvise their learning
	Learning to be fun

support learning, the server is full and bugs occur, and augmented reality has not been able to help to learn.

8 Discussion

According to the findings of the research and the table above, the majority of respondents are students who profess to favor the presence of augmented reality technology. They claim to be quite supportive of this technology as a means of enhancing their accomplishments and adding new experiences. This is in contrast to prior studies conducted by numerous academics addressing numerous students and lecturers who objected to the use of this technology as a medium for learning due to its high cost and limited accessibility. However, many schools are already using it as a mode of instruction for students, and students are pleased with the system’s existence.

Table 3 Problem to be faced

Question	Answer
In your opinion, what are the problems that will be faced with the implementation of augmented reality?	Expensive cost
	There are system problem and errors
	Features that do not support learning
	The server is full and bugs occur
	Augmented reality has not been able to help to learn
What are your hopes for augmented reality technology as a learning tool?	Augmented reality can improve student achievement
	Able to make learning fun
	Able to create a superior generation
	Learning just got easier
	Can make learning more effective and efficient

9 Conclusion

The research indicates that students accept the use of augmented reality technology as a medium for learning on campus in order to better their achievements and get new experiences. The best method to utilize augmented reality technology as a teaching tool is to build a system that assists students while they are on campus. The system can be constructed utilizing the hardware and software provided in the research results since the hardware and software are suitable for the development of augmented reality. The problems that may arise while employing augmented reality in a campus environment, as reported by respondents, include a system that is unable of assisting them in learning, system failures and bugs that impede their learning process, and features that are incapable of assisting their learning. It could become an issue in future. The respondents hope that this application will function properly and that it will enable them to better their academic performance if this technology is employed as a tool for learning.

Acknowledgements The research is funded by the Research and Technology Transfer Office of Bina Nusantara University.

References

1. Azuma RT (1997) A Survey of Augmented Reality. *Teleoperators and Virtual Environments*, 355–385
2. Furht B (2011) *Handbook of Augmented Reality*. Springer, New York

3. Gutiérrez-Jorge Martín CE-D-M (2017) Virtual technologies trends in education. *EURASIA J Math Sci Technol Educ* 469–486
4. Kiryakova Gabriela AN (2018) The potential of augmented reality to transform education into smart education. *TEM J* 556–565
5. Haller M, Billingham M (2010) *Emerging technologies of Augmented Reality, Interfaces and Design* Idea Group Publishing
6. Kose U et al (2013) An augmented reality based mobile software to support learning experiences in computer science courses. In: 2013 international conference on virtual and augmented reality in education. *Procedia Comput Sci* 25, 370–374 (2013)
7. Suprihatiningrum J (2013) *Strategi Pembelajaran Teori dan Aplikasi*. AR Ruzz Media, Yogyakarta
8. Rudi S, Cepi R (2008) *Media Pembelajaran*. CV Wacana Prima, Bandung
9. Sanaky HAH (2013) *Media Pembelajaran Interaktif-Inovatif*. Kaukaba Dirpantara, Yogyakarta
10. Sirakaya Mustafa EK (2018) The effect of augmented reality use on achievement, misconception and course engagement. *Contemp Educ Technol* 297–314
11. Sural I (2018) Augmented reality experience: Initial perceptions of higher education students. *Int J Instr* 565–576
12. Dutta K (2015) *Augmented reality for E-Learning*. Seminar augmented reality, mobile & wearable. *Augmented Reality, Mobile & Wearable*, Aachen
13. Arsyad A (2013) *Media pembelajaran edisi revisi*. PT. Rajagrafindo Persada, Jakarta
14. Alzahrani NM (2020) Augmented reality: A systematic review of its benefits and challenges in E-Learning contexts; Aug 2020; *Appl Sci* 10(16):5660. <https://doi.org/10.3390/app10165660>
15. Joshith VP (2020) Augmented reality in English language pedagogy: an innovative techno culture for contemporary classrooms-a meta review. *Int J Adv Sci Technol* 29(03)
16. Saputro B (2016) *Manajemen Penelitian Pengembangan*. Aswaja Pressindo, Yogyakarta
17. Borg WR (1983) *Education research and introduction*(fourth ed). Longman Inc., New York
18. Fitriyah IJ, Setiawan AM, Marsuki MF, Hamimi E (2020) Development of augmented reality assisted learning media to improve concept understanding of chemical bonding topics. *J Pembelajaran Sains* 4(2) (2020)
19. Gurevych R, Silveistr A, Mokliuk M, Shaposhnikova I, Gordiichuk G, Saiapina S (2021) Using augmented reality technology in higher education institutions. *Postmodern Openings* 12(2):109–132

FPGA-Based Implementation of Time-To-Digital Converter



H. Mangalam, S. Meishree, S. Mageshwari, and B. Prathiksha

Abstract Time-to-digital converters are utilized to represent time in digital format whenever the time pulse is given as an input. The vital requirements for mainstream applications include a broad measurement range, low cost, high resolution, voltage sensitivity, and temperature sensitivity. Previously, many TDCs have been designed using application specific integrated circuits (ASICs), in order to attain resolution more than 10 ps. But, implementation using ASIC does not provide reprogramming feature. Therefore, this paper proposes the new concept of time-to-digital converter (TDC) using FPGA with high resolution. In this concept, the given input signals are sampled n number of times. The timing reference is obtained by feeding the original clock signal to tapped delay lines. A single reference time period is obtained in accordance with the periodicity for achieving high resolution. Finally, this high resolution TDC is implemented in the VERILOG and synthesized using XILINX FPGA. The performance of TDC is also evaluated in terms of power, delay, and area, and a high resolution of 10 ps has been achieved.

Keywords TDC · CMOS · FPGA board · Pulse generator · Tapped delay line · VERILOG · XILINX

H. Mangalam · S. Meishree (✉) · S. Mageshwari · B. Prathiksha
Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering
College, Coimbatore, India
e-mail: meishree.1802116@srec.ac.in

H. Mangalam
e-mail: mangalam.h@srec.ac.in

S. Mageshwari
e-mail: mageshwari.1802112@srec.ac.in

B. Prathiksha
e-mail: prathiksha.1802142@srec.ac.in

1 Introduction

Temperature, mass, or time are the well-known physical properties, and they are processed through analog circuits by electrical quantities like voltage, current, charge, or frequency. Then, the VLSI technology has been evolved with technology scaling to achieve better performances such as size, speed, accurate, cost, and mainly resolution. Time domain signal processing can only be fully exploited if there is no analog conversion step in time-to-digital conversion and is being fully digital. Technological advances in CMOS are directed toward the optimization of digital circuits, and hence, the conversion of analog time value to digital value with higher resolution and accuracy is becoming important.

In applications, time intervals are measured using TDCs that can be implemented using field programmable gate arrays (FPGAs). In TDC, the obtained analog time is digitized to make it suitable for various applications. An FPGA integrated TDC with nanosecond resolution is of great performance as the technology allows precise control of the internal propagation time for a signal which is generated by a pulse generator. The original time adder has been replaced by this pulse generator, and linearity of pulse generator plays an important role, as the TDC resolution advances to a few picoseconds. In the CMOS inverter, longer channel length (L) has been adopted to optimize area and increase the dynamic range of delay line [1–4]. Finally, the time subtractor (TS) is used to reduce offset errors and obtain better accuracy. The goal of this work is to implement TDC in a Xilinx FPGA device to achieve better resolution and accuracy.

2 Related Works

- Chun-Chi et al. [1] proposed a CMOS-based converter that converts both time and temperature to corresponding digital values. For eliminating the time consumed by full-custom CMOS design, the circuit is fully designed with digital CMOS logic gates. The pulse-shrinking approach [3] with a subgate resolution is acquired with simple circuit and without any complicated process. All digital pulse mixing unit provides adequate time resolution as required. The fabrication of time-to-digital converter is done using 0.35 micron CMOS technology. The effective time resolution was approximately 45 ps.
- Chun-Chi et al. [5] proposed a highly accurate time-to-digital converter which serves as the major circuit in various specific instrumentation systems utilized for resolving the indefinite timing of signals. The cyclic delay line circuit has identical logic gates, thereby eliminating undesired shift resolution and the proposed pulse-shrinking unit provided a good resolution. The proposed TDC converter first converts the physical quantity into a time signal and then digitizes into a digital code. The implementation of CMOS TDCs at picosecond range and good resolution has widened various applications.

- Maatta et al. [6] designed a time-to-digital converter and constructed six units, and it has been tested. The stable performance of the time-to-digital converter is obtained by the construction of interpolation electronics. The developed real-time calibration procedure is used to improve the stability of the system. Synchronous digital time interval is used in this operation, in which an analog interpolation method improves the uncertainty of one clock cycle and has a better single-shot precision of few picoseconds, and range of measurement is comparatively less. The measured time interval is converted into the digital code during the implementation, and the obtained digital code is based on synchronous counting of the clock pulses of an oscillator. Here, the TDC is designed especially for pulsed time of flight laser radar applications.
- Dudek et al. [7] proposed a technique that allows to achieve more resolution with less dead-time. The resolution stabilization is achieved using a delay locked loop against process variations and other conditions. They have proposed various methods for digitizing short time intervals. For this process, fast counters have been used and CMOS tapped delay line, and a voltage ramp is generated based on analog method. An effective resolution is achieved by using a time unit as logic buffer delay. This delay locked loop scheme changes the buffer delay values that varies with temperature and power and stabilizes the delay value for the vernier line to have high effective range for providing calibration automatically. A test circuit is created using $0.7\ \mu\text{m}$ CMOS process technology. The time-to-digital converter has 128 delay stages, and 30 ps resolution has been obtained, stabilized by the delay locked loop, with the accuracy more than 1LSB.
- Malti Bansal et al. [8] discussed a taxonomical review of different designs of multiplexers and demultiplexers. They have also discussed on various techniques used in multiplexing that is employed in analog and digital electronics. This paper discussed the improvement and development that has taken place in multiplexing, and how the telephony multiplexing has taken a growth.
- Chun et al. [9] proposed a method for obtaining the frequency domain characteristics of the jitter using the vernier delay line using the various schemes proposed by different authors.
- Jansson et al. [10] proposed a high precision CMOS time-to-digital converter IC which measures the time interval using a counter and two level interpolator implemented using stabilized delay lines for improving integral nonlinearity of the interpolator.
- Mäntyniemi et al. [11] proposed a CMOS-based time-to-digital converter having a resolution of 1.2 ps. In this work also, a counter and two-stage stabilized delay line interpolation is used for time conversion. But, the total power consumption is very high, particularly dominated by delay locked loops.
- Hwan et al. [12] proposed a two-level conversion time-to-digital conversion scheme. In the first level, the multi-phase sampling technique is implemented with delay locked loop (DLL) for conversion, and in the second level, the sampling clocks and input signal are further adjusted and fed to vernier delay line (VDL) sampling circuit.

3 Methodology Existing System

Time-to-digital converters (TDCs) [1] perform the conversion of analog input time pulse into respective digital code. Several models for TDCs are done using simple microcontrollers and other methodologies like using ASIC, vernier delay line, etc. [3, 4, 9, 13, 14]. Many other frontend works have been done for the implementation of TDCs. These ideas have some limitations like less resolution, more power consumption, and the implementation using ASIC cannot be reprogrammed if needed. This motivated us to implement TDC using FPGA.

Proposed System

The paper proposes the new concept of time-to-digital converter (TDC) using FPGA with high resolution. The performance of TDC has been increased, and high resolution has been achieved. TDC implemented using FPGA has the reprogrammable capability. In this concept, the given input signals are sampled n number of times, and the original clock is fed to the tapped delay line from which a timing reference is generated. For those timing references, delays are achieved. Those delays are enfolded into a single reference period in accordance with the periodicity, and the effective TDC resolution can be achieved. The novelty of this work is that FPGA can be reprogrammed by a user, and besides it requires a smaller board space and can be more energy efficient. Finally, this concept of high resolution of TDC is implemented in the VERILOG and synthesized using XILINXFPGA.

A. Principle of Operation

In the LCD display, the digital format of the input time pulse will be displayed. The input pulse is generated using pulse generator and fed to the upcoming components like cyclic delay line for tapped delay, and the time subtractor is also used to reduce the offset error. The counter counts the input pulse, and the digital format of time will be displayed in the 7 segment LCD display. This implementation is done using Xilinx FPGA Spartan 6 board where all the required components are integrated in the board.

Thus, this project ensures the implementation of time-to-digital converter using FPGA board.

B. General Block Diagram

Figure 1 shows the basic block diagram of the time-to-digital converter implemented in our project. Here, the implementation is done using Xilinx FPGA board. The pulse generator generates the input pulse. The delay has been achieved from cyclic delay line by feeding the pulse obtained from pulse generator. The input pulse is fed to the time subtractor from where offset error has been eliminated, and the output from TS is given to counter which counts the time in accordance with the tapped delay line and the counts at the counter is displayed as digital format of time in LCD integrated in the FPGA board.

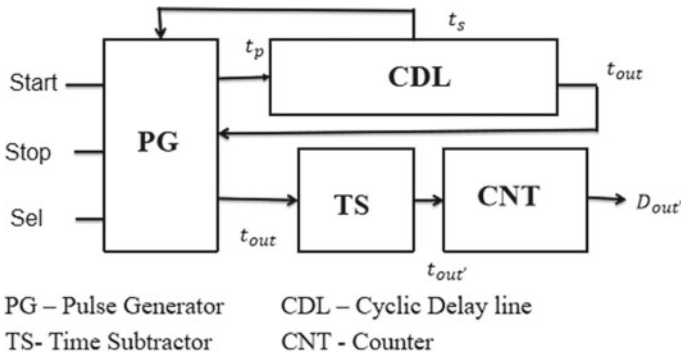


Fig. 1 Block diagram of the time-to-digital converter [1]

C. Implementation

The proposed TDC is implemented using FPGA board. All the hardware components are integrated in FPGA board and interfaced with the software. It employs pulse generator (PG), cyclic delay line (CDL), time subtractor (TS), and a counter (CNT). The various techniques like time amplification, time interpolation, and vernier principle have been employed to obtain resolution higher than the other methodologies. Software simulation has been done as a frontend process to verify the Verilog code for each submodule like pulse generator, cyclic delay line, time subtractor, and counter and also for the entire module. Then, this simulated Verilog code has been interfaced with the FPGA board for displaying the required digital format of time.

D. Development Procedure

Step 1: Verilog code has been written for each module like pulse generator, cyclic delay line, time subtractor, and counter has been written.

Step 2: The written Verilog code has been verified, RTL schematic of the entire module has been taken, and software simulation has been verified successfully.

Step 3: The verified code has been dumped into the FPGA Spartan 6 board for hardware implementation.

Step 4: On implementation, the required digital format of the time is displayed on the LCD integrated in the FPGA board.

So, by this process, the TDC has been implemented using Xilinx FPGA Spartan 6 board.

4 Results

The required digital format of time is displayed on the LCD, and the result is obtained with greater accuracy and a high resolution of 10 ps compared to TDCs developed with other methodologies. The performance of the entire system implemented using

Table 1 Area utilized by the proposed system

S. No	Components utilized	Used	% utilization
1	Number of slice registers (flip flops)	16 out of 11,440	1
2	Number of slice LUTs	12 out of 5720	1
3	Number of occupied slices	8 out of 1430	1
4	Number of MUXCYs used	8 out of 2860	1
5	Number of fully used LUT-FF pairs	13 out of 17	76
6	Logic utilization	11 out of 5720	1
7	Number of bonded IOs	12 out of 102	11
8	Number of BUFG/BUFGMUXs	1 out of 16	6

Spartan 6 FPGA board is evaluated by measuring the power, area, and delay. From the obtained delay results of the components, the maximum net delay of the system was computed to be 40.805 ns. The total power consumption of the system is 15 mW. Table 1 gives area utilized by the system in terms of the number of components utilized and the percentage of utilization.

5 Experimental Outputs

The software implementation for the system has been done using Xilinx tool, and the simulation has been verified using ISim. The hardware implementation has been done in Spartan 6 FPGA board as in Fig. 2. The bitwise measurement range of proposed TDC would be in hexadecimal, and till this hexadecimal range, the digital count will be displayed in LCD.

Fig. 2 Experimental hardware setup



The individual modules and the entire module are simulated using Xilinx tool. The RTL schematic of the entire module (TDC) is shown in Fig. 3. The simulation waveforms (ISim outputs) are shown in Figs. 4, 5, 6, 7 and 8. The software simulation has been successfully done, and the code has been verified.

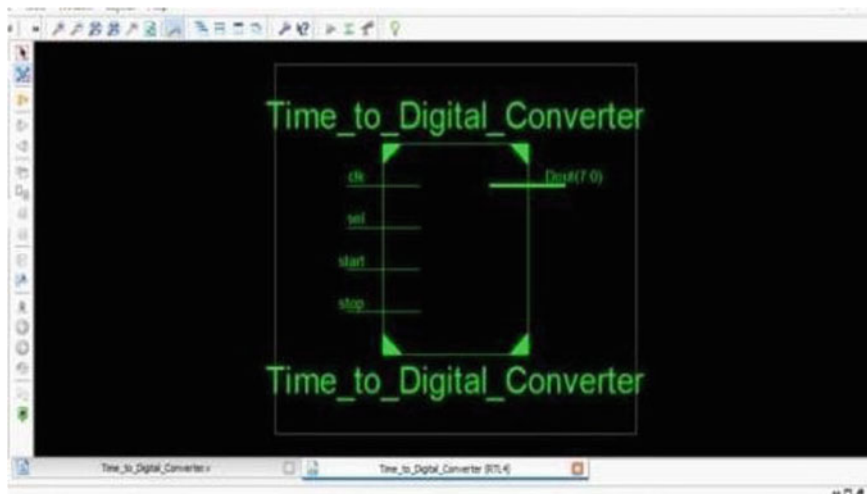


Fig. 3 RTL schematic of entire module

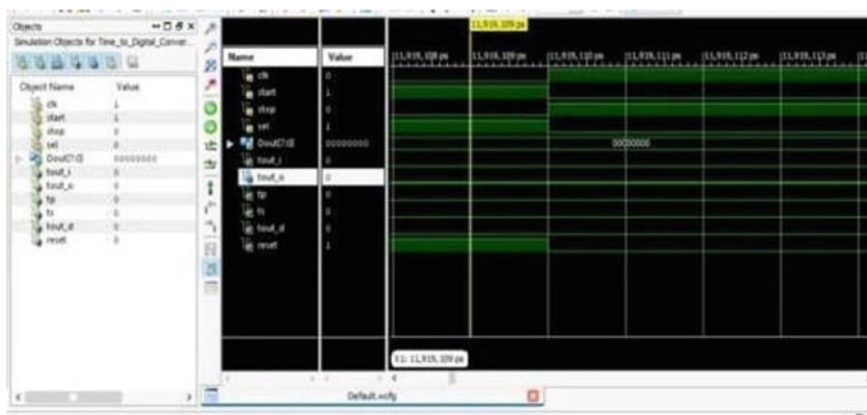


Fig. 4 ISim output of the entire TDC module

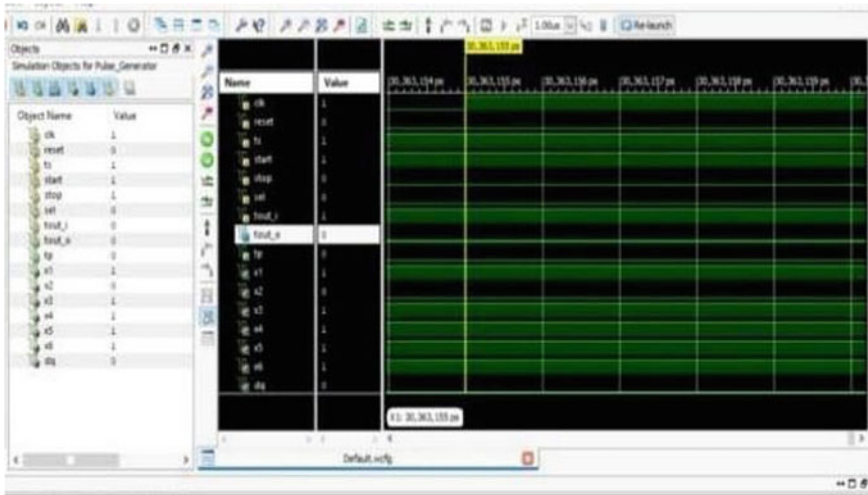


Fig. 5 ISim output of pulse generator

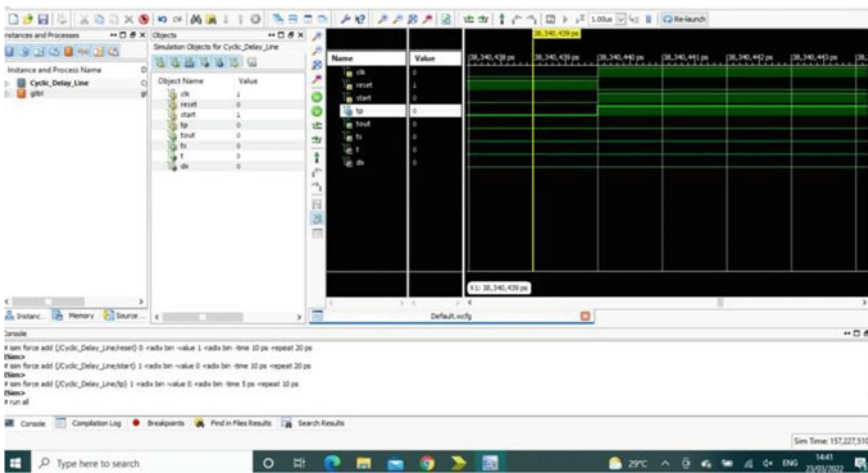


Fig. 6 ISim output of CDL

6 Conclusion

The FPGA implementation of the TDC has been done as a backend system like a smart sensor that senses the time pulse and converts it to digital code obtained from the counter. The implementation can be reprogrammed if needed as it is done using FPGA board. The frontend process of software simulation has also been done before



Fig. 7 ISim output of TS

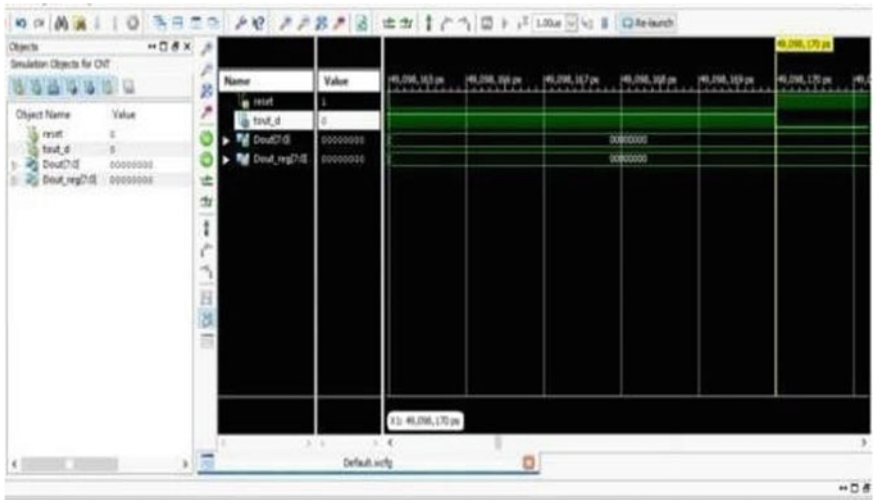


Fig. 8 ISim output of counter

hardware implementation for verification. Also, the approach presented in this work can be extended to any other application that requires TDC with reprogramming capability.

7 Future Scope

In future, it can be extended for any FPGA-related applications according to the requirements. This can be used to develop a time sensor where the digital format of time is required. Also this concept could be an example to implement time-to-digital converters using FPGA board.

Acknowledgements The authors would like to acknowledge the Management, Principal, and Head of the Department of Electronics and Communication Engineering of Sri Ramakrishna Engineering College, Coimbatore for the facilities provided for the successful completion of the project.

References

1. Chun-Chi C-LC, Fang W, Chu YC (2020) All-digital CMOS time-to-digital converter with temperature-measuring capability. *IEEE Trans VLSI Syst* 28(9), Sept 2020
2. Chen C-C, Hwang C-S, Lin Y, Chen G-H (2018) All digital pulse-shrinking time-to-digital converter with improved dynamic range. *Rev Sci Instrum* 87(4):046104-1–046104-3, Apr 2018
3. Chen P, Liu S-L, Wu J (2018) A CMOS pulse-shrinking delay element for time interval measurement. *IEEE Trans Circ Syst II, Analog Digital Signal Proces* 47(9):954–958
4. Cui K, Ren Z, Li X, Liu Z, Zhu R (2017) A high-linearity, ring oscillator-based, Vernier time-to-digital converter utilizing carry chains in FPGAs. *IEEE Trans Nucl Sci* 64(1):697–704
5. Chun-Chi C, Lin S-H, Hwang C-S (2020) An Area-efficient CMOS time-to-digital converter based on a pulse-shrinking scheme. *IEEE TCS* 61(3), Mar 2020
6. Maatta K, Kostamovaara J (2019) A high-precision time-to-digital converter for pulsed time-of-flight laser radar applications. *IEEETIM* 47(2), Apr 2019
7. Dudek P et al (2019) A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line. *IEEE JSSC*, 35(2), Feb 2019
8. Bansal M, Singh H, Sharma G (2021) A taxonomical review of multiplexer designs for electronic circuits and devices. *J Electron Inform* 03(02):77–88
9. Chan H, Roberts GW (2018) A jitter characterization system using a component-invariant Vernier delay line. *IEEE TVLSI* 12(1), Jan 2018
10. Jansson J-P, Mäntyniemi A, Kostamovaara J (2018) A CMOS time-to-digital converter with better than 10 ps single-shot precision. *IEEE J Solid-State Circ* 41(6):1286–1296
11. Mäntyniemi TE, Rahkonen, Kostamovaara J (2019) A CMOS time-to-digital converter (TDC) based on a cyclic time domain successive approximation interpolation method. *IEEE J Solid-State Circuits* 44(11):3067–3078
12. Hwang C-S, Chen P, Tsao H-W (2018) A high-precision time-to-digital converter using a two-level conversion scheme. *IEEE Trans Nucl Sci* 51(8):1349–1352
13. Brockhaus H, Glasmachers A (2018) Single-particle detector system for high resolution time measurements. *IEEE TNS* 39(4), Aug 2018
14. Ealgoo K et al (2018) Time of flight (TOF) measurement of adjacent pulses. *IEEE NSS* 36(3), Nov 2018

Implementation of Smart Poultry Farm Using Internet of Things



Bharati Ramteke and Snehlata Dongre

Abstract The world is growing expeditiously with the latest technologies taking charge of the market these days. Every single person is now probing for an automated machine that can be utilized for all the indispensable official, household, or industrial work so that they do not have to probe for laborers or workers for long. And technology has upgraded with time and provides the best solutions. IoT also known as Internet of things is a widely used technology to automate things nowadays. This project proposes an automated poultry farm predicated on IoT, and the whole farm will be covered with multiple sensors such as water level, sultriness sensor, temperature sensor, day–night sensor, along with a microcontroller to guide the system. The main aim of this concept is to automate the working of poultry farms, check the water level of the water bowl, and automatically turn on turn off lights, supply of pabulum, etc.

Keywords Automatic system · Poultry farm · Sensors · Cleaning

1 Introduction

Shelters for birds like chickens, ducks, turkey, and geese are known as poultry farms. The main motto behind the development of poultry farms is to take care of these birds so that products like eggs, meat, and feathers can be produced on a large scale and sold in the market. Poultry farming is widely done in India as an occupation, and also it has a huge profit in a one-time investment. But there is a lot of work inside the poultry farms such as cleaning of shades, supply of water and food to the birds, collection of eggs, maintaining the temperature as per the climate, and checking humidity and oxygen level from time to time which requires manpower to complete the work.

B. Ramteke (✉) · S. Dongre

Department of Computer Science and Engineering, G H Raisoni College of Engineering, Nagpur, India

e-mail: bharati.ramteke.cs@ghrce.raisoni.net

S. Dongre

e-mail: snehalata.dongre@raisoni.net

IoT is an essential platform where embedded sensors are connected to obtain the expected result. Also, we can attach the hardware with the software, an Android, or a Web application to control the functions of the system from anywhere. IoT is essentially a platform where embedded contrivances are connected to the cyber world, so they can accumulate and exchange data from an Android app. Thus, automatic work of machines is increasingly paramount in modern agriculture, abbreviating dependence on constant labor work, incrementing management scale and efficiency, consummating the precision and consistency of product control, enabling enforceable traceability enhancing aliment safety—all of which may achieve agricultural sustainability. Our proposed work is completely focused on the parameters like maintaining temperature and humidity inside the poultry farms, making the environment clean and safe for birds living there, providing them food and water on time and alerting the owner in any danger situation [1–7]. The system defined in the abstract is predicated on the Internet of things (IoT) and will perform all the indispensable tasks of a poultry farm and avail in incrementing the profit and outcomes.

2 Literature Review

Shoba et al. [8] presented an IoT-based system for viewing and controlling poultry farms, as well as monitoring the poultry farms without any physical presence. The system includes wireless sensors to keep the room temperature stable, analyze the weight of the chickens, and provide water and food.

Elham et al. [9] presented a system in which an application for the Raspberry Pi and a blockchain module is introduced. The blockchain is utilized to secure the security of customer data from chicken farms.

Soh et al. [10] proposed a proposal for automatic chicken feeding using an Arduino UNO as the main module. The other module is used to check the temperature and humidity, which is likewise done using the Arduino UNO.

Mumbelli et al. [11] proposed a novel model for monitoring and controlling all project parameters using a combination of hardware and software. All of these operations are carried out via remote access via a mobile app. Temperature value and actuators are the parameters used in this research study.

Dbouk et al. [12] provided a brief discussion on temperature control systems used in chicken farms. Their goal is to optimize the temperature of the poultry shade based on the needs. It has also been identified that maintaining an optimal temperature is a difficult process. To address this issues, the authors presented an IoT-based temperature control system based on the DHT11 sensor, which will save chicken farm owners time and money.

3 Proposed Methodology

This project includes IR sensor, temperature sensor, water-level sensor, day–night sensor, MQ-2 sensor, servo motor, DC motor, exhaust fan, motor pump, OLED display, solar plate, and the most important component of the project, i.e., ESP8266 microcontroller as in the above Fig. 1.

Module 1—In this module, the main working of the microcontroller will take place, which is the brain of the project. The microcontroller will send a signal to the water-level sensor to check the level of water in the bowl; if the level is low, the motor pump will start automatically till the maximum level of water is supplied.

Module 2—In this module, the working of food supply will take place with the help of conveyor belt and DC motor, the food will be supplied from time to time, and then, the area will be cleaned automatically soon. The conveyor belt will move slowly and stop for some time so that the chicken can eat their food, and then, the DC motor will stop rolling.

Module 3—This module focuses on automatic light turn on and off with respect to the surrounding light, and this operation will take place with the help of a day–night sensor. The working of the temperature sensor will take place here in this module, temperature inside the farm will be maintained, and if not, the cooling fan will be turned on.

Module 4—In this module, IR sensor is used to detect any suspicious entry or presence around the poultry farm to alert the owner about theft and security. A buzzer is used as an alarm here.

Module 5—In this module, MQ-2 sensor will work when the atmosphere inside the farm gets too gassy or full of harmful gasses. After detecting high risk

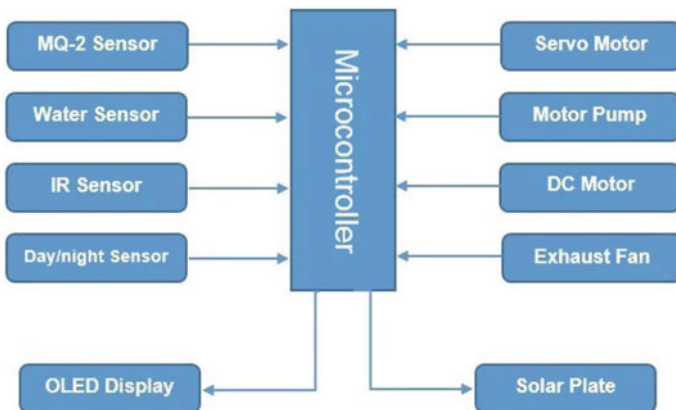


Fig. 1 System architecture of smart poultry farm using IoT

conditions, the exhaust fan will be turned on to refresh the room area and air quality.

Module 6—This module has an OLED display as its main component, all the alerts and information of the poultry farm like water level, air quality, temperature, etc., will be shown on the output screen.

Module 7—The final module of this project is solar plates which are used as a backup component; when there is no power supply inside a poultry farm, the system will get supply from solar energy.

4 System Working

The components related to the proposed work are shown from Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11.

The working of the system will be as follows:

Step1: System turn on and microcontroller sends signal to the other sensors for working.

Step 2: Water sensor is integrated with a water pump to check for water level and supply water when needed using water-level detection.

Fig. 2 IR sensor

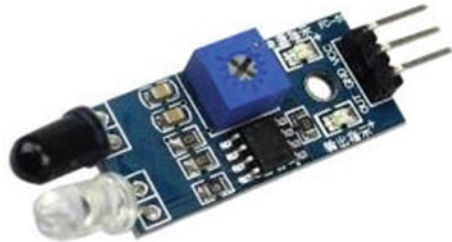


Fig. 3 DC motor



Fig. 4 Exhaust fan



Fig. 5 Temperature sensor

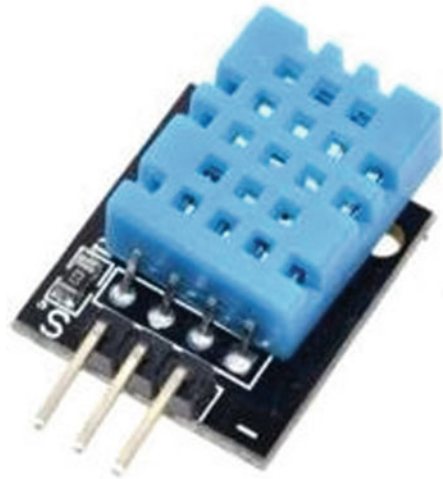


Fig. 6 Esp8266 node



Fig. 7 Water-level sensor

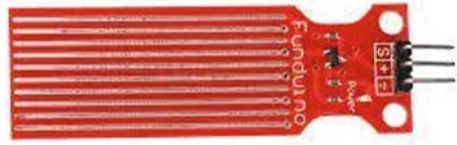


Fig. 8 LDR sensor

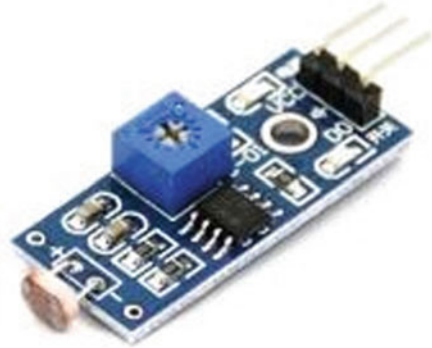


Fig. 9 Servo motor



Step 3: Conveyor belt with DC motor attached to it is doing the work of food supply from time to time. The DC motor will oven with a conveyor belt attached to it, and the food bowls will also move with it, the wait time for eating food will be set after that the area will be cleaned automatically.

Step 4: The LDR sensor will work when it gets too dark and sunny automatically, so that the farm lights will be turned on and off accordingly.

Step 5: For detection of any suspicious movement and presence of any person around the poultry farm, the system will generate an alarm to alert the owner.

Step 6: To display the complete information about the room temperature, humidity check, air quality, etc., will be shown on the OLED screen so the user gets updated information all the time.

Fig. 10 OLED display



Fig. 11 Water pump



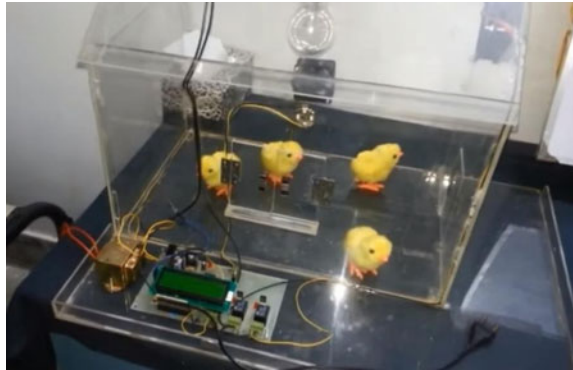
5 Result and Discussions

Result image of smart poultry farm is shown in Fig. 12.

6 Conclusion

The system is able to monitor the poultry farm and control the operations of the poultry farm. The proposed system is an amalgamation of wireless sensors to manage and monitor the poultry's work more easily. The system supersedes the human labor to aliment pabulum and dihydrogen monoxide into the containers. It surmounts the labor quandaries in the poultry industry and it with avails to control temperature, sultriness, air quality of the poultry farm.

Fig. 12 Result image of smart poultry farm



References

1. Bartere MM, Lokulwar PP (2022) Survey on edge, fog assisted IoT framework using intelligent learning techniques. In: Proceedings of the 2nd international conference on recent trends in machine learning, IoT, smart cities and applications. Lecture Notes in Networks and Systems, vol 237. Springer, Singapore. https://doi.org/10.1007/978-981-16-6407-6_17
2. Lokulwar PP, Deshmukh HR (2017) Threat analysis and attacks modelling in routing towards IoT. In: 2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), pp 721–726. <https://doi.org/10.1109/ISMAL.2017.8058273>
3. Janzal SV, Talmale GR (2016) Medicine reminder and monitoring system for secure health using IOT 3 Apr 2016, Version of Record 3 Apr 2016. <https://doi.org/10.1016/j.procs.2016.02.090>
4. Talmale G, Shrawankar U (2017) Dynamic clustered hierarchical real time task assignment and resource management for IoT based smart human organ transplantation system. Conf Emerg Devices Smart Syst (ICEDSS) 2017:103–109. <https://doi.org/10.1109/ICEDSS.2017.8073667>
5. Fulare RP, Sakhare AV (2014) Efficient sensor node authentication in wireless integrated sensor networks using virtual certificate authority. Fourth Int Conf Commun Syst Netw Technol 2014:724–728. <https://doi.org/10.1109/CSNT.2014.152>
6. Nirale P, Madankar M (2021) Analytical study on IoT and machine learning based grading and sorting system for fruits. Int Conf Comput Intell Comput Appl (ICCICA) 2021:1–6. <https://doi.org/10.1109/ICCICA52458.2021.9697161>
7. Deotale P, Lokulwar P et al (2021) Smart crop protection system from wild animals using IoT. In: International conference on computational intelligence and computing applications (ICCICA), pp 1–4. <https://doi.org/10.1109/ICCICA52458.2021.9697315>
8. Shoba K, Sushmitha V, Umashree N, Yashaswini DM, Priyanka K (2020) An effective automated monitoring and controlling of poultry farm using IoT Int Res J Eng Technol (IRJET), Apr 2020
9. Elham MN et al (2020) A preliminary study on poultry farm environmental monitoring using internet of things and blockchain technology. IEEE, June 2020
10. Soh ZHC et al (2017) Development of automatic chicken feeder using arduino uno. In: International conference on electrical electronics and system engineering (ICEESE). IEEE.
11. Mumbelli A et al (2020) Low-cost IoT based system for monitoring and remote controlling aviaries. ICICT, 2020
12. Dbouk HM, Mourad R (2020) Solar heated poultry house. In: IEEE, August 2020. International conference on energy efficient technology for sustainability (ICEETS)

Design and Implementation of a Solar Powered Floating Device for Water Quality Monitoring in Inland Aquaculture Farms Using LoRa Wireless Communication



G. Mageshkumar , Saggurthi Prabhakara Rao, S. Suthagar ,
K. Hemalatha, K. Gowtham, and T. Hariharan

Abstract Inland fish farming is one of the growing food industries in India. The potential of this sector in creating an impact on the growth of the Indian economy is not yet realized. The Government of India is constantly focusing on various opportunities to build a self-reliant nation. Fish is one of the major foods consumed by the people of Southern India. Improper maintenance of aquafarms leads to health problems for the people who consume fish from these farms. In this paper, the types of fish farming, the challenges, threats, and recent remedial actions taken to overcome the challenges are discussed. A low-cost water quality monitoring system (WQMS) based on long-range wireless communication (LoRaWC) powered by a solar panel that floats on the water is proposed to improve the yield and quality of fish from the aquafarms. The sensor node is designed to measure pH level, turbidity, total dissolved solids (TDS), atmospheric temperature, humidity, and water level. The measured values are transmitted to the cloud through a LoRa gateway for taking necessary decisions to maintain the water quality in good condition. This method improves the growth and health of the fish in the aquafarm which in turn provides healthy food for mankind.

Keywords Aquaculture · Fish farm · Internet of things (IoT) · LoRaWAN · pH sensor · Solar powered LoRa node · Total dissolved solids (TDS) · Water quality monitoring system

G. Mageshkumar (✉) · S. Suthagar · K. Hemalatha · K. Gowtham · T. Hariharan
Kongu Engineering College, Perundurai, Erode 638060, India
e-mail: magesh.me.est@gmail.com

S. P. Rao
Joginpally B R Engineering College, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes
in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_49

687

1 Introduction

1.1 Definition

The process of breeding and artificially rearing various varieties of fish for commercial food purposes is known as inland fish farming.

1.2 Fish Farming in India

For 950 million people around the world, fish and fisheries products are their primary source of protein, and they are an important part of the diet for many more [2]. India is the world's second-largest fish manufacturer [3]. The fishing sector contributes 1% to the GDP, of which 65% is contributed by inland fish farming. Inland farming from less than 0.01% of the total volume of water on the planet, fisheries, and aquaculture contributes more than 40% of global fish production [7]. In India, inland aquaculture has long been the main fish production method. Ornamental fish culture through inland fish farming can also be creating additional income for the nation [18]. Figure 1 shows the growth of fish production value in billion Indian rupees (INR) for the financial years from 2012 through 2018 [9]. As one can observe that the demand is constantly growing it will be challenging to maintain the quality of production [10].

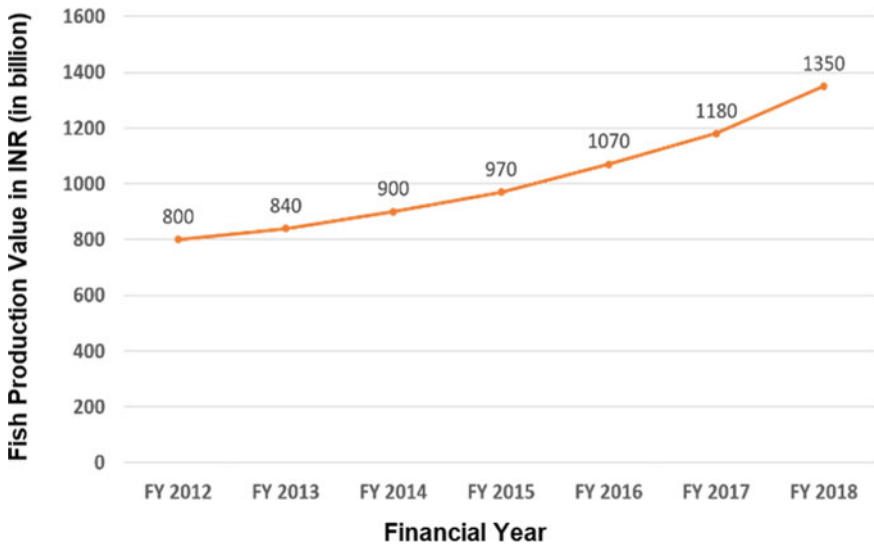


Fig. 1 Fish production value in billion Indian rupees

Table 1 Percentage contribution of fish production by inland and marine fish farming

Year	Inland	Marine
1950–51	29	71
1960–61	24	76
1970–71	38	62
1980–81	36	64
1990–91	40	60
2000–01	50	50
2010–11	61	39
2017–18	71	29

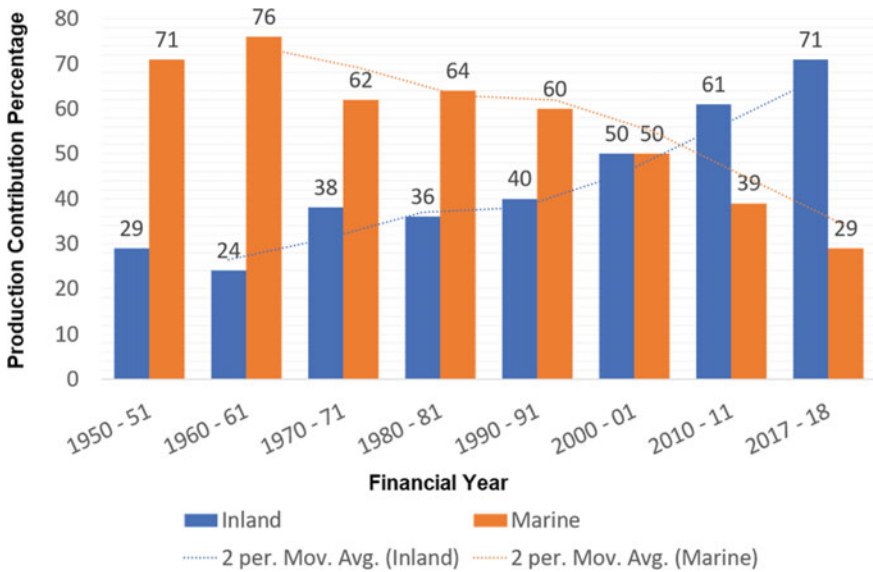


Fig. 2 Percentage contribution of inland and marine fisheries over the different period in India

Table 1 is taken from [1] and is represented as a graph in Fig. 2. The moving average of the percentage contribution of inland and marine fisheries is plotted in Fig. 2. From the plot, we can observe the trend that the inland fisheries are growing every decade, and we can predict that it will still grow for the upcoming years, but the trend in marine fisheries is falling down [9].

The most important fish species produced in pisciculture in India are Asian seabass, Pearl spot, Tilapia, and Catfish. The average fish consumption of an Indian is 5 kg per year. Indian Council of Medical Research (ICMR) recommends 12 kg of fish per year due to its high mineral value [12]. Demand for fish and fish oil is increasing day to day due to the increasing population. Those demands can be met out only by encouraging artificial fish farming.

1.3 Motivation

Fishes grew artificially in pisciculture, which satisfies the protein need of many people and fish production is also a commercial activity for the fish farmers. Healthy fish growth is essential to attain both. Proper water quality maintenance is a major thing to be ensured for this achievement [1]. Recent studies show that water pollution is the major threat to fish production. Both Federation of Indian Animal Protection Organizations (FIAPO) and All Creatures Great and Small (ACGS) combinedly surveyed 250 fish farms across the ten highest fish-producing states in India [6]. Survey result says that the consumption of unhygienic water fishes can lead several health issues in the human body. Excess lead and cadmium impurities are present in water due to antibiotics and insecticides used for cultivation. None of the fish farms had outlets for the recirculation of impure water. Due to dirt, 65% of fish farms had a low dissolved oxygen level. Disease outbreaks to fishes due to these effects lead to heavy commercial loss to farmers. In several instances, farmers selling these diseased fish to minimize their losses. The situation is worrying. Unless immediate action is taken either a commercial loss to the farmer or a life loss to a consumer may happen. Heavy metals can reach the body in small quantities through food, water, and air [6]. The excess quantity of toxic heavy metals can even cause death in living organisms by metabolic process interference. Also, there are other side effects such as fitness reduction, reproduction interference, etc. due to presence of heavy metals in water. Since heavy metals are not removed from water environments, they are contaminants for fish [2]. The source for heavy metals in aquatic farm are feeds, insecticides, and pesticides. Also, natural methods such as organic pollutants, mixing of sewage, and garbage leaching in freshwater can cause increase in heavy metal. The measure of lead (Pb), copper (Cu), and zinc in fish farms shows its excessiveness, which causes severe health issues in human health by entering into the food chain by human consumption of fish. The lead level in water depends on temperature, and the rate at which fish accumulated lead was affected by the salinity and temperature of the water.

1.4 Challenges in Fish Farm

There are several present problems in inland fishing and, and this section discusses them in brief.

- Demand is high, but the production is less.
- Seasonal variation affects the production, especially during summer water level is to be maintained for proper growth of fishes.
- Temperature increase may pull down the water level maintained in the fish farms.
- Initial investment is high to create a fish farm.
- Lack of required minimal infrastructure.
- Human mentality of polluting an open water body.
- Outbreak of diseases in fishes.

- Seed import without sufficient precautions.

There were several strategies suggested to overcome these challenges [7], diversified production through integration with the agricultural field, focused attention on producing seeds especially without disease outbreak, formulation against fish disease, infrastructure development for both production and post-harvest activities, establishing storage facilities for wastage reduction, initiating the practice of utilizing effective and affordable technology in the fish production process. Improved data collection will help the beginners to plan accordingly and decreased the chance of initial wastage and losses [21]. Management system development for addressing emerging issues, maintaining purity in water bodies, and conserving biodiversity. Institutional supports to be developed to provide genetic and biotechnological assistance in fish production and also helps the farmers in aspects such as site selection, portfolio diversification, expansion, intensification, and better integration [17]. Moreover, automation-related research is being carried out to improve the production of aquafarms.

1.5 Types of Fish Farming

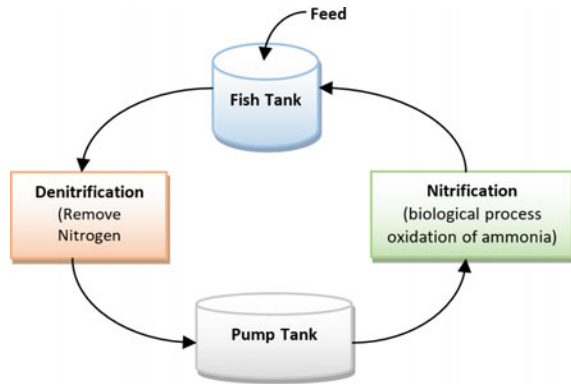
This section describes the recent trends that are followed by fish farmers all over the world. Currently, there are four types of pisciculture in trend as shown in Fig. 3.

An intensive method of fish farming (IMFF) is a well-managed type of fish farming, where all efforts are made to achieve maximum fish production from a minimum amount of water. Precision livestock fish farming (PLFF) is the use of sophisticated technology to maximize each animal’s contribution. It is a tool for livestock man-

Fig. 3 Trends in pisciculture



Fig. 4 Architecture of recirculating aquaculture system



agement that uses continuous automatic real-time animal tracking to enhance production, reproduction, health, and safety, as well as the effects on the ecosystem. It is a modern paradigm for aquaculture improvement. Open-ocean aquaculture can be optimally positioned and controlled more efficiently with new technologies. Waste is controlled by deep seas and faster winds, which hold it away from the vulnerable nearshore habitats [5]. Legal rules on this activity, though, remain unclear, and since open-ocean seas are rugged, technical challenges are complicated. Off-shore aquaculture needs to be included as one of the major activities for which spatial planning and conservation measures need to be taken care of in order to develop fish farming [16]. In a recirculating aquaculture system, water from the culture is constantly purified and reused. A recirculating aquaculture system is a closed-loop to control the nitrogen compound content by oxidation of ammonia [25] as shown in Fig. 4.

2 Technology in Current Scenario

Saha et al. [19] propose an outline for monitoring of water quality for aquaculture which is used using Arduino, Raspberry Pi and various sensors, an android application, and a smartphone camera. The parameters for water quality used in this paper are pH, color, temperature, and electrical conductivity. The use of an IoT device to measure and regulate water parameters has been suggested by the author [3]. Ammonia, water level, pH value, foul odour, temperature, and dissolved oxygen in the water are both detected and monitored by the system. Scalable approach for water quality monitoring using the LoRa module with the help of LoRaWAN protocol uses low-power wide area network technology (LPWAN) [15]. The system designed by the authors of [20] has a wireless LoRa module for sending and receiving sensor values, adding sensors to the microcontroller, and a ThingSpeak IoT platform for testing and visualizing water quality sensor values [22]. The device is made up of electrical, mechanical, and networking elements that are all connected to the Internet of things

(IoT) [3]. The electrical portion consists of a pi sensor, which offers real-time fish video data collection and Raspberry Pi B+ module interfaces. In paper by Deng, a model for the increasing status of aquaculture is created using an artificial neural network [8]. The growing state model is being developed in this work, and the use of artificial neural network technology can solve congestion problems in an expert system affect several water quality parameters [4, 11, 14]. This method is proven effective from the experimental results obtained. TNAU published research papers in fisheries which contain guidelines to be followed for rearing, breeding, and maintenance of popular commercial fishes [13]. The data provided in the website of TNAU are taken as reference for fixing the threshold values for the proposed water quality monitoring system.

3 Methodology

3.1 Parameters to Be Monitored

Water that uses for fish farming is considered suitable only based on certain parameters. These parameters include atmospheric parameters as well as water parameters that decide the growth of fishes. The list of parameters that are monitored by the proposed work is as follows,

- Temperature and humidity
- pH value
- Water level and quantity in liters
- Turbidity
- Total dissolved solids (TDS)
- Motor status.

3.2 Design of Sensor Node

The monitoring system contains five sensors for the measurement of water quality, quantity, and atmospheric condition parameters. DHT11 sensor is used for the measurement of temperature in degree Celsius and humidity in percentage [23], HC-SR04 is an ultrasonic sensor used for the measurement of water level in centimeters similar to oil level as mentioned in [24], and water quantity in liters can be calculated from water level, turbidity sensor for the measurement of turbidity in Nephelometric Turbidity Unit (NTU), TDS sensor for the measurement of total dissolved solids in parts per million, and pH sensor for the measurement of pH value. The measured parameter values are sent to the LoRa gateway which in turn forwards the data to the cloud using message queuing telemetry transport (MQTT) protocol. The parameters



Fig. 5 Inland fish farming tank and manual measurement of TDS and pH

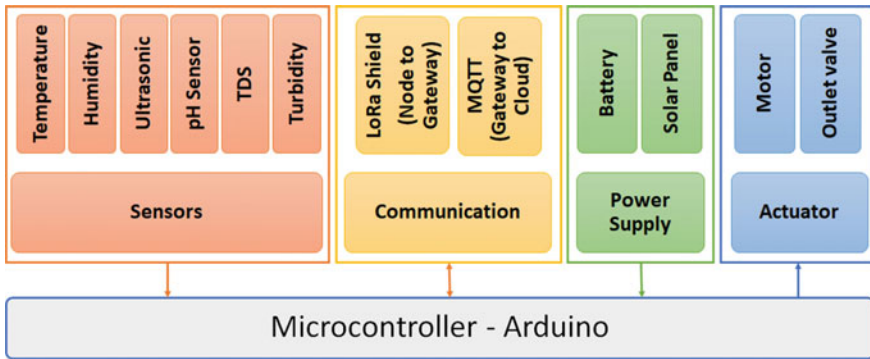


Fig. 6 Proposed architecture of the sensor node

are to be maintained in different levels for different species of fishes are available at the Tamil Nadu Agricultural University (TNAU) portal [13]. The proposed system is designed specifically to address the species that are majorly produced in southern part of India by inland fish farming. The sensor node is powered by a solar panel. LoRa uses the frequency band of 865–867 MHz in India. Arduino board with AtMega2560 microcontroller is used to interface the sensors and send the information to the LoRa gateway. The architecture of the proposed system is shown in Fig. 6. The Arduino is used as it the easy to program, and libraries are readily available for the Dragino LoRa shield used to implement the proposed idea.

3.3 Solar Panel Selection

To ensure the usage of renewable energy sources, a solar power panel is used to power up the LoRa sensor node. The total power requirement of the LoRa node is calculated, and it is found that 12 V (5 W) solar panel is sufficient enough to power the sensor node. Figure 7 shows the solar panels used for the power supply.

The total power consumption of all the components in the LoRa node is calculated with the help of a datasheet of corresponding components, and it is shown in Table 2. The design helps us to reduce the frequency of recharging and replacement of battery, thereby it reduces the power consumption cost by 1/10 times.

Fig. 7 Power supply unit

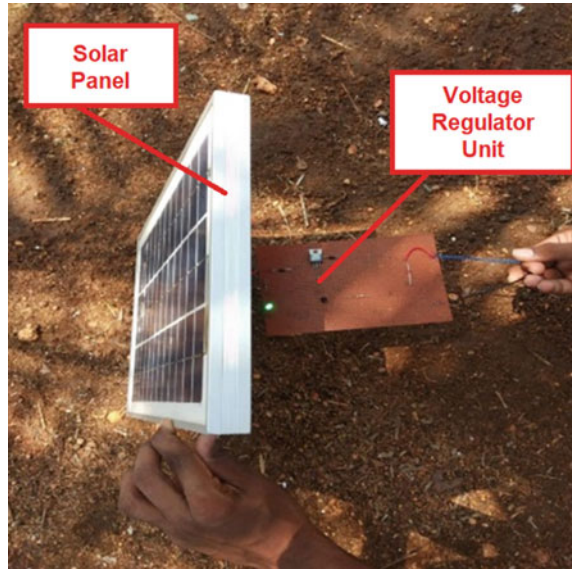


Table 2 Power requirement by the LoRa sensor nodes

S. No.	Components	Quantity	Voltage (V)	Current (mA)	Power (mW)
1.	DHT11 sensor	1	3.3/5	0.5–2.5	8.25
2.	HC-SR04 sensor	1	3.3/5	15	49.5
3.	Turbidity sensor	1	5	30 (Max)	150
4.	TDS sensor	1	3.3/5	3–6	30
5.	pH sensor	1	5	5–10	50
6.	Buzzer	1	5	<30	150
7.	LED	5	1.8–2.2	<20	22

4 Implementation

4.1 LoRa Node, Gateway, and Cloud Configuration

The work includes implementation of the hardware as per the architecture defined in the earlier section. The integration of hardware with the cloud is achieved by linking the LoRa gateway to the ThingSpeak channel to monitor the data from the nodes. In this work, two sensor nodes are designed to monitor two tanks simultaneously. There are two fish farming tanks of equal dimensions and capacity. The height of the tank is 4 ft. (1.2 m), and the radius of the tank is 4 m as shown in Fig. 5a. The volume of the tank is 15,000 L. At present, there are 3000 small Rohu fishes in tank 1. Rohu fishes have a bony skeleton. After fishes grow to a certain level, half of the fishes will

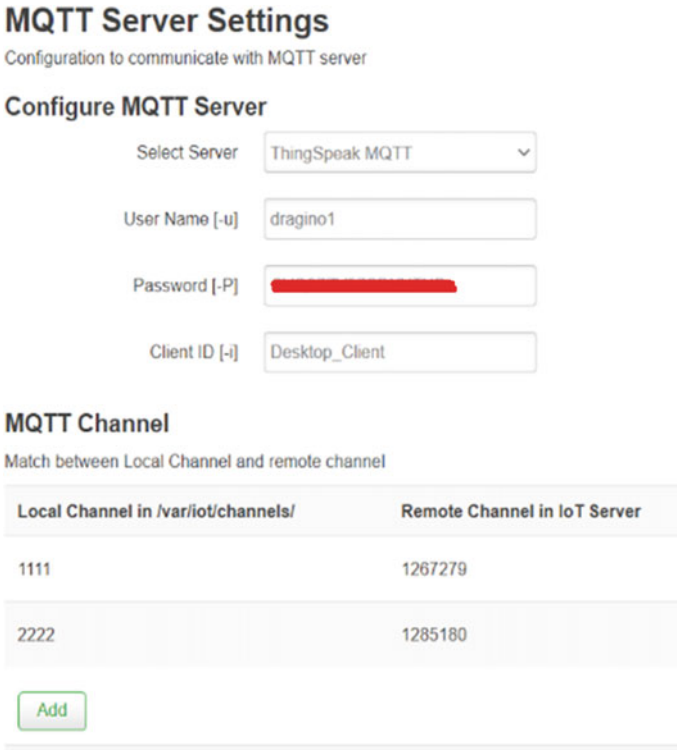


Fig. 8 Server settings in LoRa gateway to set up two MQTT channels

be transferred to tank 2. Then, both the tanks will have 1500 fishes approximately. Manual monitoring of the TDS and pH is performed using digital TDS meter and digital pH meter as shown in Fig. 5b, c respectively. The LoRa gateway has been configured to receive signals from two of the nodes with node name 1111 and 2222 to monitor two tanks in the fish farm. The write application programming interface (API) keys are also generated for each node by the gateway to communicate with the node. Two channels have been created with names fish farm node 1 (FFN1 named as 1111) and fish farm node 2 (FFN2 named as 2222) to monitor these two tanks in ThingSpeak cloud as shown in Figs. 8 and 9. Each channel has eight field to monitor and control as shown in Fig. 10. Lora nodes will transmit data packets with a spreading factor of 7, coding rate of 4/5, and signal bandwidth of 125 kHz as configured in the gateway.

Figure 11 is the setup of hardware under test in an indoor simulated environment. The setup was able to monitor the required parameters using LoRa wireless communication. The implementation of the proposed work is set to float on water in the fish farm as shown in Fig. 12. The water quality and quantity parameters received at the LoRa gateway are recorded in the logread section. The data will be in hex-

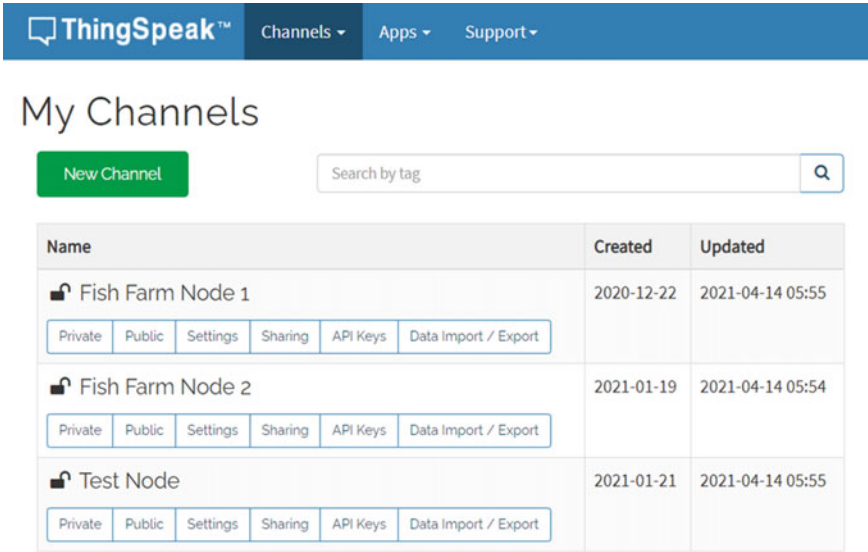


Fig. 9 Channels in ThingSpeak cloud to collect the sensor data and monitor status of motor

adecimal format here. The data can be seen in numerical values by converting that hexadecimal into a string. For every minute, one set of sensor values is updated in the cloud.

5 Monitoring System

The LoRa gateway kept in the monitoring room is connected to the Internet, and the data from the sensor node are pushed to the cloud by the gateway through the ThingSpeak Channel. Figure 13 displays the data being monitored in the cloud by the user (i.e., the farm owner).

6 Result and Discussion

The received data output of water quality and quantity parameters at the LoRa gateway under the logread section is shown in Fig. 14. The data will be in hexadecimal format here. The data can be seen in numerical values by converting that hexadecimal into a string. For every minute, one set of values is updated. The parameters monitored at the ThingSpeak cloud channels of the FFN2 are shown in Fig. 15. The data of temperature, humidity, water level, water quantity, pH value, turbidity, TDS, and motor status were obtained.

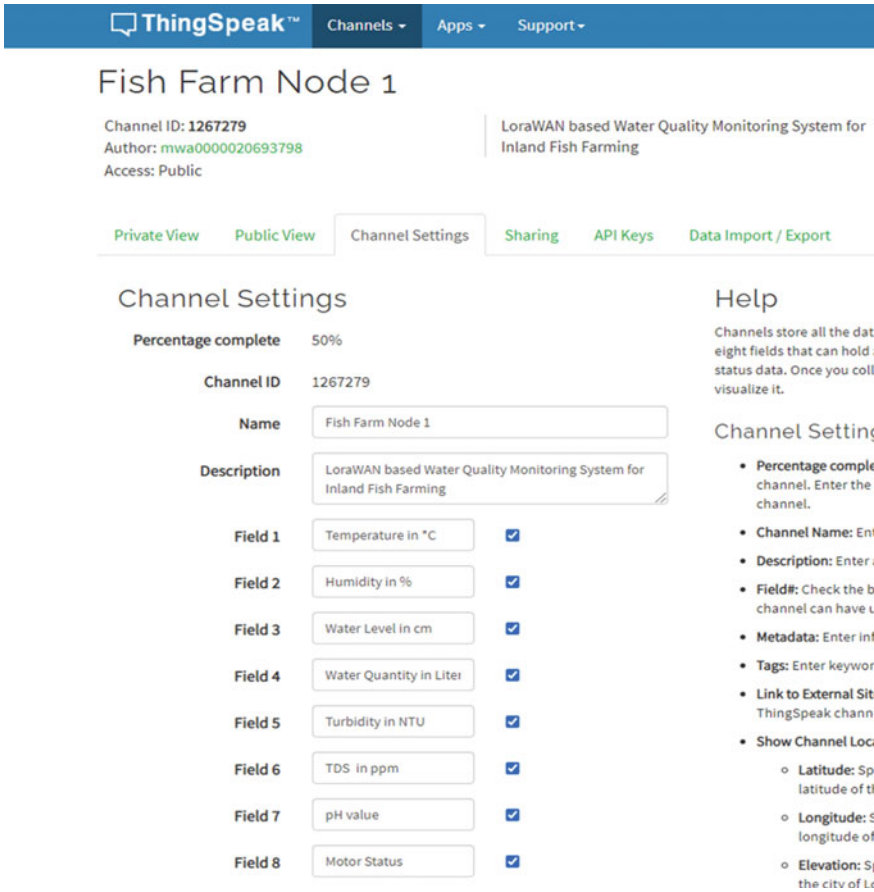


Fig. 10 Eight fields in channel 1 to monitor FFN1

6.1 Website to Monitor Fish Farm Tanks

The work includes design of a Website to monitor the data in the table format that displays the recently updated data in the cloud with an interval of thirty minutes. The data of all the parameters in multiple nodes can be seen on the developed Website (<https://fishfarming123.000webhostapp.com/index.html>). By clicking on the node at right top corner of the Webpage, the data from the corresponding LoRa node can be viewed as shown in Fig. 16.

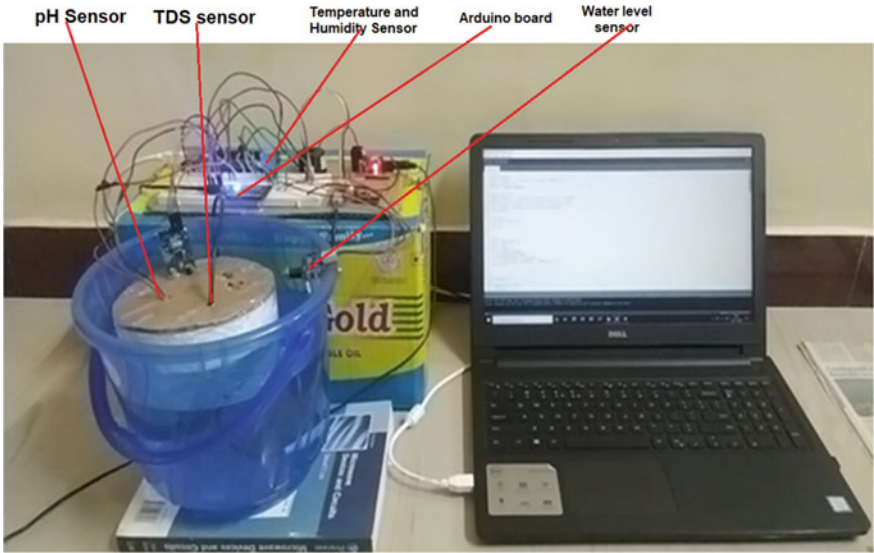


Fig. 11 Hardware implementation tested in a simulated environment

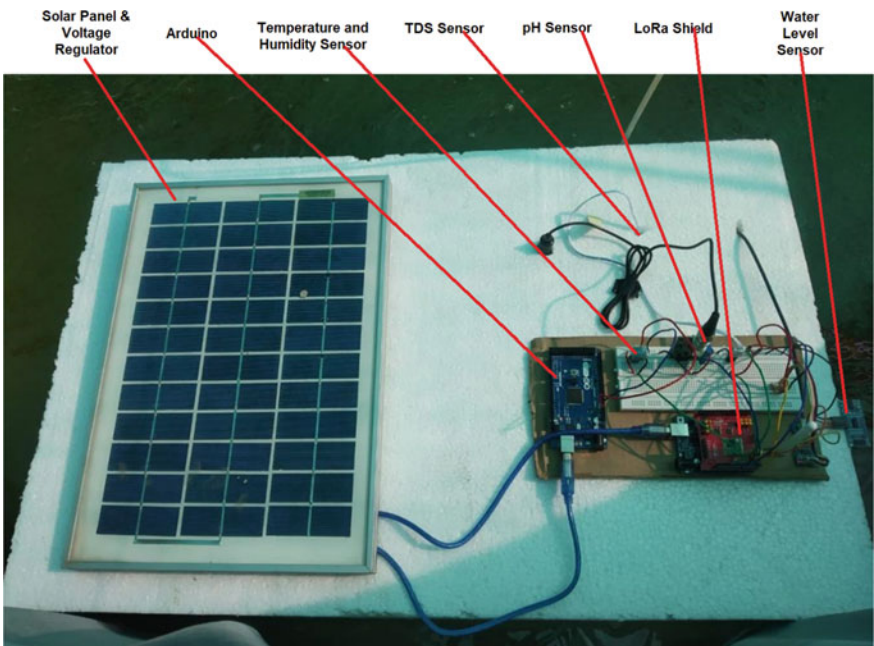


Fig. 12 Hardware implementation tested in a fish farm tank

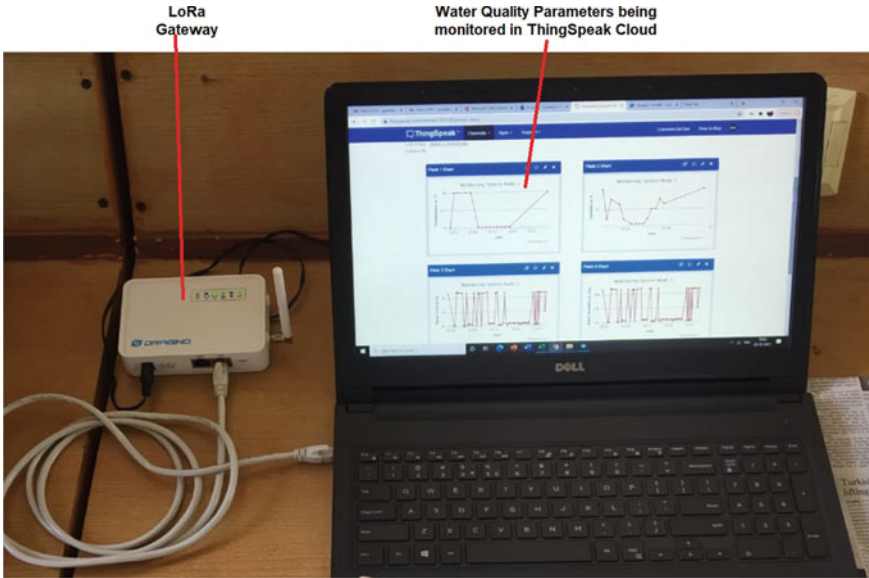


Fig. 13 LoRa gateway receives the data from two LoRa nodes and pushes the data to ThingSpeak cloud

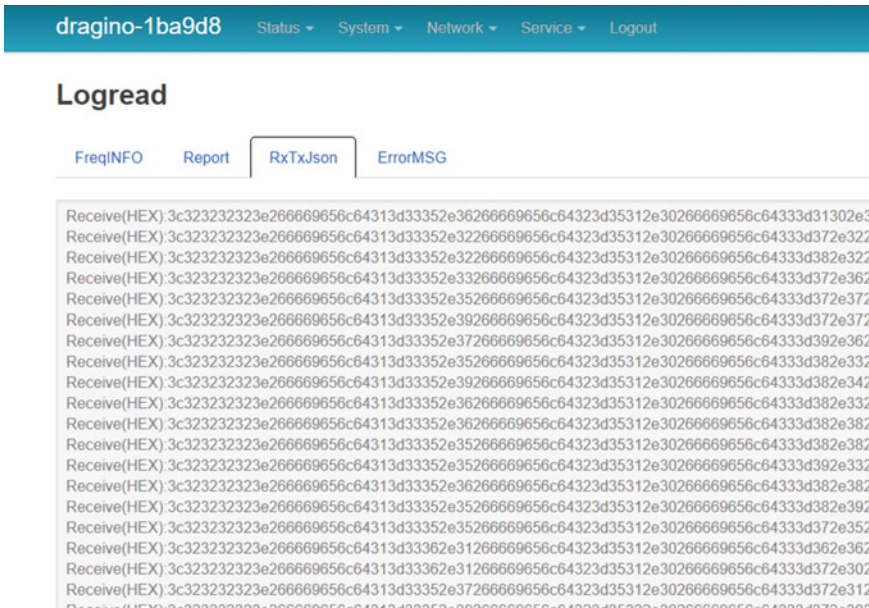


Fig. 14 Data received in LoRa gateway are observed using logread tab



Fig. 15 Plot of a temperature, b humidity, c water level, d water quantity, e turbidity, f TDS, g pH value, and h status of the motor

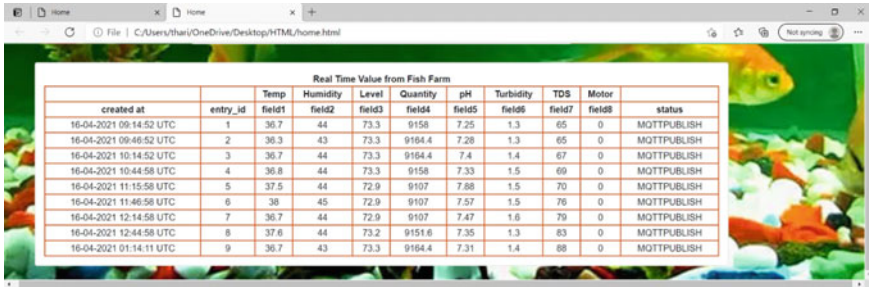


Fig. 16 Website design to monitor the fish farm

1		Temp	Humidity	Level	Quantity	pH	Turbidity	TDS	Motor		
2	created_at	entry_id	field1	field2	field3	field4	field5	field6	field7	field8	status
3	16-04-2021 09:14:52 UTC	1	36.7	44	73.3	9158	7.25	1.3	65	0	MQTTPUBLISH
4	16-04-2021 09:46:52 UTC	2	36.3	43	73.3	9164.4	7.28	1.3	65	0	MQTTPUBLISH
5	16-04-2021 10:14:52 UTC	3	36.7	44	73.3	9164.4	7.4	1.4	67	0	MQTTPUBLISH
6	16-04-2021 10:44:58 UTC	4	36.8	44	73.3	9158	7.33	1.5	69	0	MQTTPUBLISH
7	16-04-2021 11:15:58 UTC	5	37.5	44	72.9	9107	7.88	1.5	70	0	MQTTPUBLISH
8	16-04-2021 11:46:58 UTC	6	38	45	72.9	9107	7.57	1.5	76	0	MQTTPUBLISH
9	16-04-2021 12:14:58 UTC	7	36.7	44	72.9	9107	7.47	1.6	79	0	MQTTPUBLISH
10	16-04-2021 12:44:58 UTC	8	37.6	44	73.2	9151.6	7.35	1.3	83	0	MQTTPUBLISH
11	16-04-2021 01:14:11 UTC	9	36.7	43	73.3	9164.4	7.31	1.4	88	0	MQTTPUBLISH
12											

Fig. 17 Data exported from the cloud

6.2 Discussion

The proposed system shall be scaled up to monitor nearly 100 tanks of the capacity without worrying about the loss of data. When scaling for more number of tanks to be monitored, the spreading factor shall be increased accordingly between 7 and 12. As we increase the spreading factor, the area covered is increased, but the data rate will be decreased. The real-time monitored fish farm data of all the parameters are exported in an excel sheet, and it is shown in Fig. 17, along with time and entry id. One set of data is exported for 30 min, and thus, 9 sets of values were obtained for 4.5 h. The proposed work stands as a proof of concept that monitoring the water quality parameter can be implemented using LoRa wireless communication for the better maintenance of inland fish farms.

7 Conclusion

The design of LoRaWAN-based water quality monitoring and automation system for inland fish farming is implemented which is capable of monitoring the essential water quality parameters that are required for proper fish growth. The LoRa-based wireless sensor networks paved the way to build a smart monitoring system with

lower power consumption and with long-range secured data transmission. Up to 300 fish tanks can be monitored within a 1 km radius with this setup. By implementing this system in a fish farm, the water level and water quantity are always maintained at the required level. The increase in values of pH, TDS, and turbidity will be observed, and an indication for outlet pump for pumping out impure water and refilling pure water is given. Thus, proper growth and health of fishes are achieved. More yield and more profit to the fish farmer are obtained. The control mechanism provided in the system enables the fish tank to be filled with fresh water by turning on the motor; thereby, a balanced water quality is achieved to maintain a good environment for the fishes in the aquafarm tank. The strict monitoring system helps the fish farmers to avoid losses due to loss of fish lives because of poorly monitored fish farms. In future, machine learning algorithms can be developed to predict the growth and yield of fishes based on these water quality parameter values that were collected in the fish farm.

References

1. Abinaya T, Ishwarya J, Maheswari M (2019) A novel methodology for monitoring and controlling of water quality in aquaculture using internet of things (IoT). In: 2019 international conference on computer communication and informatics (ICCCI). IEEE, pp 1–4
2. Afshan S, Ali S, Ameen US, Farid M, Bharwana SA, Hannan F, Ahmad R (2014) Effect of different heavy metal pollution on fish. *Res J Chem Environ Sci* 2(1):74–79
3. Akila I, Karthikeyan P, Hari HM, Hari KJ (2018) IoT based domestic fish feeder. In: 2018 second international conference on electronics, communication and aerospace technology (ICECA). IEEE, pp 1306–1311
4. Bureau PI (2019) Press release, rrk/bk/ms/rk. Available at <https://pib.gov.in/PressReleasePage.aspx?PRID=1583635>. [Online]. Accessed 21 Sept 2021
5. Corbin JS, Holmyard J, Lindell S (2017) Regulation and permitting of standalone and co-located open ocean aquaculture facilities. In: *Aquaculture perspective of multi-use sites in the open ocean*. Springer, Cham, pp 187–229
6. Correspondent S. Study finds metal pollution in aquaculture farms. Available at <https://www.thehindu.com/news/national/study-finds-metal-pollution-in-aquaculture-farms/article33580769.ece>. [Online]. Accessed 21 Sept 2021
7. da Silva JG (2017) The future of food and agriculture: trends and challenges. Food and Agriculture Organization (FAO), Rome, Italy
8. Deng C, Gao Y, Gu J, Miao X, Li S (2010) Research on the growth model of aquaculture organisms based on neural network expert system. In: 2010 sixth international conference on natural computation, vol 4. IEEE, pp 1812–1815
9. Division FS (2019) Handbookonfs2018.pdf, Sept 2019. Available at <https://dof.gov.in/sites/default/files/2020-08/HandbookonFS2018.pdf>. [Online]. Accessed 18 Jan 2022
10. Hongskul V (1999) Into the next millennium: fishery perspective. RAP Publication, p 26
11. Iversen A, Asche F, Hermansen Ø, Nystøyl R (2020) Production cost and competitiveness in major salmon farming countries 2003–2018. *Aquaculture* 522:735089
12. Krishnaswamy K (2011) Dietary guidelines for Indians—a manual. Available at <https://www.nin.res.in/downloads/DietaryGuidelinesforNINwebsite.pdf>. [Online]. Accessed 21 Aug 2021
13. Krishnaswamy K (2011) Fish culture. Available at https://agritech.tnau.ac.in/fishery/fish_seafishes.html. [Online]. Accessed 18 Nov 2021
14. Lehane S (2013) Fish for the future: aquaculture and food security

15. Madhura S et al (2021) IoT based monitoring and control system using sensors. *J IoT Soc Mob Anal Cloud* 3(2):111–120
16. Maier RM, Gentry TJ (2015) Physiological methods, chap 11 . In: Pepper IL, Gerba CP, Gentry TJ (eds) *Environmental microbiology*, 3rd edn. Academic Press, San Diego, pp 213–243. <https://doi.org/10.1016/B978-0-12-394626-3.00011-9>. <https://www.sciencedirect.com/science/article/pii/B9780123946263000119>
17. Mizuta DD, Wikfors GH (2020) Farmed mussels in the northeastern US exclusive economic zone: an opportunity too good to ignore. *Fisheries* 45(4):207–213
18. Paul P, Basak N (2015) Present scenario, problems and prospects of inland fishing of West Bengal. *Indian J Appl Res* 5:56–58
19. Saha S, Rajib RH, Kabir S (2018) IoT based automated fish farm aquaculture monitoring system. In: 2018 international conference on innovations in science, engineering and technology (ICISSET). IEEE, pp 201–206
20. Salim TI, Haiyunnisa T, Alam HS (2016) Design and implementation of water quality monitoring for eel fish aquaculture. In: 2016 international symposium on electronics and smart devices (ISESD). IEEE, pp 208–213
21. Salin KR (2018) Inland aquaculture in Kerala; challenges and opportunities. Available at <https://krishijagran.com/agripedia/inland-aquaculture-in-kerala-challenges-and-opportunities/>
22. Simitha K, Raj S (2019) IoT and WSN based water quality monitoring system. In: 2019 3rd international conference on electronics, communication and aerospace technology (ICECA). IEEE, pp 205–210
23. Suthagar S, Tamilselvan K, Mageshkumar G, Muthupandian S, Vinod V (2019) Automated milk quantity and quality checking and vending machine. *Int J Recent Technol Eng* 8:4369–4372
24. Tamilselvan K, Murugesan G, Mageshkumar G, Suthagar S (2019) IoT based online transformer maintenance system. *Indian J Appl Res* 9(1)
25. Ward B (2013) Nitrification. In: Fath B (ed) *Encyclopedia of ecology*, 2nd edn. Elsevier, Oxford, pp 351–358. <https://doi.org/10.1016/B978-0-12-409548-9.00697-7>. <https://www.sciencedirect.com/science/article/pii/B9780124095489006977>

Afriendly: M-Health App as Strategy to Avoid Bullying



Genesis Dayana Pinto Almeida , María Cristina Páez Quinde ,
Carlos Alberto Ramos Guaña , and Carlos Fabián Martínez Vásquez 

Abstract The following research paper deals with the technology in mobile phones to identify school bullying in students who are constantly raped, which can be reflected verbally, psychologically, physically, and virtually through social networks. The study focuses on verifying the use of technology in health and harassment in schools by identifying conflicts between teachers and students and between students, in order to create an application to generate empathy in the classroom. The type of methodology is experimental–exploratory analyzing causes and effects, whilst the quantitative approach that attempts to arrive at a deductive process predominates. With regard to the collection of information, the questionnaire is used as a tool and the survey as a technique for students in the ninth grade of basic general education, as is the T.A.M. method (technology acceptance model). Therefore, the “AfriEndly” application proved satisfactory within the research according to the T.A.M. method, which revealed that the technology is accepted within the study population. Keywords: Educational harassment, Mobile applications, Mobile health, Self-esteem.

Keywords Application · Mobile health · Mental health · Bullying

G. D. P. Almeida (✉)
High School “Nuevo Mundo”, Ambato, Ecuador
e-mail: pintodayana72@gmail.com

M. C. P. Quinde
Instituto Superior Tecnológico España, Ambato, Ecuador
e-mail: mc.paez@uta.edu.ec

C. A. R. Guaña
Ez English School, Ambato, Ecuador

C. F. M. Vásquez
MecanicS.A./Ambato, Ambato, Ecuador

1 Introduction

Miranda et al. [1] in his research, “Teachers’ Training: From Pedagogical Tradition to Virtual Learning Environments” highlights the creation of virtual spaces according not only to the subject, but also to the current context, suitable scenarios for the student, teacher, and the conformity of parents. ICTs not only contribute to an improvement toward education as the simple fact of imparting knowledge, but also to a refinement to access various intelligences such as emotional intelligence, which is deeply investigated in the teacher, because by being emotionally intelligent, it favours the teacher’s learning, the educational community, providing solutions focused on technology, from the current context of its environment leading to psychological and learning problems [2].

Thus, in the face of the health emergency due to COVID 19, it has caused chronic illnesses and even death, for which reason society has been forced to lead a life through virtuality, stressing the importance in educational, employment, and even social continuity, from that time onwards. Society has become a subject of transformation moving from paper to computer or cellular. Information and communication technologies (ICT) are therefore highlighted in education as a strategy for improving quality and adapting to the current environment, and this would require changes not only in methods, but also in instruments, i.e. in the use and benefit of education, control and management of communication and information technology [1].

Additionally [3], some research has demonstrated the results of mobile well-being applications toward the subject, concluding that they often provide tolerance within distress, security within crisis, as well as future projection toward opportunities for improvement. Emotionally and most importantly feel support in front of a mobile health app. For that reason, applications motivated to contribute to the well-being of mental health as well as physical must consider factors that protect the subject, promoting personal worth, increasing self-esteem, as well as interpersonal relationships, in addition to recognizing his own achievements and recognition of the need for professional assistance, which leads to an openness to experience within one’s own communication skills. Bullying School bullying has no gender, it can be male or female; nor time, as it was evident from the biblical case of Cain and Abel, in which there was physical violence by a sibling; nor age, since it can be present in children, young people and adults, and involves anyone person who mistreats physically, through beatings; verbally, using insults to intimidate cyber, with manipulation of social networks, as most abuses occur at an early age, usually against the weakest person, implying unequal violence for obvious differences [4].

Because of the presence of school bullying, according to Smys and Raj [5], three criteria should be considered as follows: The first is when a negative event against a child occurs repeatedly; in the second, there must be instability in both the bullyer and the victim; and in the third, consequently, a discrepancy develops between the two. This investigation, after collecting data, revealed cases where there is almost no intervention by teachers and parents, with consequences of suicides, including firearm revenge attacks, on the part of the harassed.

According to Smys and Raj [5], if school bullying is to be combated, it is essential to point out the role that each teacher must play, as well as to consider that interrelationships are important so that those involved, such as the harassed, the victim, or pedagogue, can develop, put forward projects, and projects have resources that contain the aim of participatory intervention by those involved, addressing cooperative changes in the environment and context in order to avoid school violence.

Self-esteem is associated with mental health, as it is the comfort it provides to a person's personality; however, without such health, the consequences become serious, such as drug addiction and abuse, as well as school failure, loneliness, and other problems moreover, with the passing of time and without professional help, they worsen, resulting in a negative self-assessment, in which they may come to be considered insecure and incapable, provoking susceptibility to criticism and responding ineffectively to their surroundings. Social is to be an important external factor, since one person's perception of the other is estimated and influences his well-being. [6]

2 Methodology

The A.D.D. methodology was used for the development of the study, i.e.: analysis; design; implementation; and evaluation. Analysis (Fig. 1):

We analyzed as a pre-test the structured questionnaire addressed to students of ninth-grade E.B.G. applied during hours of virtual classes. For this, the Mobincube tool was an important factor to consider within the study, since it helped develop an effective and feasible free app according to various users of YouTube platforms (Fig. 2).

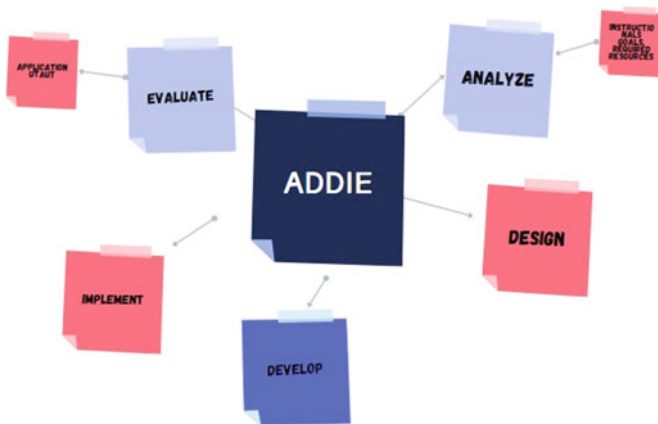


Fig. 1 Mobincube application



Fig. 2 Mobincube application

Design: The purpose of the Mobincube interface is to interact the operation of other systems in an easy way according to their use and effective between subject-computer, which was conceived of relevance in generating an interface not only feasible to use, but also optimum to any individual of different age. For this purpose, in the development of the app, a “menu” was considered with 4 important buttons, such as the harassment detector, the attacker, the victim, the suggestions the app provides, and the credits.

Development: Mobincube offered us a free and easy to use software, thanks to the fact that it is not mandatory to insert code, and whose development was with block programming, accessing interfaces with interactive images, icons, text, and tables (Fig. 3).

Harassment Detector: The main menu consists of a button called “harassment detector,” which leads to a “test” Web page, which gets seven questions that the user can answer in order to get at the end of the survey a result, i.e. the page will help you to know if you suffer from bullying or not (Fig. 4).

Suggestions: The option called “recommendations” leads to a new screen showing us two images, the young people guide us to a screen of suggestions that students can make if they suffer from harassment, whilst the teacher’s image sends us to a screen where he suggests their harassment inside the screen (Fig. 5).

Aggressor: It consists of a representative image of the aggressor, and below it is seven behavioural traits that the aggressor possesses, the instant that the text of each descriptor is crushed leads to a new page where it describes each of these characteristics. Victim: It consists of an image representative of the harassed person and beneath, and it is the behaviours that the victim usually has. Similarly, as soon as each textual feature is crushed, a new screen is opened specifying and detailing the attitude (Fig. 6).

Seek help: Within this, option is the emergency number known as 911, so that students can call when they feel intimidated.



Fig. 3 Main menu-harassment detector developed



Fig. 4 Suggestions developed



Fig. 5 “Aggressor” and “victim” behaviour

Evaluation: The development of the technology acceptance model (T.A.M) methodology was considered fundamental in order to determine whether the application was rejected or accepted within the study population, in order to investigate the impact of technology on society and how the population perceives such utility.

3 Results

1. Learning to use gamification and technology tools easy for me?

See Table 1 and Fig. 7.

Analysis and Discussion

Out of a total of 117 students representing 100% of the sample, 33.3% equivalent to 39 students agree that the use of technological tools is easy to use, 32.5% totally agree, 16.2% undecided, 10.3% disagree total disagreement with the 7.7% reported by nine pupils. This allows us to recognize within the research that the majority of E.G.B. ninth students agree that learning to use gamification and technology tools is easy.

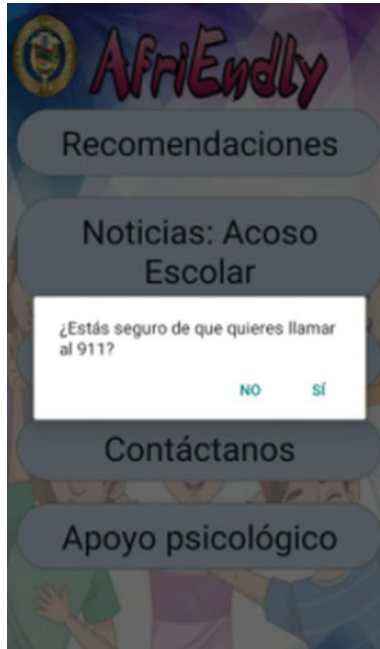


Fig. 6 Seek help developed

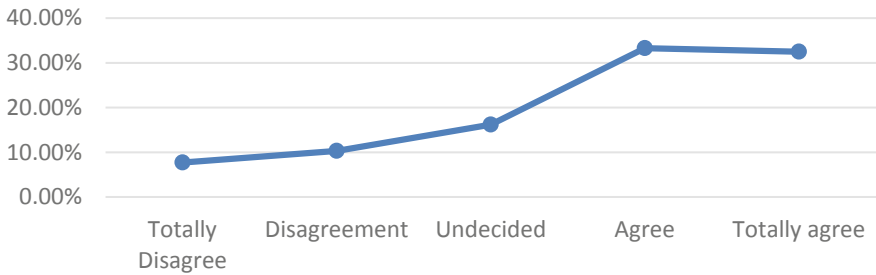


Fig. 7 Gamification and technological tools frequency. **By:** Pinto D. (2022)

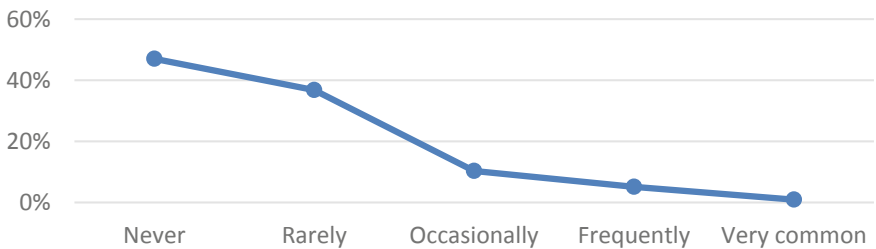


Fig. 8 Offensive comments. **By:** Pinto D. (2022)

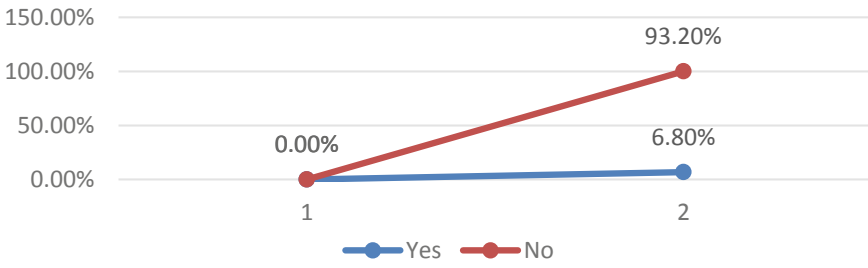


Fig. 9 Application to detect school bullying. **By:** Pinto D. (2022). **By:** Pinto, G (2021)

Table 1 Gamification and technological tools frequency percentage

	Frequency	Percentage
Totally disagree	9	7.7
Disagreement	12	10.3
Undecided	19	16.2
Agree	39	33.3
Totally agree	38	32.5
Total	117	100.0

2. How often have you been offended by comments made by your friends, either through social media or in person?

See Table 2 and and Fig. 8.

Analysis and Discussion

Out of a total of 117 students corresponding to 100% of the sample, 47.01% equivalent to 55 pupils have never been offended by any comment, 36.75% rarely, 10.26% occasionally, 5.13% frequently reported by 6 pupils, and 0.85% very often reported by 1 pupil. This allows us to recognize within the research that most of the ninth E.G.B. students have been offended by some comments made by their peers, either

Table 2 Offensive comments

Ítems	Frequency	Percentage
Never	55	47.0
Rarely	43	36.8
Occasionally	12	10.3
Frequently	6	5.1
Very common	1	0.9
Total	117	100.0

Table 3 Application to detect school bullying

	Frequency	Percentage
Yes	8	6.8
No	109	93.2
Total	117	100.0

through social media or in person, yet 43 of them indicated that they have rarely been disgusted by any comments.

3. Are you aware of an app that detects bullying in schools or identifies bullying?

See Table 3 and Fig. 9.

Analysis and Discussion

Out of a total of 117 students, representing 100% of the sample, 93.16%, representing 109 students, pointed out that they were not familiar with any application to detect harassment, and 6.84% indicated that they were aware of it. This allows us to recognize within the research that the majority of students in the ninth year of E.G.B. have no knowledge of any application that detects bullying in schools or identifies bullying.

Verification of Hypotheses

For this research, it was relevant to check through the Kolmogorov–Smirnov statistician or K-S, through the two questions representative of the study of the research; the questions are based on the object of study and the results manifested by the experiment. Question number 17: Have you heard the term mobile health? Identified with the mobile health variable, and question 5: How often have you been offended by a comment made by your colleagues, either through social networks or in person? represented by the dependent variable (school harassment) (see Table 4).

4 Conclusions

Bullying is a relevant factor to be observed in classrooms, since the consequences it causes on students such as physical, verbal, psychological, and today cyber-assaults are serious, as they can cause physical discomfort, stress, stomach, and even headaches, constant nightmares, underperformance, and exclusion in the classroom.

Usually, the victim who is subjected to various forms of violence by his or her aggressor suffers silently, as self-esteem and energy deteriorates with each aggression and humiliation. Students in the three ninth-grade parallels of general basic education do not suffer from harassment from teachers teaching classes during the academic period, however, he was able to identify that occasionally the majority of students feel intimidated by their friends.

Table 4 Kolmogorov–Smirnov test for a sample

	Have you heard the term mobile health?	How often have you been offended by some comment made by your friends, either by mobile social media or person
N	117	117
Normal parameters ^{a,b}	Media	1.94
	Standard deviation	0.238
More extreme differences	Absolute	0.539
	Positive	0.401
	Negative	− 0.539
Z de Kolmogorov–Smirnov	5.834	2.941
Sig. asintót. (bilateral)	0.000	0.000
(a) The contrast distribution is normal		
(b) They have been calculated from data		

Technology nowadays is used not only for interaction, but also for recommendations of various topics that applications typically provide. Thus, the application developed was in order to create a guide to help the victim, recommending various ways of help such as telephone numbers, and was even a guide for teachers to know the different actions they can take to avoid harassment and foster empathy.

References

1. Miranda JJS, Polo AG, Rodríguez AM (2020) Teachers’ training: from pedagogical tradition to virtual learning environments. *RIDE. Iberoam. Investig. Desarro*
2. Fuentes NIGAL (2019) Autoestima, Optimismo y Resiliencia en Niños en Situación de Pobreza. *Revista Internacional psicológica*
3. Mishra A, Aiswal A, Chaudhari L, Bodade V (2022) Health record management system—a web-based application. *J ISMAC* 4:301–313
4. Diaz YC (2019) EL ACOSO ESCOLAR: BULLYING. *Redalyc* 13
5. Smys S, Raj JS (2022) Future challenges of the internet of things in the health care domain—an overview. *J Trends Comput Sci Smart Technol* 3(4):274–286
6. González SP (2019) Metaphors for teachers’ roles in a bullying at school intervention. *Psicoperspectivas* 11

Machine Learning Techniques for Automated Nuclear Atypia Detection in Histopathology Images: A Review



Jithy Varghese and J. S. Saleema

Abstract Nuclear atypia identification is an important stage in pathology procedures for breast cancer diagnosis and prognosis. The introduction of image processing techniques to automate nuclear atypia identification has made the very tedious, error-prone, and time-consuming procedure of manually observing stained histopathological slides much easier. In the last decade, several solutions for resolving this problem have emerged in the literature, and they have shown positive incremental advancements in this field of study. The nuclear atypia count is an important measure to consider when assessing breast cancer. This work provides a comprehensive review of automated nuclear atypia scoring process which includes the current advancements and future prospects for this critical undertaking, which will aid humanity in the fight against cancer. In this study, we examine the various techniques applied in detecting nuclear atypia in breast cancer as well as the major hurdles that must be overcome and the use of benchmark datasets in this domain. This work provides a comprehensive review of automated nuclear atypia scoring process which includes the current advancements and prospects for this critical undertaking, which will aid humanity in the fight against cancer.

Keywords Breast cancer · Machine learning · Nuclear pleomorphism · Histopathological image · Nuclear atypia detection · Deep learning

1 Introduction

Cancer is a phrase that refers to diseases that cause abnormal cell proliferation in the body, resulting in tumours [1]. It can spread to other organs of the body from the initial affected site. Every year, cancer claims the lives of huge number of people

J. Varghese (✉) · J. S. Saleema
Department of Computer Science, Christ Deemed to Be University, Bengaluru,
Karnataka 560029, India
e-mail: jithylijo@caias.in

J. Varghese
Christ Academy Institute for Advanced Studies, Bengaluru, Karnataka 560083, India

throughout the world. According to the statistics report of World Health Organization (WHO), the second major reason for mortality in the world is cancer [2].

According to the GLOBOCAN 2018 [3] report, 24.2% of cancer diagnosed amongst women all over the world is breast cancer, and 15% of all cancer-related death is due to breast cancer. Cancer patients' chances of survival can be greatly improved by early detection and treatment [2]. For better early detection of cancer, intensive research investigations are being done worldwide [4]. Many medical multi-imaging modalities are utilized screening and classification of breast cancer, including histopathological images taken through biopsy, digital mammography, sono-mammography, magnetic resonance imaging, and computerized thermography [5].

As biopsy is a clinical procedure of tissue analysis for the screening of cancer [5]. Tissues from vulnerable areas are carefully removed and mounted on microscope slides. These images are called histopathological images. These images can be safely preserved in digital format and transmitted over the Internet for future study [6]. Digital pathology has helped researchers to apply computer-assisted technology in detection of biomarkers in cancer research [2]. The tedious and time-consuming tasks of pathologists can be delegated to computers, and it can be accomplished with high amount of accuracy [7–9]. Whole slide imaging (WSI) technology was introduced to read and store the whole slide at very high magnification [2, 10]. It also aided in the creation of many region of interest (ROI) images that were used to diagnose breast cancer as malignant or non-cancerous [11]. A sample of a high-power field (HPF) image of a breast cancer biopsy is shown in Fig. 1.

The cancer detection method in histopathological images usually includes classifying the image biopsy as malignant or non-cancerous. Pathologists classify the images based on criteria such as form, colour, cytoplasm proportion, and size of the nucleus of the cell. The contrast between normal and malignant cells is depicted in Fig. 2 [12]. Score of nuclear atypia provides an estimate of the prognosis of disease

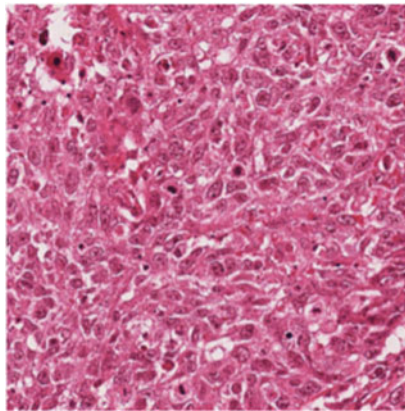


Fig. 1 HPF image of breast cancer histopathological image from MITOSIS-ATYPIA dataset

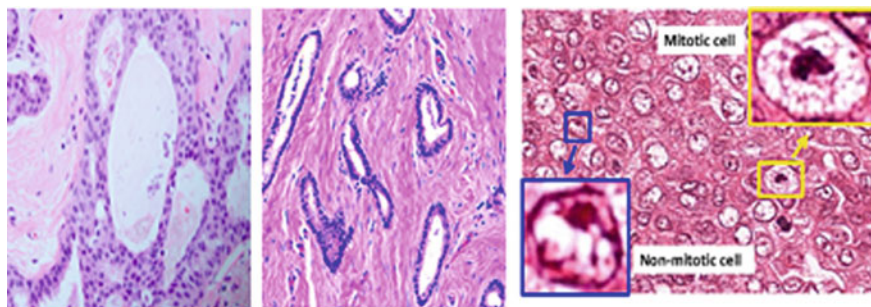


Fig. 2 a Nuclear atypia, b tubule formation, c mitotic cell [13]

in the patient and aids in the development of specific treatment strategies for the patient. According to the number of mitotic cells, tube formation, and nuclear pleomorphism, the cancer is classified as first, second, and third grade breast cancer. Computerized nuclei delineation and classification, which are usually required for cancer diagnosis and grading, are a time-consuming operation made more complex in abnormal images by the majority of nuclei's complex and irregular shape and size.

1.1 Grading and Staging

Grading and staging are the two major steps in the cancer treatment plan [14]. The degree to which malignant tumours resemble the surrounding tissue is graded. The main purpose of cancer grading is to identify the disease's aggressiveness. The malignancies which can be differentiated well have a better prognosis than poorly differentiated ones because they are less aggressive [2]. The tumour's aggressiveness is a measure of its rate of growth and extension to other organs in the body. Staging in cancer determines up to what extent the growth has advanced from the initial affected organ to other organs of the body. Oncologists use grading and staging to finalize the treatment options and anticipate disease progression [2].

The stage of the breast cancer can be understood using tumour, node, metastasis (TNM) staging [2]. The stage of a tumour is determined by its size, spread, and location. Stage T describes the size and spread of the tumour in the primary site and adjacent organs. Stage N indicates the existence of cancer cells in the lymph node [15]. This is an indication that the cancer is spreading. The lymphatic system is the primary carrier which moves cancer cells to other organs of the body. Stage M specifies that the cancer has extended to an organ other than the one where it was diagnosed first [2].

According to Nottingham grading system (NGS) [2, 16], mitotic count, nuclear pleomorphism (atypia), and tubule development are the three major characteristics utilized for grading breast cancer. Nuclear atypia count is subjective in nature where

Table 1 Breast cancer grading characteristics by NGS [2]

Parameter	Score	Criteria
Mitotic count	1	< 10 mitotic cells
	2	10–19 mitotic cells
	3	≥ 20 mitotic cells
Nuclear pleomorphism (nuclear atypia)	1	Small and similar type nuclei
	2	Moderately variable nuclei shape and size
	3	Many nuclei with clear variation
Tubule formation	1	Tubule forms (> 75%)
	2	Tubule forms (10–75%)
	3	Tubule forms (< 10%)

the grading depends on the proficiency of the pathologist. Table 1 shows breast cancer grading characteristics by NGS [2, 16].

1.2 Need for Automated Nuclear Atypia Detection

A pathologist analyzes stained histology slide under a microscope and manually assigns a score to every criteria in NGS. A human grading approach is difficult, tedious, and erroneous due to the inequality in the appearance of the cells and numerous cells per high magnification field (HPF) [2]. These factors contribute to a high amount of diversity in pathological results between observers [17]. Due to lack of qualified pathologists in many nations, early cancer detection is challenging. As a result, cancer death rates in developing countries are high [3, 18]. In this case, automated diagnostic methods can help with a quicker and more effective diagnosis. In this case, automated detection is the best option. In this case, automated diagnostic technologies can help with rapid diagnosis and accurate grading in order to define the optimal treatment strategy, lowering the cancer-related death rate significantly [19].

1.3 Challenges in Nuclear Atypia Detection

In histopathology image analysis, detecting nuclear atypia automatically is a difficult task. Dissimilarity in texture shape and size of the nuclear atypia makes the detection of nuclear atypia detection tough [2]. There are several cellular formations that resemble nuclear atypia, but they are not atypia. Such resemblance increases the false positive rate and thereby decrease the accuracy of classification [2, 20]. In the manual staining of histology slides, there may be variation in the staining due

to the human involvement. Apart from this, the colour intensity of the images varies depending on the scanner used. Aperio and Hamamatsu are the most common type of scanners used in histological images, and they have a significant amount of variation in colour intensity. So if we use colour component in feature extraction, then it might affect the overall accuracy of classification process. To avoid this, various colour normalization techniques are used in nuclear atypia detection approaches [2].

1.4 Dataset in the Public Domain

The researchers started working more in the field of nuclear atypia detection due to the public dataset availability through open challenges. The most common histopathological datasets available in the public domain for nuclear atypia scoring as per the literature are given in the Table 2 [2].

Despite the fact that several datasets for breast cancer histopathological image analysis such as TUPAC 2016, MITOSIS2012, MITOSIS ATYPIA14, AMIDA, BRAEKHIS DATASET10, and BreCaHAD 2019 but none of them contain labelled images for nuclear atypia detection. The MITOS-ATYPIA14 challenge dataset is the best suited dataset for our topic of interest in this regard. As a result, in our comparative study, we employed the MITOSIS-ATYPIA14 dataset.

In this review paper, the major emphasis is given to nuclear atypia characteristics which is used to grade breast cancer. The paper focusses on the various techniques and procedures for grading breast cancer, assesses the challenges that these techniques encounter, and explores the strategy adopted by image analysis methods to overcome these challenges. These tools assist in determining topics that have not been thoroughly investigated in the field of nuclear atypia detection, which is where future study will be focussed. The paper also discusses the various evaluation metrics

Table 2 Public domain datasets for nuclear atypia scoring [2, 21]

Public dataset	Year	Explanation	Image size
MITOS-2012 [22]	2012	50 HPF images at 40 × magnification	A:2084 × 2084 H:2250 × 2252 M:1360 × 1360
AMIDA [23]	2013	311 HPF of 23 subjects at 40× magnification	2000 × 2000
MITOS- ATYPIA14 [24]	2014	1136 HPF images at 40× magnification	A:1539 × 1376 H:1663 × 1485
TUPAC 16 [25]	2016	73 breast cancer of different patients at 40× magnification	2000 × 2000
BREAKHIS DATASET10[26]	2018	9109 WSI of 82 patients	700 × 460
BreCaHAD 2019[27]	2019	162 WSI of 82 patients	1360 × 1024

used in analyzing the detection of nuclear atypia. The following sections make up the document. The first section contains an introduction, breast cancer grading, and the need for nuclear atypia detection, the second section contains a literature review in the field of nuclear atypia detection, the third section summarizes the different evaluation performance measures used for the process of nuclear atypia scoring, the fourth section does an analysis of existing techniques used in the field, and the fifth section gives future prospects in the field of grading of breast cancer [2, 27].

2 Literature Review

For the detection of nuclear atypia, there are two types of histopathological image analysis techniques. They are hand crafted feature-based algorithms that have been constructed and feature-based algorithms. In handmade feature-based approaches, image features must be retrieved precisely. Pre-processing of image, nucleus segmentation, feature extraction, and classification operations is all part of the handcrafted feature-based method.

Pre-processing techniques such as noise smoothing, intensity normalization, colour separation, thresholding, stain normalization, reduction, and augmentation are used to improve the characteristics of histopathological images. After pre-processing, nucleus segmentation is done [21]. For nuclei segmentation, many techniques, such as active contour model, mean shift, and Gaussian model watershed, [28–31] are frequently utilized. One of the most essential requirements for grading breast cancer in histopathology images is extraction of features of biological structures such as lymphocytes and cancer nuclei. The morphological characteristics and shape and size of these structures are frequently utilized as indicators of disease severity. The collected features are given into the classification step, which analyzes them statistically and typically uses machine learning techniques to categorize them into multiple classes. K-means clustering, SVMs, Bayesian classifiers, and artificial neural networks (ANNs) are some of the most often used classifiers. The steps of handcrafted feature-based nuclear atypia detection are shown in Fig. 3.

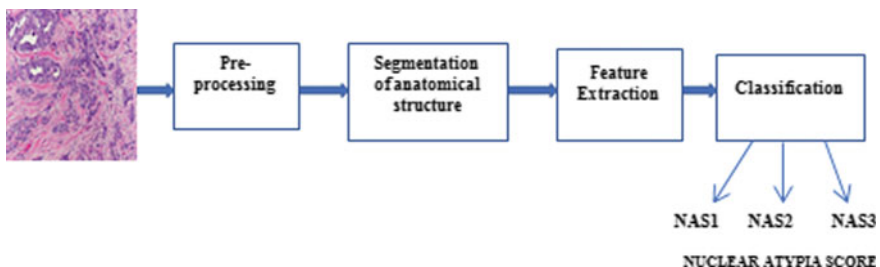


Fig. 3 Steps of handcrafted feature-based nuclear atypia detection

The second set of histopathological image classification methods includes trained feature-based algorithms that extract high-level features without direct feature extraction operation from histopathology images. Since the initiation of convolutional neural networks (CNNs) learned feature-based approaches received a lot of attention of researchers. The majority of newly developed algorithms for breast cancer grading [21] use convolutional neural networks (CNNs), transfer learning (TL), and residual networks (RN) concepts.

2.1 Handcrafted Feature-Based Algorithms

2.1.1 Pre-processing Techniques

Pre-processing removes undesired distortions from images and improves very relevant feature required for grading of histopathological images. The initial step in the pre-processing is elimination of noises and artefacts from the images which has got in at the time of slide preparation [32]. Apart from this stain normalization, colour separation is also done to improve the quality of the image [30, 33]. Prior knowledge on the stain vectors is used for stain normalization. Otsu thresholding, clustering, and trust region optimization techniques are used for stain optimization. After stain normalization, de-convolution technique is used for colour separation [34].

An adaptive technique was proposed in [35] to decrease colour variation. Colour unmixing or colour de-convolution, a special instance of spectral unmixing, is used by Veta et al. [31] and Basavanhally [36] to separate the stains. In paper [30], stain normalization and colour separation are the pre-processing steps used. Wan et al. [37] perform stain normalization which used non-linear mapping technique to adjust image intensity discrepancies due to variations in tissue preparation.

2.1.2 Segmentation of Anatomical Structures

Nuclei identification and segmentation are important processes in cancer diagnosis. Different features of nuclei, such as their morphological structure, size, and mitotic nuclei count, are crucial for identifying the disease and understanding the malignancy and intensity of the disease. Mainly, there are three types of nuclei segmentation algorithms utilized in breast cancer grading. They are threshold-based approaches, region growing-based techniques, and boundary-based techniques.

In an image, threshold can be used for distinguishing foreground objects from the background objects. Choosing a right threshold value helps to convert the image into a binary image which give adequate knowledge about the shape and size of the nucleus [38]. This makes the feature extraction and classification process comparatively easy due to the reduced complexity of the images. Due to its simple concept, many breast cancer classification systems are used thresholding approach for nuclei segmentation [21].

In [39], Petushi et al. use a combination of local morphological methods and optimum adaptive thresholding. Later in [21, 40], Petushi et al. combine edge leveling and morphological filling technique to determine the appropriate threshold value. Weyn et al. [41] combined background smoothing and thresholding based on the median of histogram intensity to get the contour of nuclei. Moncayo et al. [42] use the hematoxylin contribution map and maximally stable extreme regions to segment the nuclear region (MSER) where different thresholds are applied to the image, and the areas which reflect the least change are labelled as maximally stable extreme regions. The nuclei and the surrounds are then distinguished using open and close morphological operations.

Another method for image segmentation is region growth, which segregates neighbourhood pixels based on homogeneity and adds them to a region where the similarity criteria of the class are met. This method is applied for each of the pixels in the neighbourhood region. The technique of region growing begins by choosing a seed location based on certain criteria [21]. These discovered seed sites are then used to expand these regions depending on some conditions. In noisy images with difficulty to detect boundaries or edges, region growing algorithms are the best options.

After segmentation, [29] performs neoplasm localization using a multi-resolution approach. Cell segmentations are performed on high-resolution pictures using Gaussian colour models, whilst [31] uses a marker controlled watershed segmentation approach. To discover the prospective nuclei, pre-processed images with indicated locations with strong radial symmetry and local minima are used. The water shed segmentation algorithm is then utilized to determine the nuclei's contour. In [43], the external margins of nuclei are segmented using a convex grouping technique that is best suited for patchy and open vesicular nuclei found in breast cancer with a high aggressiveness [21]. This irregular nuclei structure is segmented using the K-means clustering algorithm [44].

The edge and intensity discontinuity in image are important features that provide knowledge about object contours and are extensively utilized in image segmentation and item identification. For nuclei segmentation of breast histopathology image, many edge-based segmentation techniques have been used [21].

Cassato et al. [35] apply a difference of Gaussian (DoG) filter to the image. The resultant edge map is then extracted using the Hough transform. With the edge map, the boundary of the nuclei is defined using an active contour model. Faridi et al. [45] used morphological feature extraction procedures and DOG filtering to detect nuclei's centre and the distance regularized level set evolution (DRLSE) technique to recover nuclear border [21]. In paper [46] to find the candidate cell nuclei, the author has employed morphological techniques and a distance transform. The thresholding technique is applied to a gamma-corrected image to generate a binary image before being subjected to morphological procedures like dilatation and erosion. Paper [47] uses thresholding and morphological filtering techniques for pre-processing and later a snake-based method on an image to retrieve nuclear edges [21]. A polar space transformation is conducted, and then, the nuclei boundary is identified with iterative snake technique. For nuclei segmentation, Basavanhally et al. [36] developed a colour

gradient-based geodesic active contour (CGAC) technique. In the gradient-based technique, segmentation is achieved by minimizing an energy function that includes a typical geodesic active contour [21].

For automatic segmentation of nuclei, many algorithms have been developed and experimented. Due to the errors at the time of image capture and the diverse shape and size of the nucleus, finding the region of interest and segmenting is considered as a difficult task. The accuracy of the nuclear atypia detection technique is strongly dependent on the success of the segmentation technique.

2.1.3 Feature Extraction

A crucial phase in the handcrafted feature-based technique, in which the most significant features of an image are retrieved from a big number of features, is called feature extraction. Excess cell proliferation in cancer tissues is caused by disruptions in the cell life cycle, resulting in impaired cellular function [21]. The various grades of cancer have different cellular characteristics. The majority of these characteristics are graph-based topographical, morphological, and textural features. Various feature extraction methods available in the literature for breast cancer grading are covered in this section [21]. When compared to normal cells or nuclei, malignant cells differ in size and shape, and pathologists utilize this difference to grade malignancy. The form and size of a cell are frequently revealed by morphological trait. Smoothness, symmetry, length of axes, roundness, and concavity are used to describe their shape, whereas the radius, perimeter, and area of segmented nuclei are commonly used to describe their size. For cancer identification and grading, morphological aspects of the nucleus structures have been extracted and used by various algorithms.

For each of the images, Petushi et al. [39] calculated standard deviation, area, intensity mean, the minimum intensity value, major and minor axes of the nuclei that have been segmented. The feature vector generated is given for clustering along with a binary decision tree. The features like mean value of intensity, area, and count of nuclei are extracted from the source image in paper [45]. For cancer grading and prediction, Veta et al. [31] considered the standard deviation and mean of the area of the segmented nuclei for feature extraction.

Textural characteristics provide crucial information on the intensity variation of pixel over a surface with respect to smoothness and regularity. Statistical, spectral, and structural approaches are commonly used to extract textural information. This section discusses about the papers which use various texture features for cancer grading and classification.

In [40], texture features which help to calculate the density of cell nuclei are used in supervised classification approaches such neural networks, decision trees, linear classifiers, and quadratic classifiers. For nuclear grading of breast tumours, Khan et al. [48] developed a textural-based characteristic called the geodesic mean of region covariance descriptors (gmRC). In the next phase, K-nearest neighbour (KNN) classifier is applied on gmRC matrix. Ojansivu et al. [49] used descriptors such as local phase quantization (LPQ) and local binary patterns (LBPs) to create

histograms that describe the image's statistical textural qualities. Later, support vector machine classifier is applied on the pre-processed image data for categorizing the nuclear atypia score. A bag of features (BoF) with multi-scale descriptors is used to represent the discovered nuclei in [42], and later, the K-means clustering algorithm is utilized to split the extracted descriptors, which are then employed as the dictionary's atoms. Rezaeilouyeh in [50] used textural characteristics like shearlets for cancer tissue classification using convolutional neural network (CNN) classifier. For nuclear grading, Lu et al. [30] retrieved a set of morphological and textural parameters like size of nuclei, mean, standard deviation, sum, and entropy of image features. The histogram of each of these attributes is used for grading using an SVM classifier. [43] uses features to reflect the size differences between mitotic nuclei and normal nuclei. The feature set is made up of the mean and standard deviation of ten parameters like contrast, skewness, solidity, grey value, eccentricity, entropy, diameter, area, smoothness, and symmetry retrieved from the segmented nuclei. Gandomkar et al. [22] used texture-based feature extraction method to extract attributes from histopathology slides which can identify the grade of the cancer. These features are later analyzed and coupled with the ensemble of trees to evaluate and to calculate the atypia count.

Topological features in a tumour tissue provide information on the structure and spatial arrangement of nuclei. The spatial interdependence of the cells is represented in this way using various forms of graphs, from which the required information for classification is derived. For cancer grading, graph-based criteria are frequently combined with morphological or textural features.

Naik et al. [51], Das et al. [21] extracts many morphological and graphical representations of features using Voronoi diagram (VD), Delaunay triangulation (DT), and minimal spanning tree (MST) for automated breast cancer grading. Apart from that morphological features as well as boundary features derived from nuclear structure are subjected to principal component analysis (PCA) for feature reduction, and classification is carried out later using an SVM classifier. Textural features and graphical representation of features using MST, VD, and DT are used to represent the architecture of nuclei and grade cancer in Doyle et al. [21, 52]. The textural features like standard deviation, minimum to maximum ratio, and average are included in the study along with graph-based features, and these two feature vectors are subjected to spectral clustering (SC) techniques to reduce the number of features, which is then supplied to the SVM classifier. In paper [53], Wan et al. apply Gabor filters, kirsch filters to access pixel level features, object-based features are extracted through MST, DT, VDs, and convolutional neural networks (CNNs) helps to extract semantic level features, and this helps to identify heterogeneity of cancerous tissue. Later, graph embedding technique is used to reduce the dimensionality of the image, and the resultant image is given to SVM classifier [21]. [48] presents a novel image level descriptor based on area covariances for assessing breast full slide images. The image is divided into non-overlapping area and region covariance (RC) which is calculated for each area. To extract features of this area, maximum response 8(MR8) filter banks are used [21]. The extracted RC descriptors form points on Riemannian manifold.

This is integrated to get single image descriptor called geodesic geometric mean of region covariance, and later, geodesic K-nearest neighbour classifier (GKNN) is used for grading the breast cancer [21].

2.1.4 Image Classification

The features retrieved from tumour tissue are required for cancer classification and grading. Classifiers are divided into learning and testing phase. The characteristics retrieved from digital annotated slides are used to train the classifier during the learning phase [21]. These classifiers are then put to the test with previously unseen data. Algorithms such as Gaussian mixture models, K-nearest neighbourhood (KNN) algorithm, decision tree, random forest classifier (RF), Bayes classifiers, and supervised learning techniques are used for nuclear atypia detection widely. Table 3 gives details of the few most frequently used machine learning techniques in the domain of nuclear atypia scoring. The most widely used deep learning (DL) techniques can avoid explicit feature extraction process [21] and organize discriminative information of the data as its in built functionality.

The cancer features derived from histopathology images are used in an automated breast cancer grading system. This could necessitate precise identification and segmentation of biological structures, which is a difficult operation due to the complex structure of histologic data. As a result, computerized intelligent systems for nuclear atypia scoring are in high demand.

2.2 *Learned Feature-Based Algorithms*

Deep neural networks (DNNs) have been widely used to solve a variety of medical image analysis problems like genetic disorder identification, speech recognition, Alzheimer's disease classification, etc. DL methods have outperformed traditional methods in several areas. This does not necessitate the feature extraction phase. The layers are capable of learning an implicit representation of the raw input by themselves. Several deep learning algorithms for nuclear atypia scoring have recently appeared in the literature. In this section, we will go through the existing deep learning techniques used in the domain of nuclear atypia detection. Figure 4 shows the learned feature-based nuclear atypia detection process.

Rezaeilouyeh et al. [50] employed the magnitude and phase shearlet coefficients as secondary information for the neural network to train CNN and classify cancer. For the classification of breast cancer images, Rakhlin et al. [55] employ a gradient boosting method pre-trained CNN on Image Net. Bardou et al. [56] used CNN with a fully connected classifier layer to classify cancer subtypes using handcrafted features [21].

To extract details from nuclei in different scale in paper [57], Araujo et al. applied CNN architecture. In [58], author proposed a deep convolutional activation feature

Table 3 Nuclear atypia scoring using handcrafted feature-based machine learning technique

Paper	Pre-processing technique	Segmentation technique	Feature extraction technique	Classification technique
Cosatto et al. [35]	Adaptive thresholding technique	Active contour using Hough transform	Texture and shape features	SVM
Naik et al. [51]	–	Bayesian classifier	Morphological and graph-based features	SVM
Dalle et al. [29]	–	Region growing	Morphological features	–
Dalle et al. [46]	–	Boundary	Size, shape, and texture features	Gaussian mixture model
Huang et al. [47]	–	Snake-based algorithm	Size and textural features	Bayesian classifier
Veta et al. [21, 31]	Colour normalization	Watershed segmentation	Morphological features	–
Basavanahally et al. [21, 36]	Orthonormal transformation of RGB vectors	Snake-based segmentation	Textural and graph-based features	Random forest
Lu et al. [30]	Stain normalization and colour de-convolution	Region growing	Morphological features	SVM
Maqlin et al. [21, 43]	–	K-means clustering	Mean and standard deviation of textural and morphological features	Artificial neural network
Moncayo et al. [21, 42]	Colour de-convolution	Threshold	Textural features	SVM
Faridi et al. [45]	Unmixing of colour channels	Boundary	Morphological features	SVM
Wan et al. [37]	Stain normalization	Snake-based segmentation	Textural, graph-based, and CNN derived features	SVM-based cascaded ensemble classifiers
Gandomkar et al. [21, 54]	Stain normalization and colour de-convolution	Threshold	Textural features	Multiple regression trees
Khan et al. [21, 48]	–	–	Geodesic mean of region covariance descriptors	Geodesic K-nearest neighbour classifier(GKNN)

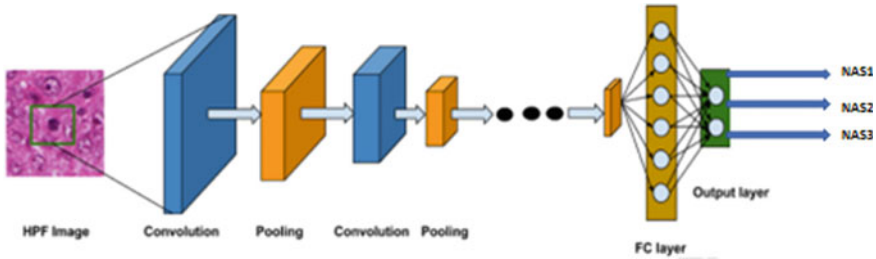


Fig. 4 Learned feature-based nuclear atypia detection process

(DCAF) for classification of histopathological images. In this, a pre-trained CNN for feature extraction is used, and later, it is fed into a new classifier that has been trained on problem specific data. Nahid et al. [59] categorize histopathological images using long short-term memory (LSTM) and an integrated CNN + LSTM model, and a softmax and SVM layers employ the extracted global and local information to determine the class. Guo et al. [60] used a CNN architecture that combines patch level voting, merging module and reduces generalization of error using hierarchy voting approach and bagging technique [21].

Golatkar et al. [61] used Inception-V3 CNN technique on patches extracted based on nuclei density for histological breast cancer classification as the initial step and in the later stage of the process, majority voting technique was used for the final classification [21]. For grading breast cancer, Jannesari et al. [62] employed pre-trained ResNet, ImageNet, and inception model networks. For the multiclass histopathological image classification, Jiang et al. [63] integrated the squeeze and excitation patch and the ResNet model, which reduced the number of parameters and thereby avoiding the problem of model over fitting [21]. In [33] in order to adaptively identify the sample, the author has proposed batch mode active learning on the Riemannian manifold for breast cancer nuclear atypia scoring [21, 64]. In [65], the author proposed sparse coding and dictionary learning on symmetric positive definite (SPD) matrices. The sparse coding problem on the SPD manifold is a convex problem in the higher dimensional reproducing kernel Hilbert space (RKHS), which allows for greater use of the reparability of histopathological images [21, 65].

Recently, generative adversarial networks (GANs) are recently widely employed in biomedical area of research [26] such as medical image reconstruction, image generation, segmentation, classification, and histopathological image analysis since its beginnings [66]. GAN can be used for artificially staining digital pathology images which helps to avoid the problems created through manual staining and thereby reduce the effort and expense of staining process [66]. GAN models are also used to eliminate unknowing and intentional discolorations in stained histopathology pictures that could impair analysis using deep learning algorithm. On breast histology, Xu et al. [67] employed a multi-resolution convolutional network with plurality voting approach for nuclear atypia detection [66].

Conditional GANs [68] are a type of GAN conditioned on labelled images whilst attempting to match the target data directly. This necessitates precisely labelled and registered dataset for training, which might be costly and laborious in histopathological dataset [66]. A coupled GAN (CoGAN) [69] is a GAN version that share weights of layers and combines two GANs to learn a joint distribution from only the residuals [66]. Disadvantage of this GAN model is mode collapse and the gradient growth and instabilities during the process of training. GAN variant [53] recommends that the Wasserstein distance can be used to improve the model. The convergence of these methodologies is still uncertain, its implications are still unknown, and the level of computational complexity is likewise quite high [66].

In [66], the author proposed NAS-SGAN to overcome the shortage of labelled images. It is a stacked feature matching semi-supervised generative adversarial network (GAN) technique used for breast cancer grading. The author used the ability of semi-supervised GAN (SGAN) for representing the distribution of data by utilizing the unlabelled and labelled samples [66]. A combined training of the discriminator with the use of feature matching and layered architecture also makes the model much more stable than other GAN models. Deep learning methods for nuclear atypia scoring are summarized in Table 4

2.3 *Evaluation Metrics*

The evaluation metrics are used to assess the trained classifier model's generalization ability as well as a model evaluation assessor. They are used to evaluate and summarize the efficiency of the classifier model when it is put to the test with data that has not been seen before. For quantitatively summarizing the outcomes, the majority of the approaches we looked at used confusion metrics as the validation metrics.

Performance and quality of classification model are interpreted with confusion matrix table. The true negative (TN), true positive (TP), false positive (FP), and false negative values of the confusion metrics are used to calculate the following metrics like accuracy (Acc), precision (Pr), specificity (TNR), sensitivity or recall (Re), and F1 score (FS), Matthew's correlation coefficient (MCC), false negative rate (FNR), true negative rate (TNR), true positive rate (TPR), false positive rate (FPR), and negative predictive value (NPV) to evaluate the performance of the model classifier.

Receiver operating characteristics (ROC) graphs are an excellent way for analyzing classifiers and visualizing their performance. The relative trade-off between the true positive rate (TPR) and the false positive rate (FPR) at various thresholds is depicted by a ROC curve. The area under the curve (AUC) can be measured to see how well the extracted features can distinguish between different categories of breast cancer. The system is better if the area beneath the curves is larger. Table 5 shows the most commonly used validation metrics in the nuclear atypia detection papers which discussed in the literature review.

Table 4 Nuclear atypia scoring using deep learning techniques

Paper	Classification technique used	Pros of the technique	Cons of the technique
Araujo et al. [57]	CNN	Extract details from nuclei in different scale	Imbalances in the class is not considered
Rakhlin et al. [55]	CNN	Transfer learning using ImageNet	No pre-processing
Bardou et al. [56]	CNN	Combination of handcrafted and learned features	Multi-class classification is not efficient
Spanhol et al. [58]	DL	Pre-trained deep convolutional activation feature	Complicated structures in the image will pull down the accuracy
Nahid et al. [59]	DL	Integrated CNN and long short-term memory (LSTM) model guided by structural and statistical features	Small dataset
Guo et al. [60]	DL	CNN combined with hierarchy voting and bagging technique which shows good performance	Sensitivity value is low
Golatkar et al. [26]	CNN variants	Pre-trained CNN Inception-v3 is used	Grading accuracy is low
Rezaeilouyeh et al. [50]	DL	Improved accuracy due to DNN fed with shearlet coefficients	Very complex process
Xu et al. [67]	DL	Multi-resolution convolutional network combined with plurality voting	Multiple levels of training required
Jannesari et al. [62]	CNN variants	High sensitivity achieved through combination of residual network and inception network	Requires fine tuning of parameters
Jiang et al. [63]	CNN variants	ResNet with reduced set of parameters	Very complex process

(continued)

Table 4 (continued)

Paper	Classification technique used	Pros of the technique	Cons of the technique
Asha Das et al. [64]	Batch mode active learning(BMAL)	Active learning on Riemannian manifold helped to reduce manual effort in labelling the images	Very complex process
Asha Das et al. [65]	Kernel-based sparse coding and dictionary learning(KSCDL)	A sparse coding and dictionary learning on SPD matrices provide a better discrimination	Very complex process
Gulrajani et al. [53]	WGAN-GP	GAN model which use Wasserstein distance	Very complex computation process
Liu et al. [69]	CoGAN	Coupled GAN which share weights and learn with joint distribution	Gradient expansion, vanishing, and mode collapse aresues whist training
Mirza et al. [68]	CGAN	GAN model which is conditioned on labelled images	Needs perfectly annotated data
Asha Das et al. [66]	CNN variant NAS-SGAN model	A semi-supervised generative adversarial training, which improves accuracy and robustness with limited annotated data	Not using clustering properties for improved classification accuracy

2.4 Performance Analysis of Nuclear Atypia Scoring Algorithms

The histopathological image classification continues to be a difficult task due to the huge storage requirement of the image data. Furthermore, the processing required for the image analysis and classification is huge and can take several hours. As a result, computer analysis of histopathological images frequently necessitates the employment of highly efficient computing technologies such as multi-core processors and graphics processing units (GPUs) that are required to speed up the processing. A comparison of the many strategies detailed in the literature is time-consuming because each method uses its own unique dataset, and the findings are given using different assessment measures. This section discusses a comparative examination of some papers which had worked on breast cancer nuclear atypia scoring [24].

The papers included in the comparative analysis are geodesic-distance-based K-nearest neighbour (GKNN) classifier [48, 66], support vector machine [30], deep neural network shearlet transform [50], multi-resolution convolutional network with plurality voting (MR-CN-PV) model [66, 67], batch mode active learning (BMAL)

Table 5 Validation metrics in the nuclear atypia detection

Metrics	Formula	Description
TP		Count of nuclear atypia cells classified correctly
TN		Count of non-nuclear atypia cells classified correctly
FP		Count of nuclear atypia classified incorrectly
FN		Count of non-nuclear atypia classified incorrectly
Accuracy (Ac)	$\frac{TP+TN}{TP+TN+FN+FP}$	It calculates the proportion of right classifications to the total number of test data analyzed
Precision (Pr)	$\frac{TP}{TP+FP}$	Count of correctly identified positives
F1-score (Fs)	$\frac{2*Precision*Recall}{Precision+Recall}$	Harmonic mean of sensitivity gives F1-score
Specificity	$\frac{TN}{TN+FP}$	Count of correctly identified negatives
Sensitivity	$\frac{TP}{TP+FN}$	Calculates the count of correctly identified positives
Error rate	$\frac{FP+FN}{TP+TN+FP+FN}$	The ratio of inaccurate classifications to the total number of test data analyzed
MCC	$\frac{(TP*TN-FP*FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$	The observed classifications' correlation coefficient with the projected classifications
NPV	$\frac{TN}{TN+FN}$	The percentage of negatives properly detected out of the total samples in the negative class
FPR	$\frac{FP}{FP+TN}$	The proportion of negative samples that were incorrectly classified as positive to the count of negative samples
FNR	$\frac{FN}{FN+TP}$	The count of positive samples that was incorrectly forecasted as negative to the count of negative samples

[64, 66], kernel-based sparse coding and dictionary learning (KSCDL) [65, 66], Wasserstein GAN-gradient penalty (WGAN-GP) [53], coupled GAN (CoGAN) [69], conditional GAN (CGAN) [66, 68], and nuclear atypia scoring semi-supervised generative adversarial network (NAS-SGAN) [66]. Tables 6 and 7 show analysis of performance of nuclear atypia scoring on the Aperio and Hamamatsu dataset.

The KSCDL technique was chosen for analysis because they investigate histological breast images in a non-Euclidean framework, and its performance was good [65]. In addition, the SVM, DNN-shearlet, GKNN, and MR-CN-PV were also taken into consideration for comparison because of their outstanding results in textural, morphological, deep learning, and transfer learning-based algorithms. Furthermore, batch

Table 6 Performance analysis of nuclear atypia scoring algorithms on Aperio dataset [21]

Paper	Algorithm	Acc	Re	TNR	Pr	FS	MCC
Khan et al. [48]	GKNN	0.797	0.562	0.782	0.631	0.565	0.393
Lu et al. [30]	SVM	0.780	0.455	0.771	0.445	0.447	0.262
Rezaeilouyeh et al. [50]	DNN-shearlet	0.720	0.333	0.666	0.250	0.285	0.163
Xu et al. [67]	MR-CN-PV	0.800	0.325	0.672	0.330	0.317	0.030
Asha et al. [64]	BMAL	0.853	0.785	0.874	0.762	0.790	0.665
Asha et al. [65]	KSCDL	0.826	0.762	0.855	0.721	0.741	0.621
Gulrajani et al. [53]	WGAN-GP	0.875	0.813	0.902	0.786	0.813	0.701
Liu et al. [69]	CoGAN	0.873	0.811	0.900	0.785	0.812	0.698
Mirza et al. [68]	CGAN	0.857	0.792	0.880	0.767	0.790	0.667
Asha et al. [66]	NAS-SGAN	0.982	0.964	0.974	0.965	0.963	0.940

Table 7 Performance analysis of nuclear atypia scoring algorithms on Hamamatsu dataset [21]

Paper	Algorithm	Acc	Re	TNR	Pr	FS	MCC
Khan et al. [48]	GKNN	0.812	0.571	0.779	0.627	0.569	0.395
Lu et al. [30]	SVM	0.756	0.406	0.726	0.423	0.402	0.181
Rezaeilouyeh et al. [50]	DNN-shearlet	0.747	0.333	0.666	0.249	0.285	0.173
Xu et al. [67]	MR-CN-PV	0.702	0.476	0.726	0.520	0.492	0.224
Asha et al. [64]	BMAL	0.864	0.797	0.880	0.769	0.781	0.653
Asha et al. [65]	KSCDL	0.829	0.761	0.844	0.722	0.742	0.621
Gulrajani et al. [53]	WGAN-GP	0.885	0.815	0.904	0.795	0.814	0.677
Liu et al. [69]	CoGAN	0.868	0.812	0.901	0.793	0.813	0.674
Mirza et al. [68]	CGAN	0.868	0.799	0.881	0.770	0.782	0.654
Asha et al. [66]	NAS-SGAN	0.984	0.975	0.982	0.975	0.974	0.956

mode active learning on Riemannian manifold (BMALR) approach was considered for the comparison due to the active learning used in it [64, 66]. The method was used to reduce the requirement for a labelled dataset. The problem of getting precisely labelled and registered data which is often costly and laborious in the case of histopathology images is resolved to an extent by the use of by the generative adversarial network [66]. Three papers which use GAN concept were chosen for comparative analysis due to exceptional accuracy level when compared to their techniques. Out of the three techniques, complexity and issue of convergence were the disadvantages of CoGAN and CGAN when compared to NAS-SGAN. Amongst all the three techniques which use GAN concept, the NAS-SGAN which use semi-supervised GAN concept was able to resolve the stability issue to an extent and SGANs using its stacked feature matching it was able to collect the data distribution more effectively using both labelled and unlabelled data [66]. Figures 5 and 6

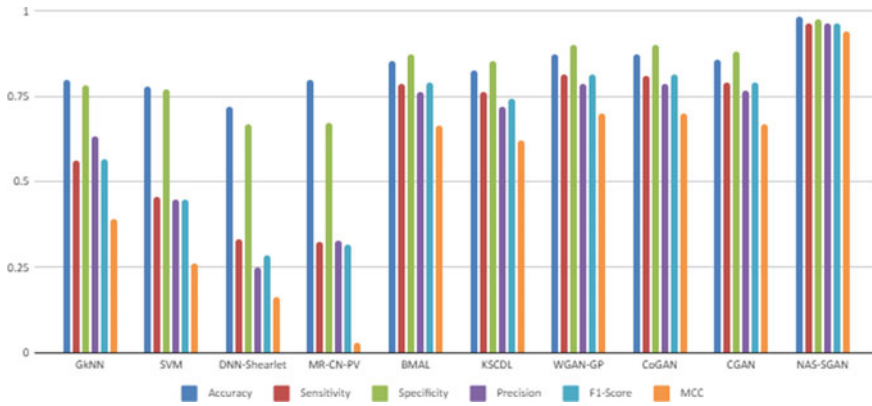


Fig. 5 Performance analysis of nuclear atypia scoring algorithms on Aperio dataset

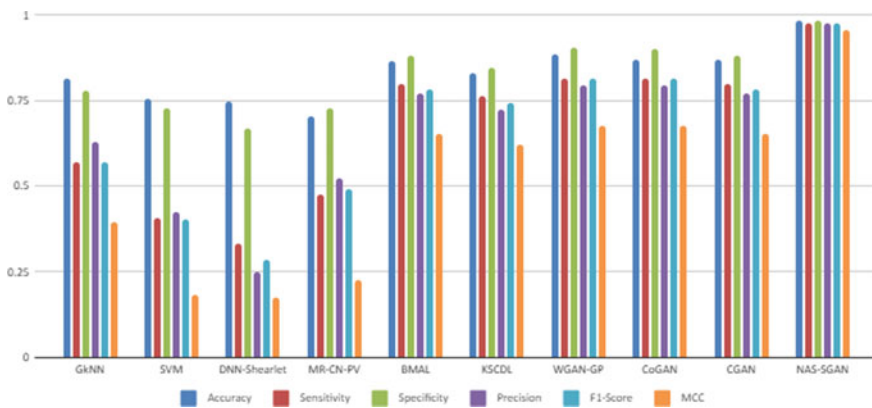


Fig. 6 Performance analysis of nuclear atypia scoring algorithms on Hamamatsu dataset

show the performance analysis of nuclear atypia scoring algorithms on Aperio and Hamamatsu dataset [21]

2.5 Conclusion

The significant challenges of nuclear atypia detection in histopathological breast tissue image analysis are distortions of image caused by inaccurate slide preparation, complicated structure of the underlying biological pattern, lack of stability as well as inter- and intra-observer variation, which can affect diagnostic accuracy and therapy planning. These challenges can be overcome with the help of computer aided analytical techniques. This study reviews the methodologies used for nuclear

atypia detection process in breast cancer grading, obstacles in existing computerized automated grading process, prospective techniques of computer-assisted grading for the disease diagnosis. This research is an attempt to synthesize information to learn about the most recent changes in analyzing characteristics for breast cancer grading and provide a summary of the precision and reliability of various methods in the domain of nuclear atypia detection. The findings shows that deep learning has exploded in popularity in the field of nuclear atypia detection as a result of its superior results. Approaches based on deep learning offer higher prospects in the future. In deep learning technique, generative adversarial network (GAN) outperforms non-adversarial and deep learning networks because adversarial training uses the discriminator network to train the generator network, resulting in improved model performance. GAN also provides provision for automatic staining, and it performs well even without perfectly annotated dataset. We also found that public datasets and open challenges have a good impact on research on histopathological image analysis, and most of the studies in the literature used public dataset. In the future variations, of GAN technique can be developed with acceptable performance levels for clinical applications. This study could serve as inspiration for future research.

References

1. WHO—cancer key facts; 2018. <https://www.who.int/news-room/fact-sheets/detail/cancer>. Accessed 20 Nov 2020
2. Mathew T et al (2020) Computational methods for automated mitosis detection in histopathology images: a review. *Biocybern Biomed Eng*. <https://doi.org/10.1016/j.bbe.2020.11.005>
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin* 68(6):394–424
4. Wardle J, Robb K, Vernon S, Waller J (2015) Screening for prevention and early diagnosis of cancer. *Am Psychol* 70(2):119
5. Mahmood T, Li J, Pei Y, Akhtar F, Imran A, Rehman KU (2020) A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. *IEEE Access* 8:165779–165809. <https://doi.org/10.1109/ACCESS.2020.3021343>
6. Bejnordi BE, Zuidhof G, Balkenhol M, Hermsen M, Bult P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak J (2017) Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J Med Imag* 4(4):1–8
7. Al-Janabi S, Huisman A, Van Diest PJ (2012) Digital pathology: current status and future perspectives. *Histopathology* 61(1):1–9
8. Higgins C (2015) Applications and challenges of digital pathology and whole slide imaging. *Biotechn Histochem* 90(5):341–347
9. Manoharan JS (2021) Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm. *J Soft Comput Paradigm (JSCP)* 3(02):83–95
10. Al-Janabi S, Huisman A, Vink A, Leguit RJ, Offerhaus GJA, Ten Kate FJ et al (2012) Whole slide images for primary diagnostics of gastrointestinal tract pathology: a feasibility study. *Human Pathol* 43(5):702–707

11. Shibusawa M, Nakayama R, Okanami Y, Kashikura Y, Imai N, Nakamura T, Kimura H, Yamashita M, Hanamura N, Ogawa T (2016) The usefulness of a computer-aided diagnosis scheme for improving the performance of clinicians to diagnose non-mass lesions on breast ultrasonographic images. *J Med Ultrason* 43(3):387–394
12. Kumar R, Srivastava R, Srivastava S (2015) Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. *J Med Eng* 2015(2015):1–14. <https://doi.org/10.1155/2015/457906>
13. Wan T et al (2014) Wavelet-based statistical features for distinguishing mitotic and non-mitotic cells in breast cancer histopathology. In: 2014 IEEE international conference on image processing (ICIP) 2290–2294
14. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI et al (2002) Revision of the American joint committee on cancer staging system for breast cancer. *J Clin Oncol* 20(17):3628–3636
15. O’Sullivan B, Brierley J, Byrd D, Bosman F, Kehoe S, Kossary C, Piñeros M, Van Eycken E, Weir HK, Gospodarowicz M (2017) The TNM classification of malignant tumours—towards common understanding and reasonable expectations. *Lancet Oncol* 18(7):849–851. [https://doi.org/10.1016/S1470-2045\(17\)30438-2](https://doi.org/10.1016/S1470-2045(17)30438-2). PMID: 28677562; PMCID: PMC5851445.
16. Elston CW, Ellis IO (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 41(3a). 151. Elston CW, Ellis IO (2002) Author commentary. *Histopathology* 19:403–410
17. Fuchs TJ, Buhmann JM (2011) Computational pathology: challenges and promises for tissue analysis. *Comput Med Imaging Graph* 35(7–8):515–530
18. Hamdan YB (2021) Construction of statistical SVM based recognition model for handwritten character recognition. *J Inf Technol* 3(02):92–107
19. Paul A, Mukherjee DP (2015) Mitosis detection for invasive breast cancer grading in histopathological images. *IEEE Trans Image Process* 24(11):4041–4054
20. Raj JS, Joe MCV (2021) Wi-Fi network profiling and QoS assessment for real time video streaming. *IRO J Sustain Wirel Syst* 3(1):21–30
21. Das A, Nair MS, Peter SD (2020) Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: a review. *J Digit Imaging* 33(5):1091–1121. <https://doi.org/10.1007/s10278-019-00295-z>. PMID: 31989390; PMCID: PMC7573034
22. MITOS – dataset; 2012. http://ludo17.free.fr/mitos_2012/dataset.html. Accessed 20 Nov 2020
23. Veta M, Van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A et al (2015) Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal* 20(1):237–248
24. MITOS-ATYPIA—dataset; 2014. <https://mitos-atypia-14.grand-challenge.org/Dataset/>. Accessed 20 Nov 2020
25. TUPAC – dataset; 2016. <http://tupac.tue-image.nl/node/3>. Accessed 20 Nov 2020
26. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. *IEEE Trans Biomed Eng* 63(7):1455–1462. <https://doi.org/10.1109/TBME.2015.2496264>
27. Aksac A, Demetrick DJ, Ozyer T et al (2019) BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis. *BMC Res Notes* 12:82. <https://doi.org/10.1186/s13104-019-4121-7>
28. Beevi KS, Nair MS, Bindu GR (2016) Automatic segmentation of cell nuclei using Krill Herd optimization based multithresholding and localized active contour model. *Biocybernetics Biomed Eng* 36(4):584–596. <https://doi.org/10.1016/j.bbe.2016.06.005>
29. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC (2008) Automatic breast cancer grading of histopathological images. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 3052–3055. <https://doi.org/10.1109/IEMBS.2008.4649847>

30. Lu C, Ji M, Ma Z, Mandal M (2015) Automated image analysis of nuclear atypia in high-power field histopathological image. *J Microsc* 258(3):233–240. <https://doi.org/10.1111/jmi.12237>
31. Veta M, Kornegoor R, Huisman A, Verschuur-Maes AHJ, Viergever MA, Pluim JPW, van Diest PJ (2012) Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol* 25(12):1559–1565. <https://doi.org/10.1038/modpathol.2012.126>
32. Man R, Yang P, Xu B (2020) Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. *IEEE Access* 8:155362–155377. <https://doi.org/10.1109/ACCESS.2020.3019327>
33. Niethammer M, Borland D, Marron J, Woosley JT, Thomas NE (2010) Appearance normalization of histology slides. In: *MLMI*, Springer, Berlin, pp 58–66
34. Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, Thomas NE (2009) A method for normalizing histology slides for quantitative analysis. In: *Proceedings—2009 IEEE international symposium on biomedical imaging: from nano to macro*. ISBI 2009, pp 1107–1110. <https://doi.org/10.1109/ISBI.2009.5193250>
35. Cosatto E, Miller M, Graf HP, Meyer JS (2008) Grading nuclear pleomorphism on histological micrographs. 2008 ICPR 2008 19th international conference on (August 2016) *Pattern Recognition*, pp 1–4. <https://doi.org/10.1109/ICPR.2008.4761112>
36. Basavanahally A, Ganesan S, Feldman M, Shih N, Mies C, Tomaszewski J, Madabhushi A (2013) Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. *IEEE Trans Biomed Eng* 60(8):2089–2099. <https://doi.org/10.1109/TBME.2013.2245129>
37. Wan T, Cao J, Chen J, Qin Z (2017) Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features. *Neurocomputing* 229:34–44. <https://doi.org/10.1016/j.neucom.2016.05.084>
38. Xing F, Yang L (2016) Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a Comprehensive review. *IEEE Rev Biomed Eng* 9:234–263. <https://doi.org/10.1109/RBME.2016.2515127>
39. Petushi S, Katsinis C, Coward C, Garcia F, Tozeren A (2004) Automated identification of microstructures on histology slides. In: *2004 IEEE international symposium on biomedical imaging: nano to macro*, IEEE, pp 424–427
40. Petushi S, Garcia FU, Haber MM, Katsinis C, Tozeren A (2006) Large-scale computations on histology images reveal gradedifferentiating parameters for breast cancer. *BMC Med Imaging* 6(1):14. <https://doi.org/10.1186/1471.2342.6.14>
41. Weyn B, Van De Wouwer G, Van Daele A, Scheunders P, Van Dyck D, Van Marck E, Jacob W (1998) Automated breast tumor diagnosis and grading based on wavelet chromatin texture description. *Cytometry* 33(1):32–40. [https://doi.org/10.1002/\(SICI\)1097-0320\(19980901\)33:1.32:AID-CYTO4.3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0320(19980901)33:1.32:AID-CYTO4.3.0.CO;2-D)
42. Moncayo R, Romo-Bucheli D, Romero E (2015) A grading strategy for nuclear pleomorphism in histopathological breast cancer images using a bag of features (bof). *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* 9423:75–82. <https://doi.org/10.1007/978.3.319.25751.8.10>
43. Maqlin P, Thamburaj R, Mammen JJ, Manipadam MT (2015) Automated nuclear pleomorphism scoring in breast cancer histopathology images using deep neural networks. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* 9468:269–276. <https://doi.org/10.1007/978.3.319.26832.3.26>
44. Xue M, Shivakumara P, Zhang C et al (2019) Curved text detection in blurred/non-blurred video/scene images. *Multimed Tools Appl* 78:25629–25653. <https://doi.org/10.1007/s11042-019-7721-2>
45. Faridi P, Danyali H, Helfroush MS, Jahromi MA (2016) Cancerous nuclei detection and scoring in breast cancer histopathological images. *arXiv:161201237*
46. Dalle Jr, Racoceanu D, Putti TC (2009) Nuclear pleomorphism scoring by selective cell nuclei detection. In: *IEEE workshop on applications of computer vision: 7–8*

47. Huang CH, Veillard A, Roux L, Lomenie N, Racoceanu D (2011) Time-efficient sparse analysis of histopathological whole slide images. *Comput Med Imaging Graph* 35(7–8):579–591. <https://doi.org/10.1016/j.compmedimag.2010.11.009>
48. Khan AM, Sirinukunwattana K, Rajpoot N (2015) A global covariance descriptor for nuclear atypia scoring in breast histopathology images. *IEEE J Biomed Health Inform* 19(5):1637–1647. <https://doi.org/10.1109/JBHI.2015.2447008>
49. Ojansivu V, Linder N, Rahtu E, Pietikainen M, Lundin M, Joensuu H, Lundin J (2013) Automated classification of breast cancer morphology in histopathological images. *Diagn Pathol* 8(1):S29
50. Rezaeilouyeh H, Mollahosseini A, Mohammad MH (2016) Microscopic medical image classification framework via deep learning and shearlet transform. *J Med Imaging* 3(4):044,501. <https://doi.org/10.1117/1.JMI.3.4.044501>
51. Naik S, Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J (2008) Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: 2008 5th IEEE international symposium on biomedical imaging: from nano to macro, proceedings, ISBI, pp 284–287. <https://doi.org/10.1109/ISBI.2008.4540988>
52. Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J (2008) Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In: 2008 5th IEEE international symposium on biomedical imaging: from nano to macro, proceedings, ISBI, pp 496–499. <https://doi.org/10.1109/ISBI.2008.4541041>
53. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A (2017) Improved training of Wasserstein GAN's. arXiv preprint [arXiv:1704.00028](https://arxiv.org/abs/1704.00028), 2017
54. Gandomkar Z, Brennan PC, Mello-Thoms C (2019) Computer-assisted nuclear atypia scoring of breast cancer: a preliminary study. *J Digit Imaging*. <https://doi.org/10.1007/s10278-019-00181-8>
55. Rakhlin A, Shvets A, Igllovikov V, Kalinin AA (2018) Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol 10882 LNCS, pp 737–744. <https://doi.org/10.1007/978-3-319-93000-8-83>, 1802.00752.
56. Bardou D, Zhang K, Ahmad SM (2018) Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access* 6:24,680–24,693
57. Ara'ujo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, Polonia A, Campilho A Classification of breast cancer histology images using convolutional neural networks. *PloS one* 12(6), 2017.
58. Spanhol FA, Oliveira LS, Cavalin PR, Petitjean C, Heutte L (2017) Deep features for breast cancer histopathological image classification. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp 1868–1873
59. Nahid AA, Mehrabi MA, Kong Y (2018) Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *BioMed research international* 2018.
60. Guo Y, Dong H, Song F, Zhu C, Liu J (2018) Breast Cancer Histology Image Classification Based on Deep Neural Networks. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol 10882 LNCS, pp 827–836. <https://doi.org/10.1007/978-3-319-93000-8-94>, 1803.04054
61. Golatkar A, Anand D, Sethi A (2018) Classification of breast cancer histology using deep learning. In: *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, vol 10882 LNCS, pp 837–844. https://doi.org/10.1007/978-3-319-93000-8_95
62. Jannesari M, Habibzadeh M, Aboulkheyr H, Khosravi P, Elemento O, Totonchi M, Hajira-souliha I (2019) Breast cancer histopathological image classification: a deep learning approach. In: *Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018*, pp 2405–2412. <https://doi.org/10.1109/BIBM.2018.8621307>.
63. Jiang Y, Chen L, Zhang H, Xiao X (2019) Breast cancer histopathological image classification using convolutional neural networks with small SE-resnet module. *PLoS ONE* 14(3). <https://doi.org/10.1371/journal.pone.0214587>.

64. Das A, Nair MS, Peter DS (2020) Batch mode active learning on the Riemannian manifold for automated scoring of nuclear pleomorphism in breast cancer. *Artif Intell Med* 103:101805. ISSN 0933-3657. <https://doi.org/10.1016/j.artmed.2020.101805>
65. Das A, Nair MS, Peter SD (2019) Sparse representation over learned dictionaries on the Riemannian manifold for automated grading of nuclear pleomorphism in breast cancer. *IEEE Trans Image Process* 28(3):1248–1260. <https://doi.org/10.1109/TIP.2018.2877337>
66. Das A, Devarampati VK, Nair MS (2021 Nov 26) NAS-SGAN: a semi-supervised generative adversarial network model for atypia scoring of breast cancer histopathological images. *IEEE J Biomed Health Inform.* <https://doi.org/10.1109/JBHI.2021.3131103>. Epub ahead of print. PMID: 34826299
67. On breast histology, Xu et al. [76] describe a deep learning–based approach for computerized Nuclear Atypia Grading (NAS). The computerized scoring is done via a Multi-Resolution Convolutional Network with Plurality Voting
68. Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv preprint [arXiv:1411.1784](https://arxiv.org/abs/1411.1784)
69. Liu M, Tuzel O (2016) Coupled generative adversarial networks. CoRR abs/1606.07536. [Online]. Available <http://arxiv.org/abs/1606.07536>

Security Tradeoff in Network Virtualization and Their Countermeasures



Nayeem Ahmad Khan and Mohammed Al Qurashi

Abstract To minimize the ever-growing infrastructure within data centers and other critical concerns such as availability and scalability, network virtualization technology has gained favor not just among the technology profession but also among infrastructure managers. Network virtualization is the abstraction of network resources from their conventional delivery in hardware to software. Numerous physical networks may be combined into a single virtual, software-based network using network virtualization, or a single physical network can be divided into multiple distinct and independent virtual networks using network virtualization. Network virtualization is considered a lifeline for cloud computing for distributed services. Network virtualization is widely used to increase the efficiency of computing services. Like other technologies, network virtualization encounters significant security issues that make it vulnerable to attacks such as distributed denial of service (DDoS) or denial of service (DoS). Therefore, it is critical to formulating an effective solution to address these security threats. This study proposes an effective solution to minimize the security threats to network virtualization. The proposed solution is based on the use of Confidence-Based Filtering, Hop Count Filtering, and Cloud Trace.

Keywords Virtualization · Virtual machines · Attacks · Distributed denial of service · Confidence-based filtering · Hop count filtering · Cloud trace back

1 Introduction

The demand for network function virtualization and network virtualization is increasing with the trend of transforming new IP infrastructures on today's networks.

N. A. Khan (✉) · M. Al Qurashi
Faculty of Computer Science and Information Technology, AlBaha University, AlBaha, Saudi Arabia
e-mail: nayeem@bu.edu.sa

M. Al Qurashi
e-mail: malqurashi@bu.edu.sa

Network virtualization allows the creation of a virtual network over a single physical network, while network function virtualization enables virtual machines to perform the network functions in a cloud infrastructure [1]. Security challenges to both techniques make them vulnerable to DoS and DDoS threats. A DoS attack attempts to make a computer or network resource unavailable to its intended users [2]. Network function virtualization is a technique that enables the creation of virtual network functions in the cloud infrastructure. The virtual network function can be deployed on the cloud infrastructure and can be used to provide network functions such as firewalls, load balancers, and VPNs. The virtual network functions can be used to create a virtual network that is indistinguishable from a physical network. Virtual network functions can also be deployed on a private cloud infrastructure that is owned and operated by an enterprise. Virtualized network functions can provide a number of advantages over physical network functions. Virtualized network functions can be deployed more flexibly than physical network functions and can be scaled up or down in response to changing network requirements. Virtualized network services can also be moved from one location to another, allowing network functions to be deployed in an on-demand manner and provisioned on an as-needed basis. This contrasts with traditional network services, which are typically provisioned at the time of deployment [3].

Network virtualization encounters common security threats include disclosure, deception, disruption, and usurpation [4]. Disclosure occurs when private information leaks through the physical resources that virtual networks share. Deceptions include identity fraud, loss of registry entries, and reply attacks. Disruption happens due to the mismanagement of resources and physical device failures. Usurpation attacks allow attackers to access sensitive information on virtual routers. NFV, on the other hand, faces network function-specific security problems that refer to attacks occurring directly on the network resources [5]. The private use of NFV can narrow down the attackers to insiders who have access to the system, but public use of NFV poses greater threats as attackers can use public or third-party network attach like DDoS and DoS to access the system.

DoS is a common threat to any organization. It occurs when attackers flood the server with fraudulent contents, causing it to crash. On the other hand, DDoS is more dangerous than DoS, as attacks on the system come from various locations and through a different Internet connections [6]. Both can cause temporary or permanent crashes on the server, making it unavailable to use. Recently, DoS and DDoS are becoming massive threats to virtualization, causing its system and network to crash and thus become unavailable to users. Some defense mechanisms have been constructed to address these threats, such as Intrusion Detection System (IDS), Hybrid Intrusion Detection System (H-IDS), Confidence-Based Filtering (CBF), Hop Count Filtering (HOC), and Cloud Trace Back (CTB) [7]. Despite their effectiveness, each mechanism has weaknesses that still allow attackers to enter the network. Therefore, this study proposes a combination of CBF, H-IDS, and CTB approaches to provide better security for virtualization. The purpose of combining these mechanisms is to cover each other's loopholes through their strengths, thus, providing full-coverage protection for the network.

2 Existing Solutions

2.1 Intrusion Detection System (IDS)

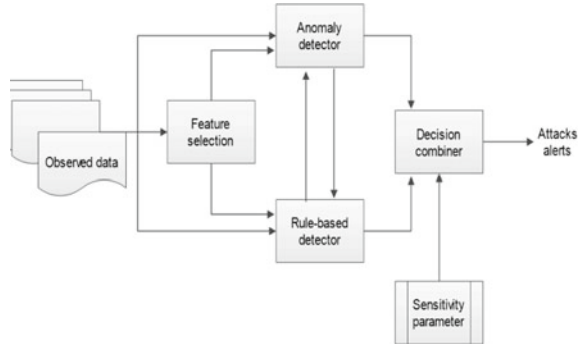
The IDS detects DoS attacks present on the virtual machine, both through the internal and external machines. Real-time IDS involves the sniffer, feature extraction process, and a machine learning classifier such as One-Class Support Vector Machine (SVM) [8]. The feature extraction process is used to extract the features from the packets. The machine learning classifiers are used to classify the packets into normal and DoS attacks [9]. Then, the feature extraction script takes the input from the captured IP header fields and selected traffic features to calculate the receiving IP address. The traffic flow is then distributed into the UDP, ICMP, TCP, and others based on their protocol field. The classifier applies these features to determine the abnormal and normal data instances. Through the transformation method used by One-Class SVM, the decision boundary within a higher dimension space is formed [10]. An alert system then sends a notification to the cloud administrator for the detected DoS attack.

2.2 Hybrid Intrusion Detection System (H-IDS)

The H-IDS improves the detection of DDoS attacks by shortening detection delay and increasing detection accuracy.

As shown in Fig. 1 [11], a DDoS attack is processed to obtain features of the normal traffic in the observed data. Signature-based and anomaly-based detection blocks are linked to the data in order to identify attacks. The decision combiner examines the detector's output together with the sensitivity parameter before an alarm notifies the administrator about the attack. The H-IDS consists of feature extraction and activity model computation, an anomaly-based detector, a signature-based detector, and a hybrid detection engine (HDE) [12]. In the feature extraction stage, the H-IDS extracts feature from data streams. The activity model computation stage computes the activity model for each user. First, the H-IDS forms an activity model and removes some features such as packet interarrival time. Then, the anomaly-based detector, which differentiates abnormal from normal traffic, and the signature-based detector, which works with predefined signature sets to enable the alarm upon detection, are implemented after feature extraction. Lastly, the HDE in the decision combiner uses the combined output from both detectors to improve detection performance before it notifies the administrator.

Fig. 1 H-IDS model



2.3 Confidence-Based Filtering (CBF)

CBF works on both the attack and non-attack periods with the aim of preventing the occurrence of DDoS at the network and transport layer [13]. During the non-attack period, CBF secures the collection of legitimate packets and evaluates them to create a normal profile from the attribute pair in their TCP and IP headers. The computed confidence value shows the patterns in the attribute pair; a higher occurrence means a higher confidence value. In the attack period, computed CBF scores are compared with the standard value to determine the legitimacy of the packet; when the CBF score is higher than the threshold value, the packet is legitimate. This mechanism promotes high computational efficiency and speed, which is ideal for large network traffic [14].

2.4 Hop Count Filtering (HCF)

The main purpose of HCF is to spoof and classify the legitimate packet. HCF can remove almost all spoofed IP packets from a traffic attack; thus, protecting the server from further damage. The HCF makes it difficult for attackers to forge the information header of IP addresses and determine between spoofed and legitimate packets [15]. This mechanism obtains packets from the attacked traffic at the Window Matrix through the genetic algorithm. The GA reduces the overhead computational for HCF since this solution still needs to calculate the Hop Counting Information per packet, resulting in increased computational overhead [16].

2.5 Cloud Trace Back (CTB)

CTB traces the source of all attacks and is commonly used to determine the source of DDoS attacks accurately. Figure 2 [17] shows how CTB is placed before the

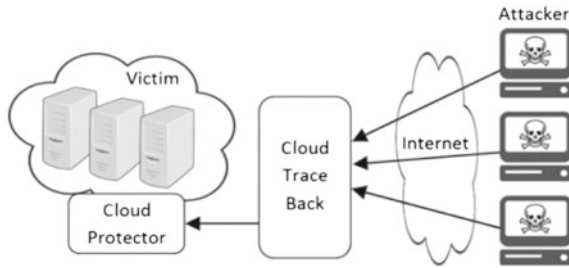


Fig. 2 CTB model

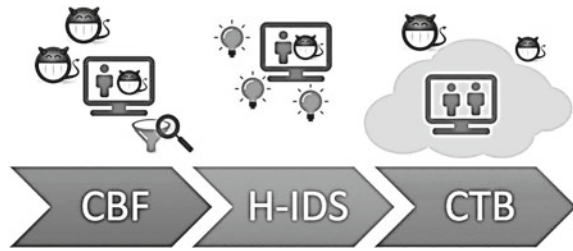
web server, making the Cloud Trace Back Mark its header. Through this, all service requests will be marked in CTB, which will remove the service provider address and hinder direct attacks on the requests. When an attack occurs, the user will send a request message based on the service description to the CTB, and the mechanism will be placed on the header. The service will then be sent to the web server for reconstruction. If the message is found abnormal, a notification will be given to the request owner for further messages and processing requests [18].

3 Proposed Solution

The proposed solution is to combine three solutions presented in Sect. 2: CBF, H-IDS, and CTB. The role of the CBF is to prevent DDoS/DoS attacks at the network and transport layer. It is used as the first line of defense because its computational time is shorter than H-IDS. After CBF detects abnormal data packets, the H-IDS will follow; if not, the system will end. To provide improved performance at the CBF level, a more efficient hash algorithm is necessary to help in boosting the speed and accuracy of CBF’s filtering method. The weakness of CBF that H-IDS covers is the low accuracy of the solution for detection. Through the anomaly-based detector, the H-IDS can accurately detect unidentified attacks. H-IDS also performs machine language techniques using the genetic algorithm and artificial neural network, enhancing the accuracy of detecting the system’s attackers and reducing false alarms. However, the solution also needs to remove the feature extraction since this will already be the task of the CBF layer.

H-IDS’s weakness is that it often misclassifies minority classes of attacks. The CTB covers this weakness as it can detect most of the sources of attack messages accurately within a few seconds. CTB also addresses the weakness of H-IDS in detecting attacks only at the network layer DDoS and not on the application layer. CTB detects an attack on the application layer DDoS while CBF detects those from the network layer DDoS. The CTB’s main role is tracing and identifying the IP address of the attacker after receiving the abnormal network from the H-IDS layer. Figure 3 shows the process of the proposed solution.

Fig. 3 Proposed model



The proposed solution fulfills the following criteria:

- Provides quick response during the attack phase
- Reduces the false alarm rate
- Utilizes anomaly-based approach for detecting unidentified attacks
- Accurately identifies the source of an attack
- Traces back the attacker's IP address.

4 Discussion and Analysis

As noted, the proposed solution is composed of three models that start with CBF, then H-IDS, and lastly, the CTB. The combination of these approaches resulted in an increased level of security for virtualization. A detailed step for the proposed solution is presented below.

Steps for proposed solution:

1. Gather data packets from the system of the receiver
2. Filter and process the data packets
3. Implement anomaly-based and signature-based techniques if data packets are abnormal
4. Remove data packets and end system if data packets are normal
5. Use traceback if both anomaly and signature return as abnormal
6. Use traceback if an anomaly is abnormal and the signature is normal
7. Use traceback if an anomaly is normal and the signature is abnormal
8. Don't implement traceback if both anomaly and signature return as normal
9. The attacker should be found after a successful traceback.

The presented steps are proven and justified through a mathematical solution named resolution. First, we formulated a hypothesis or the expected result from the model, which is finding the attacker. This hypothesis needs to be negated as the first step of the resolution technique. Next, the steps for the proposed solution were transformed into mathematical statements. Then, we applied the rules and regulations of the resolution rule on the mathematical statements. The resolution ends if the result ends up with an empty set; this means that the hypothesis is proven true. The proposed

solution is logically correct and acceptable if the resolution technique results in an empty set.

Solution

Step 1: Contradict the hypothesis

The negated hypothesis is that the attacker is not found.

Step 2: Assign mathematical values on the steps

- A. Gather data packets from the receiver system
- B. Process data packets
- C. Abnormal result
- D. Use anomaly technique
- E. Use signature technique
- F. Anomaly is abnormal
- G. Signature is abnormal
- H. Use trace back
- I. Attacker is found

Step 3: Resolution

The resolution 1 to 9 is gained from the proposed steps 1–9.

1. A
2. B
3. $C \rightarrow D \wedge E$
4. $\neg C \rightarrow \neg D \wedge \neg E \wedge \neg A \wedge \neg B$
5. $F \wedge G \rightarrow H$
6. $F \wedge \neg G \rightarrow H$
7. $\neg F \wedge G \rightarrow H$
8. $\neg F \wedge \neg G \rightarrow \neg H$
9. $H \rightarrow I$
10. $\neg I$ Negation of hypothesis
11. $C \vee (\neg D \wedge \neg E \wedge \neg B)$ Modus Tollens Step 1 & 4
12. $C \vee (\neg D \wedge \neg E)$ Negation using Step 2 & 11
13. \square Negation using Step 3 & 12
14. \square Negation using Step 5 & 8
15. $\neg F \vee G \vee H$ Modus Tollens using Step 6
16. $F \vee \neg G \vee H$ Modus Tollens using Step 7
17. H Negation using Step 15 & 16
18. I Modus Tollens using Step 9 & 17
19. \square Negation using Step 10 & 18

The results of the resolution technique show an empty set which means that the hypothesis is true and that the proposed model is logical and can be accepted as a solution for security threats in virtual environments. After the attacker is found and the traceback method gains their IP address, the packet from this IP address will be blocked, thus, preventing it from entering the virtualized network.

5 Conclusion

Virtualization is a highly demanded technology that constantly faces security threats such as DoS and DDoS. This paper identified and reviewed existing solutions that address these security issues, such as the IDS, H-IDS, CBF, HCF, and CTB. Considering that each of the current solutions has weaknesses, we proposed a model that strengthens virtualization security. The proposed solution combines CBF, H-IDS, and CTB approaches in one system. CBF acts as the first line of defense that filters data packets. H-IDS, then, accurately detects DDoS/DoS attacks. The CTB is then used to trace back the source of the attack and determine its IP address. Through a resolution rule, the proposed model was evaluated, and its result proves that our proposed solution is logically true and can be accepted as an effective model for protecting the virtualized environment from DDoS/DoS attacks.

References

1. Khan N, Alzaharani MY (2018) Establishing EndVisor and quarantine approach in solving security issues of virtualization. *Indian J Sci Technol* 11:47
2. Khan N, War TA (2018) A Deep Study on security vulnerabilities in virtualization at cloud computing. *Int J Comput Appl* 975, 8887
3. Dildar MS, Khan N, Abdullah JB, Khan AS (2017, March) Effective way to defend the hypervisor attacks in cloud computing. In: 2017 2nd international conference on anti-cyber crimes (ICACC). IEEE, pp 154–159
4. Bodeau D, Graubart R, Heinbockel W, Laderman E (2015) Cyber resiliency engineering aid-the updated cyber resiliency engineering framework and guidance on applying cyber resiliency techniques. MITRE CORP BEDFORD MA, Bedford United States
5. Firoozjaei MD, Jeong JP, Ko H, Kim H (2017) Security challenges with network functions virtualization. *Futur Gener Comput Syst* 67:315–324
6. Khan N, Abdullah J, Khan AS (2015, August) Towards vulnerability prevention model for web browser using interceptor approach. In: 2015 9th international conference on IT in Asia (CITA). IEEE, pp 1–5
7. Modi C, Patel D, Borisaniya B, Patel H, Patel A, Rajarajan M (2013) A survey of intrusion detection techniques in cloud. *J Netw Comput Appl* 36(1):42–57
8. Shon T, Moon J (2007) A hybrid machine learning approach to network anomaly detection. *Inf Sci* 177(18):3799–3821
9. Suresh M, Anitha R (2011 July) Evaluating machine learning algorithms for detecting DDoS attacks. In: International conference on network security and applications. Springer, Berlin, Heidelberg, pp 441–452
10. Khan N, Abdullah J, Khan AS (2017) Defending malicious script attacks using machine learning classifiers. *Wirel Commun Mobile Comput*. <https://doi.org/10.1155/2017/5360472>
11. Khan NA (2022) PKI-Based security enhancement for IoT in 5G networks. In: inventive computation and information technologies pp 217–225. Springer, Singapore
12. Nathiya T, Suseendran G (2019) An effective hybrid intrusion detection system for use in security monitoring in the virtual network layer of cloud computing technology. In: Data management, analytics and innovation. Springer, Singapore, pp 483–497
13. Dou W, Chen Q, Chen J (2013) A confidence-based filtering method for DDoS attack defense in cloud environment. *Futur Gener Comput Syst* 29(7):1838–1850

14. Chen Q, Lin W, Dou W, Yu S (2011, December) CBF: a packet filtering method for DDoS attack defense in cloud environment. In: 2011 IEEE ninth international conference on dependable, autonomic and secure computing. IEEE, pp 427–434
15. Wang H, Jin C, Shin KG (2007) Defense against spoofed IP traffic using hop-count filtering. *IEEE/ACM Trans Netw* 15(1):40–53
16. Blanco R, Malagón P, Cilla JJ, Moya JM (2018, July) Multiclass network attack classifier using CNN tuned with genetic algorithms. In: 2018 28th international symposium on power and timing modeling, optimization and simulation (PATMOS). IEEE, pp 177–182
17. Masdari M, Jalali M (2016) A survey and taxonomy of DoS attacks in cloud computing. *Secur Commun Netw* 9(16):3724–3751
18. Joshi B, s AS, Joshi BK (2012, January) Securing cloud computing environment against DDoS attacks. In: 2012 international conference on computer communication and informatics. IEEE, pp 1–5

Current Waste Management Issues in Developing Countries and IoT-Based Solutions in Prospect of Smart City Development



Mohd Anjum, Sana Shahab, Ummatul Fatima, and M. Sarosh Umar

Abstract In the last half century, waste generation has grown multi-fold worldwide, so solid waste management has emerged as a ubiquitous problem. This unmanaged waste affects human health, creates environmental pollution, and is an obstacle to the sustainable development of urban regions. The study has identified issues related to the existing SWM service framework, especially in developing countries. It has briefly discussed the concept of smart and sustainable urban development, i.e., smart cities, to develop an understanding of SWM services in the context of smart city services. It also explains that the SWM services are crucial for creating the smart environment, one of the smart city development vision components. The literature is reviewed in the context of technical opportunities in SWM. The study has uncovered various technologies, namely spatial technologies, identification technologies, data acquisition technologies, data communication technologies, and artificial intelligence, to develop smart solutions for different services related to SWM systems. These identified technologies have a broad range of practical applications in various phases of the SWM service process, from its inception to the final disposal; therefore, they can be effectively used to overcome the identified problems. The research has built a smart bin using an ultrasonic sensor to overcome the issue of real-time waste monitoring. It has also implemented the RFID-based waste collection system and demonstrated the results.

M. Anjum (✉) · M. S. Umar

Department of Computer Engineering, Aligarh Muslim University, Aligarh, India
e-mail: mohdanjum@zhcet.ac.in

M. S. Umar

e-mail: saroshumar@zhcet.ac.in

S. Shahab

Department of Business Administration, College of Business Administration, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
e-mail: sshahab@pnu.edu.sa

U. Fatima

Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh, India

Keywords Information and communication technology · Internet of Things · Municipal solid waste · Smart city · Smart solid waste management system · Solid waste management

1 Introduction

The huge amount of municipal solid waste (MSW) generated has emerged as an attention-seeking issue, especially in developing countries, due to fast urbanization, remarkable socio-economic growth, and unprecedented population growth. Solid waste management (SWM) is a ubiquitous problem impacting the whole world in regard to environmental degradation and human and animal health. According to the world bank statistics [1], 2.01 billion tons (5.5 million tons per day) of MSW are produced annually at the global level and are expected to grow to 3.40 billion tons by 2050. Approximately 33% of total MSW is not managed scientifically to protect the environment and human health. In developing countries, this percentage is significantly higher than 33%. These statistics have also predicted that the volume of global waste generated by 2030 and 2050 will be approximately 2.59 and 3.40 billion tons per annum [1]. Furthermore, the urban areas of developing countries have shown an exponential increase in waste generation. Nowadays, most of these areas have been implementing Internet of Things (IoT)-enabled services to build smart cities.

The fundamental problems related to MSW, especially in developing countries, start at its inception and at the starting point of the SWM service process. The study has identified various issues related to the SWM system in developing countries. One of these major issues is that people do not possess a serious attitude toward waste and a lack knowledge about the consequences of open disposal, so they do not take proper care when throwing waste. Consequently, the waste flood seems around a lot of bins. Secondly, improper collection schedules, lack of bin monitoring, and inadequate size of bins to accommodate the waste dropped between two consecutive collections also give birth to the same problem. Inappropriate bin allocation and positioning push people to dump the waste at the roadside and in residential areas. Predicting waste generation is also necessary to develop a sustainable SWM system for managing future waste generation. These are the many factors that instigated the development of an advanced technology-based and cost-effective SWM system for proper management. Moreover, with the existing SWM systems, it is not feasible to collect, transport, and dispose of such waste in the future. The study has identified and analyzed various technologies to develop a smart SWM system. The paper has also implemented a smart bin using an ultrasonic sensor and an RFID-based waste collection system. A conceptual architecture of an RFID-based waste collection and monitoring system is shown.

The study is structured as follows: Sect. 2 sketches the picture of a typical SWM system, and various issues related to the existing SWM system are explained in Sect. 3. Section 4 defines the concept of a smart city and explains SWM services in the context of smart city development. Section 5 demonstrates the smart SWM

system and various technologies involved in its development. Section 6 illustrates the experimental setup and testing results of a simplified smart bin and RFID-based waste collection system. The study is concluded in Sect. 7.

2 Solid Waste Management

Generally, the term “Solid waste management” is interchangeably used for municipal solid waste management. SWM is observed as one of the integral components of urban services that comprise a set of different tasks from collecting waste from households, streets, and marketplaces to final disposal as landfill or recycling. It involves various processes and actions needed to handle the waste from its inception to final disposal. SWM problems are significantly more severe in developing countries compared to developed countries. Based on the literature survey and reports, it is concluded that the major factors behind the waste management problems are inefficient collection processes, high cost, and no monitoring of waste [1]. These factors lead to poor waste management practices that directly depreciate the resources, energy extracted, and recycled materials from a significant portion of the MSW. MSW is most commonly pronounced as “solid waste”, “waste”, “garbage,” or “trash.” SWM is needed to obtain sustainable urban development by collecting, transporting, recycling, and disposing of solid waste at the appropriate time to reduce the harmful impact on humans, the environment, and animals [2]. The SWM processes mainly consist of the deployment of bins, the collection of waste, and transportation and recycling for disposal. Recycling is the process of converting waste into useful raw materials. Figure 1 depicts the various stages involved in the SWM service framework.

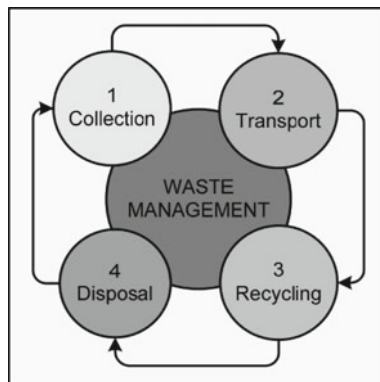


Fig. 1 Different stages in the SWM service framework

3 Issues in Current SWM System

Based on the literature analysis, published reports, and surveys, several issues are identified in the existing SWM system, especially in developing countries like India. These issues are enumerated as follows:

3.1 No Prediction About Future Waste Generation

The existing SWM system does not monitor waste, so municipalities lack data about waste generation. As a result, municipalities cannot predict the amount of waste generation in the future. This lack of prediction will lead to improper SWM-related infrastructure development and the need for more resources as municipalities do not know how much waste they will manage in the future.

3.2 Scarcity of Data Analytics Tools Applications

The regular SWM system neither involves any real-time waste data nor performs regular offline surveys to collect valuable data related to waste generation. Consequently, municipalities lack data and cannot apply any data analytics tool to determine the waste generation patterns in a particular area or bin. This scarcity leads to inappropriate planning and poor services.

3.3 Low Efficiency in Waste Collection

The regular SWM system comprises a fleet of waste collection vehicles and their drivers, who follow a predefined, fixed path without any information about the status of the bin. This system does not measure the current waste level (empty, semi-empty, and full) in the bin. Therefore, drivers visit all the bins. As a result, semi-empty bins are emptied, and in contrast, already-filled ones need to wait until the following collection round. Furthermore, since drivers visit empty bins, waste collection through predefined fixed paths yields a waste of time, fuel consumption, and excessive use of resources. The system performs inefficiently and requires high operational costs. Overall, this system does not have the tools and techniques to optimize schedules and waste collection routes. Low efficiency is considered a major drawback of the current system.

3.4 No Real-Time Monitoring

The existing system does not perform real-time tracking of bin and waste collection vehicles. Besides the current waste level monitoring, the system does not measure environmental and hazardous indicators such as temperature and methane gas concentration; therefore, it is impossible to know when a fire occurs in the bin. In addition, another drawback is being unable to track if any movement happens in the bin.

3.5 Unavailability of Online Platform or Web Portal

The current system does not have any type of platform or web portal to keep track of the drivers operating on the field. Furthermore, there is no such platform or web portal to enable the citizens to participate in SWM services such as registering a complaint and urgent services to remove waste scattered or dumped in residential areas.

3.6 No Illegal Dump Monitoring

The existing SWM system does not comprise any process to identify and localize the illegal waste dumps on roadsides and in residential areas. Also, it does not perform regular surveys and checks to determine these dumps. But the employees from the municipality manually verify it if someone makes a complaint regarding it and remove it if it is found. This traditional method of illegal dump identification and collection comprises humans who move around the city to identify the illegal dumps. Nowadays, the frequency of waste dumping on streets and roadsides has significantly increased as a consequence of growing waste generation and the scarcity of dumping spots. This open dumping primarily impacts the cleanliness of the city and deteriorates the surrounding environment and the inhabitants' health.

3.7 No Analysis of Citizens' Behavior Toward Waste

The existing SWM system does not assess the role of citizens in successful waste management. People play an essential role in making the waste collection process highly efficient in time and energy. If people are attentive and know about the consequences of waste scattering in the open, they will take proper care when dropping waste so that it cannot drop outside the bin. But according to the survey performed, people have very little awareness about the consequences of waste scattering. They

do not possess a positive attitude toward waste, especially in developing and economically weak countries. Therefore, they do not take proper care and drop the waste around the bin. Due to people's behavior, a flood of waste seems to be in the vicinity of the bins. This attitude leads to environmental pollution, damages the cleanliness of the city, and makes the process of waste collection from the bin very time-consuming, tedious, and inefficient.

4 Concept of the Smart City

Nowadays, the paradigm of development has completely shifted from conventional methodology to the approach of smart and sustainable development, a most popular catchphrase in contemporary development. The incorporation of innovation in sustainable development has evolved due to the pervasiveness of IoT technology and ubiquitous Internet connectivity. The development of urban areas innovatively and sustainably is called a "smart city." Now, the term "smart" is defined as an urban region incorporated with intelligent systems or a city with plans, concepts, and people with intelligent understanding. Here, smart systems are not bound to information and communication technology (ICT) based ones; the term "smart" refers to the intelligence to create new designs, organizations, services, etc. According to this viewpoint, the smartness of a city can be interpreted as the ability to integrate all its resources to effectively attain the goals and fulfill the set objectives and purposes [3]. A smart city is an urban area that incorporates spatial technologies, IoT technologies, and ICT to transfer information for regulating the resources and managing the city services in an efficient way [4]. One of these services is waste collection and transportation to recycling and landfill disposal facilities [5].

Furthermore, a smart city is also defined as a city that functions competently and progressively in the following fundamental elements; smart economy, smart mobility, smart environment, smart people, smart living, and smart governance. It is built over the combination of intelligent features and activities of self-decisive and has self-sufficient and sustaining well-informed people [6]. This definition of a "smart city" implies that it is a concept used to depict the novel plans and goals to develop and build urban regions. It focuses on ICT and IoT technologies for information exchange and significantly highlights the applications of spatial technologies, data acquisition, and identification technologies to persuade smart and sustainable urban development. It exploits the existing and new resources in an optimal way and collectively applies the human resources and technological infrastructure in the development process [7]. It primarily manages urban infrastructure, such as urban transportation, energy distribution and utilization, governance, civic services, environmental monitoring, resource conservation and utilization, healthcare services, etc., in smart and sustainable ways. It also critically analyzes the different application aspects of technologies to optimally utilize the physical infrastructure for developing the novel and smart SWM system [8]. Figure 2 displays the components of a smart city with their focused areas.

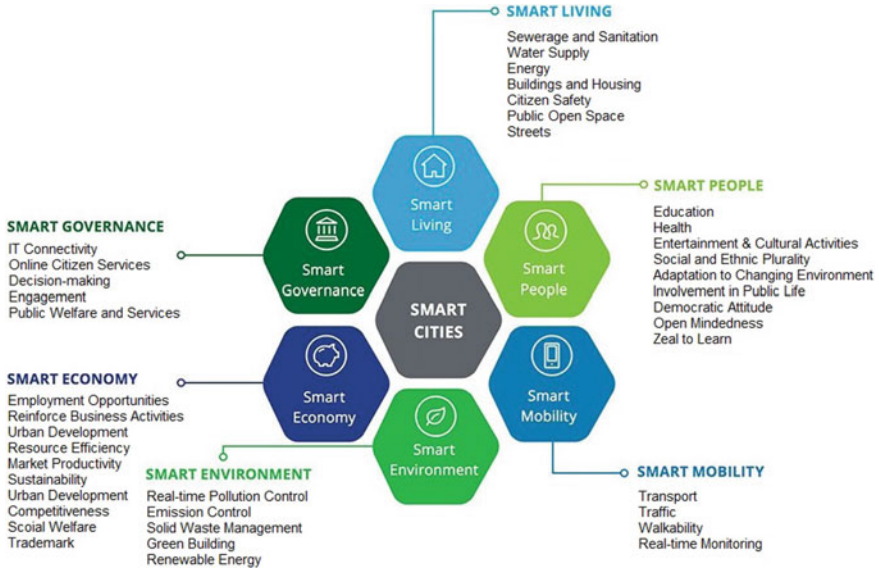


Fig. 2 Main components of a smart city

All the above definitions imply that the smart city focuses on six major challenges: *“Providing an economic base; building efficient urban infrastructure; improving the quality of life and place; ensuring social integration; conserving natural and environmental qualities; and guaranteeing good governance [9].”*

4.1 Smart Economy

It incorporates technology and innovation to reinforce business activities, employment opportunities, and urban development [10]. It adopts innovation, new entrepreneurial initiatives, resource efficiency, sustainability, trademarks, increased labor market productivity, competitiveness, high social welfare, flexibility, and integration with the global market to improve citizens’ quality of life [11].

4.2 Smart Mobility

It is an intelligent transport and mobility network with multiple modes of transportation. It incorporates advanced technology in transportation, such as real-time monitoring and control systems. It is just a rethinking of the existing mobility infrastructure used in daily life and business [12, 13].

4.3 *Smart Environment*

It connects computers and intelligent devices to daily life activities and tasks. It allows users to engage and interact seamlessly with their immediate surroundings for better understanding and control. It has evolved with the introduction of intelligent technologies such as IoT, software-based services, etc. It comprises smart cities, smart manufacturing, and smart and green homes. It is also used to protect and manage natural resources such as SWM systems, emission control, renewable energy real-time pollution, and atmosphere parameter monitoring [14].

4.4 *Smart People*

They possess enhanced creativity, a lifelong zeal to learn, open innovation with open-mindedness, and social and ethnic plurality. They also have the flexibility to adapt to the changing environment and the desire to contribute to the knowledge pool. They believe in democratic principles and involvement in public life [15].

4.5 *Smart Living*

It is the contemporary trend of living incorporating innovation to enhance the living standard in urban spaces. It provides the opportunity for people to benefit from the new ways of living life. It comprises indigenous and innovative solutions to make life easier and more effective through social integration [16].

4.6 *Smart Governance*

It is established in the smart city and involves technology and innovations for engagement, welfare, and services. It effectively utilizes cutting-edge technologies, especially ICT, enhancing decision-making through greater collaboration among various stakeholders. It allows more significant involvement of citizens in government approaches [17, 18].

Besides the six components listed above, two additional components are also essential for a smart city vision.

4.7 *Smart Infrastructure*

It integrates the city infrastructure, such as water supply and energy distribution networks, streets, buildings, etc., with intelligent technologies, namely sensors, smart grids, surveillance systems, etc. It is termed a “cyber-physical system” that integrates the management of all elements in a single window and incorporates various technological tools to compile and analyze data for improving efficiency, sustainability, productivity, and safety objectives [19].

4.8 *Smart Services*

It uses ICT, IoT, and other advanced technologies to deliver health care, hospitality education, safety and security, and other services in the entire urban space. It is a digital service that reacts and performs decisions actions on the analysis of data collected from networks, intelligent technical systems, and various platforms [20].

Moreover, these components are interconnected through ICT infrastructure and require data collection. They are also integrated with the hard-urban infrastructure to provide intelligent services to all residents of a city. In addition, governance and control over the subsystems are essential to becoming a smart city vision and mission success. Figure 3 displays the conceptual architecture of a smart city. It is deduced that a smart city comprises a broader range of services in the urban space to provide life with dignity and prosperity on a healthy planet. Smart SWM is one of the vital services to create a clean and green city that ultimately contributes to building a smart environment.

5 *Smart SWM System*

In the current era of urban development, the concept of a “smart city” has significantly emerged to improve the quality of life through the integration of ICT, IoT, and other cutting-edge technologies. Therefore, the traditional SWM system must be migrated to an intelligent SWM system in urban areas for waste collection, transportation, disposal, planning, and policymaking to achieve the goal of a smart environment, one of the components of the concept of the smart city mission. Figure 4 demonstrates the service framework of an ICT and IoT-enabled smart city, where SWM is one of the services.

Now, a smart SWM system can be defined as a system that involves the utilization of ICT and other cutting-edge technologies to make the waste collection and transportation process more efficient in terms of running costs, time, and resource utilization in an environmentally safer manner. Generally, these systems are equipped with IoT and ICT components used to capture and transfer data in real time. The



Fig. 3 Conceptual architecture of a smart city



Fig. 4 Conceptual service model of an ICT and IoT-enabled smart and sustainable city

collected data helps optimize waste collection routes and schedules and prompts future innovation. Figure 5 presents a schematic smart SWM system built based on the literature analysis.

This study primarily concentrates on the progress of applications of advanced technologies in SWM systems. A thorough analysis of the literature is performed to achieve this goal, and different ICT, IoT devices, and image processing technologies are identified that have been applied in developing an intelligent SWM service framework. The identified technologies are an effective solution to overcome the

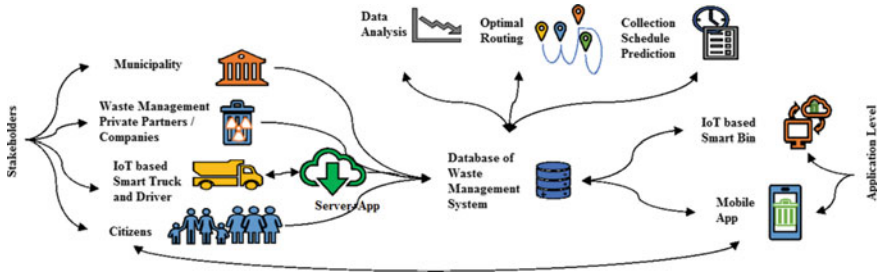


Fig. 5 Schematic diagram of a smart SWM system

identified issues in the existing SWM systems. The identified technologies and their applications are listed in Table 1. A detailed discussion about each technology with its functions, applications, and limitations is presented in the subsequent sections.

6 Experimental Prototypes and Results

The previous section has demonstrated a schematic diagram of an IoT-based SWM framework. This framework consists of a smart bin that incorporates an ultrasonic sensor, a GSM module, and an RFID tag. The following subsections explain the physical architecture of the smart bin and the experimental results.

6.1 Smart Bin Architecture and Testing Results

Figure 6 depicts the schematic block diagram of a simplified smart bin with the installation of different devices. The implemented architecture consists of an Arduino Uno microcontroller board in its core, which is connected with an ultrasonic sensor to measure the waste inside the bin, a light-emitting diode (LED) indicator to show the certain levels of waste in the bin, and a GSM module to transfer the captured data.

Figure 7 shows the hardware connections of different peripheral devices with the microcontroller. The ultrasonic sensor HC-SR04 needs a trigger signal to start measuring the distance. As soon as the ultrasonic sensor receives the trigger signal from the microcontroller, it generates an ultrasonic wave and starts the timer. When the echo signal is received after being reflected from the target, the timer stops, and at the same moment, a high output signal from the ultrasonic sensor is transmitted to the microcontroller. The duration of this output signal is the time difference between the transmitted ultrasonic signal and the received echo signal. Therefore, the trigger and echo pins of the ultrasonic sensor are coupled with the input/output pins of the microcontroller. The GSM module SIM900 is in serial with the microcontroller,

Table 1 Different technologies and their applications in the SWM service system

Category	Examples	Potential applications in SWM system
Spatial technologies	GIS	Disposal sites identification and selection, bin allocation and management, and route identification and optimization for waste collection
	GPS	Bin and collection vehicle location tracking
Identification technologies	Barcode quick response (Q.R.) code RFID	Bin identification
Data acquisition technologies	Ultrasonic and infrared sensor	Waste level measurement inside the bin
	Load cell sensor	Waste weight measurement
	Accelerometer and proximity	Compute the bin lid state
	Biosensor gas sensor	Detect the hazardous gas and chemicals inside the bin to stop the diffusion in the environment
	Temperature	Measure the temperature to control the fire
	Humidity	Measure the humidity inside the bin to prevent the decomposition of biodegradable material
Data communication technologies	ZigBee, Wi-Fi, Bluetooth	Long-range communication
	VHFR, LoRa, and GSM	Short-range communication
Artificial intelligence (artificial neural network)	Convolutional neural network (CNN) Long short-term memory (LSTM)	Waste detection and classification—glass, metal, trash, cardboard, plastic, medical, recyclable, non-recyclable, e-Waste, polyethene, organic, inorganic, etc. Object detection—battery, waste bags, etc. Waste forecasting and gas prediction inside the bin

as the RX and TX ports of the SIM900 are coupled with the TX and RX ports of the Arduino Uno board. Both LEDs are attached to the input/output pins of the microcontroller.

The testing is done to confirm that the built smart bin is measuring the waste level properly. To perform the testing, two threshold levels of waste are fixed as TL1 and TL2. TL1 is fixed approximately at 75%, while TL2 is fixed approximately at 90% of the height of the bin from the bottom. When the waste reaches level TL1, the first warning message is sent by the microcontroller through the GSM module to the municipal authority, and the green LED turns ON. The second warning is

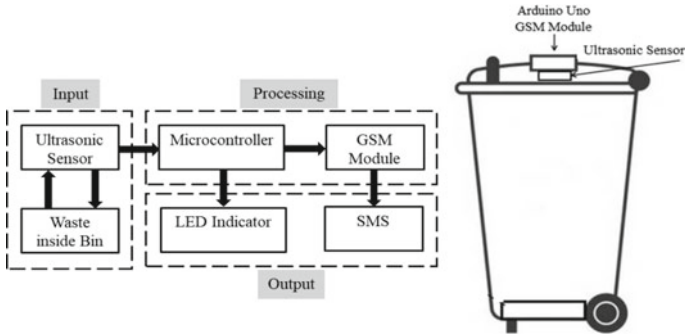


Fig. 6 Schematic block diagram of a simplified smart bin

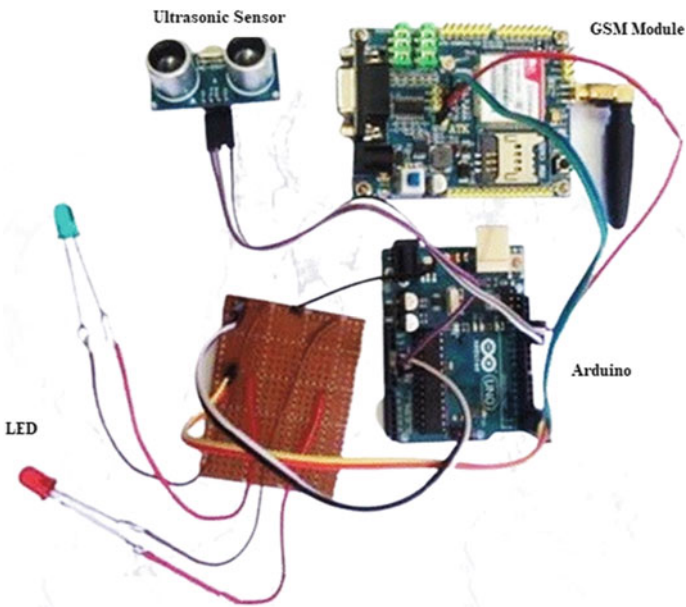


Fig. 7 Typical connection diagram of smart bin components

sent as soon as the waste level crosses level TL2 and the red LED turns ON. LEDs are incorporated to show the waste level in the bin to the people. The testing was also performed for the empty bin or the level of waste below TL1. In this situation, neither a warning message was sent, nor any LED turned ON. The received warning messages are shown in Fig. 8.

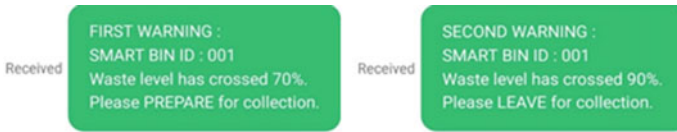


Fig. 8 Received warnings after crossing the threshold TL1 and TL2

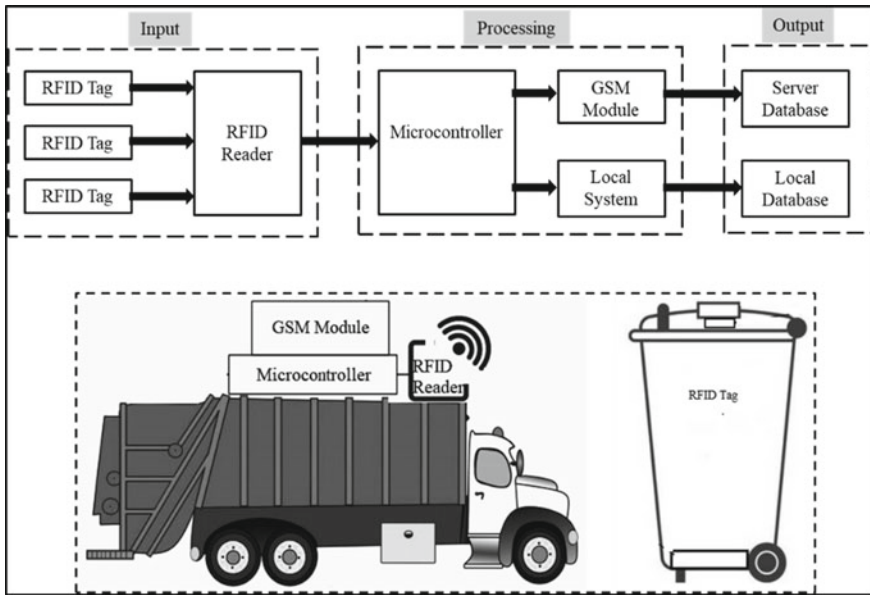


Fig. 9 Schematic diagram of bin identification using RFID reader and tag

6.2 RFID-Based Waste Collection from Smart Bins

Figure 9 displays the system architecture of the bin identification for waste collection. The placement of devices and communication flow are also depicted. The prototype incorporates a microcontroller in its core, an RFID reader, a GSM module at the peripheral, and an RFID tag mounted on the outer surface of the bin.

6.3 System Workflow and Testing Results

When the RFID reader mounted on the waste collection vehicle comes under the frequency range of the RFID tag on the bin, then the reader sends a request signal to the tag. Then the tag responds, and the reader starts reading data from the tag. The captured data is sent to the microcontroller for further processing. The valid data is

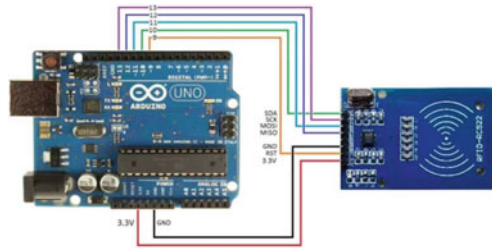


Fig. 10 Connection interface of microcontroller with the RFID reader

Table 2 Information of visited bins in the database

S. No.	Bin ID	Tag No.	Date and time	Status
1	B1	000007EC2AB4	19/12/2018 11:05:18	Collected
2	B2	00000AD31DF6	19/12/2018 11:35:39	Collected
3	B3	000008A4002E	19/12/2018 11:59:02	Collected
4	B1	000007EC2AB4	20/12/2018 10:08:38	Collected
5	B2	00000AD31DF6	20/12/2018 11:01:21	Collected
6	B3	000008A4002E	20/12/2018 11:40:09	Collected
7	B1	000007EC2AB4	21/12/2018 11:10:11	Collected
8	B2	00000AD31DF6	21/12/2018 11:49:56	Collected
9	B3	000008A4002E	21/12/2018 12:20:29	Collected

converted into a pre-specified format; otherwise, it is discarded. This data is sent to the central server using the GSM module and updated in the database. The hardware interface connection setup of the Arduino microcontroller with RFID reader is shown in Fig. 10, and testing results are given in Table 2.

The above RFID-based waste collection system does not incorporate any system to monitor the collection truck and find the waste in the bin and its surroundings. Due to physical infrastructure constraints, these functions are not involved in the experimental setup. Therefore, a conceptual system that involves the above features and its architecture is depicted in Fig. 11. This architecture comprises bins with an RFID tag, a truck, and a monitoring station. The truck is mounted with data acquisition devices, namely an RFID reader to read data from the tag, a camera to capture images before and after waste collection from the bin, GPS to determine the truck’s real-time location, and a GSM module to send the captured data by acquisition to the central server. The waste monitoring station consists of a GSM receiver, a GIS program, and a database and a webpage for the user interface.

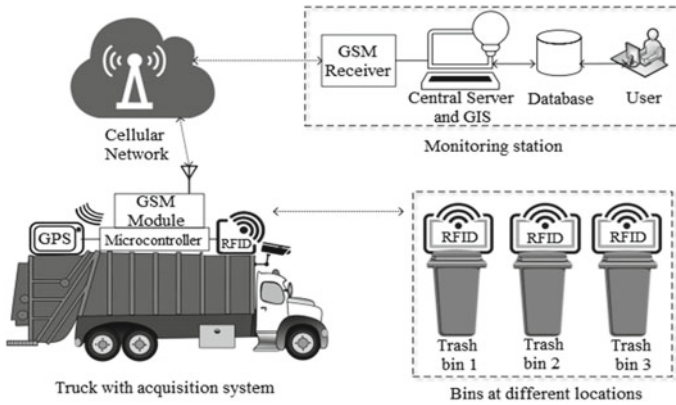


Fig. 11 Conceptual architecture of RFID-based waste collection and monitoring system

7 Conclusion

The study concludes that the SWM service framework has shifted from the traditional approach to an advanced technology-based smart system. These technologies have played a vital role in increasing collection efficiency, minimizing the adverse consequences of waste on the environment, and overall improving the cleanliness of the city. Spatial technologies enhance planning, route identification and optimization, collection vehicle tracking and scheduling, and environmental impact assessment. Identification technologies such as barcodes and RFID have transformed the waste segregation process from manual to automatic sorting. They have also used recycling process intelligence, reducing disposal landfills, and identifying bins and trucks. Various sensors have been incorporated to measure the waste level in the bin, and environmental parameters, namely humidity, temperature, and gas concentration. Image processing has been utilized for waste sorting, waste dump detection, street cleanliness assessment, and various activity identification and classification in the proximity of bins. Finally, different types of captured data are sent to a central server through various communication technologies, making the system accessible in real-time.

The study has demonstrated a smart bin using an ultrasonic sensor for real-time monitoring of the waste level inside the bin. The study has also implemented an RFID-based waste collection system and illustrated the results. The implemented system resolves the issue related to real-time monitoring.

References

1. Kaza S, Yao LC, Bhada-Tata P, Van Woerden F (2018) What a waste 2.0: a global snapshot of solid waste management to 2050. World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, Washington. <https://doi.org/10.1596/978-1-4648-1329-0>
2. Ayilara MS, Olanrewaju OS, Babalola OO, Odeyemi O (2020) Waste management through composting: challenges and potentials. <https://www.mdpi.com/2071-1050/12/11/4456/htm>. <https://doi.org/10.3390/su12114456>
3. Anthopoulos LG (2017) The rise of the smart city. In: Public administration and information technology. Springer, Berlin, pp 5–45. https://doi.org/10.1007/978-3-319-57015-0_2
4. Abdullah N, Alwesabi OA, Abdullah R (2019) IoT-based smart waste management system in a smart city. Springer International Publishing. https://doi.org/10.1007/978-3-319-99007-1_35
5. Medvedev A, Fedchenkov P, Zaslavsky A, Anagnostopoulos T, Khoruzhnikov S (2015) Waste management as an IoT-enabled service in smart cities. In: In the Internet of Things, smart spaces, and next generation networks and systems. Springer, Cham, pp 104–115. <https://doi.org/10.1007/978-3-319-23126-6ss>
6. Giffinger R, Pichler-Milanović N Smart cities: ranking of medium-sized European cities. <http://www.smartcities.eu>
7. Ali T, Irfan M, Alwadi AS, Glowacz A (2020) IoT-Based smart waste bin monitoring and municipal solid waste management system for smart cities. Arab J Sci Eng 45:10185–10198. <https://doi.org/10.1007/s13369-020-04637-w>
8. Sanchez L, Munoz L, Galache JA, Sotres P, Santana JR, Gutierrez V, Ramdhany R, Gluhak A, Krco S, Theodoridis E, Pfisterer D (2014) SmartSantander: IoT experimentation over a smart city testbed. Comput Netw 61:217–238. <https://doi.org/10.1016/j.bjp.2013.12.020>
9. Yigitcanlar T, Lee SH (2014) Korean ubiquitous-eco-city: a smart-sustainable urban form or a branding hoax? Technol Forecast Soc Change 89:100–114. <https://doi.org/10.1016/j.techfore.2013.08.034>
10. Cornetta G (2020) Social, legal, and ethical implications of IoT, cloud, and edge computing technologies. <https://doi.org/10.4018/978-1-7998-3817-3>
11. Kézai PK, Fischer S, Lados M (2020) Smart economy and startup enterprises in the Visegrád Countries—a comparative analysis based on the crunchbase database. Smart Cities 2020, vol 3, pp 1477–1494. <https://doi.org/10.3390/SMARTCITIES3040070>
12. Benevolo C, Dameri RP, D'Auria B (2016) Smart mobility in smart city action taxonomy, ICT intensity and public benefits. In: Lecture notes in information systems and organisation. Springer Heidelberg, pp 13–28. https://doi.org/10.1007/978-3-319-23784-8_2
13. Freitas A, Brito L, Baras K, Silva J (2017) Smart mobility: a survey. In: Internet of Things for the global community, IoTGC 2017—Proceedings. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/IoTGC.2017.8008972>
14. Nugent CD, McClean SI, Cleland I, Burns W (2014) Sensor technology for a safe and smart living environment for the aged and infirm at home. In: Comprehensive materials processing. Elsevier, pp 459–472. <https://doi.org/10.1016/B978-0-08-096532-1.01319-4>
15. Gupta S, Ziaul Mustafa S, Kumar H (2017) Smart people for smart cities: a behavioral framework for personality and roles. In: Advances in smart cities. Chapman and Hall/CRC, pp 23–30. <https://doi.org/10.1201/9781315156040-3>
16. Kozłowski W, Suwar K (2021) Smart city: definitions, dimensions, and initiatives. Eur Res Stud J XXIV 509–520
17. Bernardo MDRM (2018) Smart city governance: from E-Government to smart governance. In: Smart cities and smart spaces: concepts, methodologies, tools, and applications. IGI Global. <https://doi.org/10.4018/978-1-5225-7030-1.ch009>
18. Pereira GV, Parycek P, Falco E, Kleinhans R (2018) Smart governance in the context of smart cities: a literature review. <https://doi.org/10.3233/IP-170067>

19. Nishi H, Nakamura Y (2020) IoT-based monitoring for the smart community. In: Urban systems design. Elsevier, pp 335–344. <https://doi.org/10.1016/b978-0-12-816055-8.00010-5>.
20. Alt R, Demirkan H, Ehmke JF, Moen A, Winter A (2019) Smart services: the move to customer orientation. <https://link.springer.com/article/https://doi.org/10.1007/s12525-019-00338-x>. <https://doi.org/10.1007/s12525-019-00338-x>

Image Forgery Detection: A Review



Preethi Vijayakumar, Elizabeth Mathew, M. Gayathry Devi, P. T. Monisha, C. Anjali, and Jisha John

Abstract Images can now be easily modified due to advances in digital image processing. Image forgery is the process of manipulating or changing a digital image in order to conceal vital information. Nowadays, as technology advances and the performance of low-cost computers improves, forgery in images has rapidly increased. Fraudulent images are quite impossible to recognize and detect through human eyes. The process of modifying fake images has been made extremely easy with powerful editing tools that are available for free. The software used works so well that it's tough to tell the difference between legitimate and faked photos. Because digital photographs and images are used in so many locations, detecting digital image counterfeiting is critical as it can be a misuse of technology and image authenticity is essential in various social platforms. There are several forgery techniques developed to find the authenticity of images. This paper will benefit researchers in comprehending current algorithms and strategies in this sector, with the goal of eventually developing new and more efficient detection algorithms.

Keywords Image forgery · Copy-move · Splicing

1 Introduction

Forgery in images means manipulation of the digital image to hide some substantive or useful information of the image. Image forgery is not new. As powerful image processing software is already widely available, digital images are now simpler to manipulate and edit. These tools are versatile and come with a user interface. Manip-

P. Vijayakumar · E. Mathew (✉) · M. Gayathry Devi · P. T. Monisha · C. Anjali · J. John
Department of CSE, Mar Baselios College of Engineering and Technology,
Trivandrum, Kerala, India
e-mail: aanya.elizabeth@gmail.com

C. Anjali
e-mail: anjalichandrika@gmail.com

J. John
e-mail: jisha.john@mbcet.ac.in

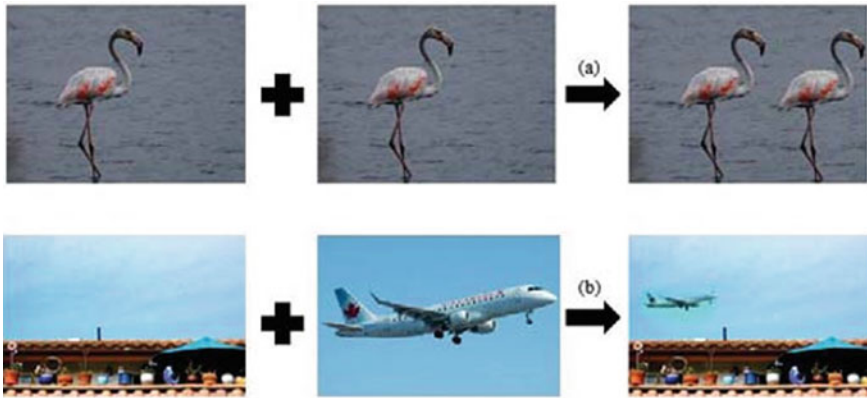


Fig. 1 Copy-move and spliced image forgery

ulation and tampering of images may now be done efficiently not just by experts but even by inexperienced users. These manipulated images are unrecognizable and appear so genuine in perception that they lose their legitimacy. The authenticity of digital images cannot be detected by the human eyes as it does not leave any visual clue indicating the forgery. Scientific advances in the realm of computer graphics have led to development of many picture editing applications, including Photoshop, Lightroom, and others. This modifies the image's content without creating any irrefutable evidence of fraud. Today's advanced image editing software allows anyone to make quick and easy changes to their photographs and images. As a result, detecting these alterations and adjustments becomes difficult for people and the reliability and credibility of digital images come under scrutiny (Fig. 1).

Abuse of digital forgeries has been regarded as a severe concern in several fields. In a variety of social situations, image authenticity is significant. Establishing techniques to confirm the validity and veracity of digital photographs is critical, especially when these photos are used as evidence in court, as part of medical reports, as reports with large financial expenditures, as part of recovery records, or even images of crime scenes for investigation. As a result, one of the most essential aspects of image forensics is the identification of digital image falsification. Splicing and copy-move forgeries examples are depicted above from paper [1].

1.1 Types of Forgery

There are two basic approaches of tampering with digital images: active and passive which can be further divided as shown in Fig. 2.

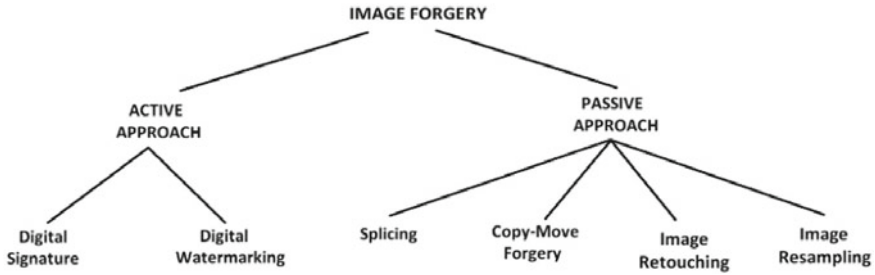


Fig. 2 Image forgery classification

Active forgery requires prior information about the image. It typically involves coming up with varied types of inserting watermarks or fingerprints of the image content into the digital image. Active forgery can be categorized as:

- A. Digital Signature—Signatures are one of the most important methods of getting a person’s validation for the information in a document since they are made up of a mix of individual textual elements. It is a mathematical strategy that helps in preventing image alteration, such as spoofing, and maintaining image integrity.
- B. Digital Watermarking—It is the practice of embedding a digital code into digital information such as a picture, video, or other type of file. It is mostly employed in copyright and to conceal certain important information.

Passive forgery acts in the form of a blind detection, that is, there is no prior knowledge about the image with this procedure. It identifies digital image fraud by examining unique inherent clues or patterns that appear during the manipulation or creation of digital images. Types of passive forgery include:

- A. Splicing—This type of forgery means copying one or more sections from a source image and pasting them onto a destination image.
- B. Copy-Move Forgery—This sort of forgeries entails copying and pasting one or more elements of a picture onto the same image.
- C. Image Retouching—This is the process of making minor alterations to an image in order to prepare it for a final presentation, such as polishing.
- D. Image Resampling—This is the forgery where the image’s pixel count is modified.

2 Literature Review

The system’s first section focuses on detecting copy-move forgeries. The method’s second part is recolored picture detection. The proposed method in [1] uses the structural similarity index and a threshold value to determine the degree of similarity between altered and original images. A rectangle mark is used to highlight the altered

area. Detection of recolored images uses Interchannel correlation and the illumination consistency for feature extraction. When the system was employed to detect picture alteration using copy-move and recoloring, it produced a high-quality result for both types of image forgeries.

The authors of [2] proposed a system where the Expectation-Maximization (EM) rule is used to analyze a digital image. The first phase involves extraction of features using the Scale Invariant Feature Transform (SIFT) and key comparison. Phase two was dedicated to data collecting and forgery detection, while the third phase was dedicated to determining the manipulation area. SIFT aligns the image with the feature key points, then applies geometric transformations to rotation and scaling. The EM rule is a two-step method. In E-step, it estimates the probability of every sample to each model and within M-step, it will evaluate a precise correlation between the samples. The tampering of the images is displayed through a red rectangular region which is determined through the variation of frequencies. The image is cast if there is a smooth frequency. The image has been tampered with if the output indicates a frequency.

The spliced image's textural examination is employed in paper [3] to detect image forensics. Ground truth's feature set masks Find Gray level regional Maxima (FGM) and Entropy-based Edge (EbE) is extracted from the forged picture's localized entropy of the median filter residual (MFR). For the ground truth mask, the local range is also used to derive the feature set. For solely the performance evaluation of the suggested technique, a cubic SVM classifier was used to train the given novel feature vector. On five distinct kinds of images, the proposed image forensics detection technique (IFD) was put to the test: median filtered, JPEG compressed, average filtered, and so on. An input image is used to identify the Cut-Paste zone in the proposed system for identification of spliced image forensics, eliminating the necessity for a trained SVM classifier.

A convolutional neural network is used in this paper [4] to present a new deep learning-based picture forgery detection method (CNN). This method detects counterfeit medical pictures by using a CNN. To generate an estimated noisy image, a Weiner filter-based noise reduction is applied to an input image. The Convolutional Neural Network is given a regression filter and normalization. This information is sent to the Residual Network, which determines whether the image is altered or real.

The five stages of the Image Copy-Move Forgery Detections (IC-MFDs) proposed in paper [5] are: pre-processing of the image, image partitioning into overlapping blocks, evaluating each block's mean and standard deviation, lexicographically cataloguing feature maps, and feeding the feature vector to the Support Vector Machine (SVM) classifier to assess whether the image is authentic or forged. Experiments using a standard dataset of copy-move forged pictures, MICC-F220, are used to test the proposed approach.

Kanwal et al. [6] provided two approaches for detecting forgeries. Two kinds of features are obtained from the first method, Local Binary Pattern (LBP) and Local Ternary Pattern (LTP). LBP was first presented for extracting local image characteristics and has a two-valued code matrix, however LTP was offered as a better alternative to LBP because it generates three-valued code. However, LTP was

offered as a better alternative to LBP because it generates three-valued code, the Support Vector Machine (SVM) receives this vector. The second method uses Fast Fourier Transform (FFT) and Enhanced Local Ternary Pattern (ELTP). The Discrete Fourier transform (DFT) can be computed using the FFT method and ELTP is an extended variant of LTP that achieves superior picture features. In this approach, the resulting feature vector containing each FFT block of the picture is assigned an ELTP code, which is subsequently given to SVM for classification as forged or legitimate.

The authors of [7] suggested a spliced image detection technique based on Markov characteristics. The DCT domain extracts two types of Markov features: coefficient-wise and block-wise Markov features. The first used coefficient correlations between consecutive blocks, while the second used correlations between adjacent coefficients. The two Markov characteristics are merged to create a feature vector, which is then used to detect spliced and authentic images using the Support Vector Machine (SVM).

The research [8] proposed a method for detecting spliced forgeries by using Discrete Wavelet Transform (DWT) and histograms of discriminative robust local binary patterns. A color image is transformed to the YCbCr color space, and then DWT is applied to the Cb and Cr components of the digital image. The texture variation in each DWT subband is described using the Dominant Rotated Local Binary Pattern (DRLBP). The feature vector is created by concatenating the DRLBP from each subband. The Support Vector Machine (SVM) classifier is used to obtain and assess the performance of the DWT-DRLBP descriptor. Experimentation and evaluation was done using three publicly available benchmark datasets.

The authors of [9] combined block-based and keypoint-based forgery detection algorithms with adaptive overlapped segmentation and feature point matching. The image gets decomposed into different super pixels by the adaptive overlapped segmentation algorithm. The feature points are extracted using the Speeded Up Robust Features (SURF) technique. The extracted features get paired to one another and if the match exceeds the predefined limit, the regions suspected of forgery will be highlighted, and combined regions will be created. To depict the discovered forgery regions, morphological techniques are applied to the merged regions.

Abdalla et al. [10] proposed a system that uses Generative Adversarial Networks (GAN) and Convolutional Neural Networks (CNN). The GAN is composed of a generator and a discriminator where the generator creates an image from random noise input and the discriminator detects whether the image is genuine. CNN was used for feature extraction. The data is fed into two networks (GAN and CNN) before being sent to a linear classifier and being localized.

The paper [11] is split into two parts. To begin, a hybrid method is provided, which combines existing block-based and non-block-based copy-move forging algorithms. Secondly, the performance of the proposed technique is evaluated using a range of picture characteristics such as SIFT, SURF, MSER, MinEigen, FAST, and Harris. The results of the study indicate that the proposed picture forgery detection system is able to detect copy-move forgery with high accuracy and speed. Apart from that, the proposed technique works well with graphics that have been smoothed and brightened.

Using the Local Binary Pattern (LBP) and the Discrete Cosine Transform (DCT), the paper [12] proposed a method for detecting splicing image frauds. The input image's chrominance component is first separated into overlapping blocks. The LBP is then calculated for each block and converted into the frequency domain utilizing 2D DCT. Finally, the standard deviations of DCT coefficients of all the blocks was determined and used as features in a Support Vector Machine classification (SVM).

A technique for recognizing duplicate regions in photographs that takes advantage of statistical characteristics of the image was given in this study [13]. The mean and variance were used to divide the image into pixel blocks. The variance is used to determine how each pixel in a block differs from its neighbors, while the mean is used to determine what factors contribute to the pixel intensity of each individual element in the image. The proposed method is assessed and compared using the Discrete Wavelet Transform (DWT) of an image, and then similarity and threshold are determined. It has been proven that the algorithms presented outperform existing approaches.

The authors of [14] developed a deep learning-based solution for identifying image splicing. The input image was initially preprocessed using the 'Noiseprint' approach, which suppressed the image content, to get the noise residual. The popular ResNet-50 network was then used as a feature extractor. Finally, using the Support Vector Machine (SVM) classifier, the obtained features were classed as spliced or authentic indicating class 0 and 1 respectively with a RBF kernel.

The image is separated into square pieces that overlap and have a fixed side length, and each of the blocks is broken into evenly spaced k subblocks using the suggested algorithm in paper [15]. A k -dimensional vector is created by combining the pixel intensities of each sub-block, which is then used as an attribute for each block, using a sliding window. A KD-tree is used to hold the combined features of all blocks. Each node's block in the KD-tree is compared to its nearest neighbor's block. The two blocks are considered clones if their correlation is greater than a predetermined threshold. The following are the important steps of the algorithm: first, divide the image to be studied into overlapping blocks first. Generate feature vectors from blocks in the second step. Third, construct a KD-tree using the feature vectors. Finally, in the KD-tree, look for near duplicate nodes and use a block matching algorithm to detect forgeries. A KD-tree is used to hold the combined features of all blocks. The block that corresponds to every node in the KD-tree is compared to the block that corresponds to the node's nearest neighbor.

3 Comparative Study

Table 1 summarizes a comparison study of the various methodologies reported in the publications along with their limitations.

Some of the issues encountered in these existing approaches was data set images were less leading to less accuracy. Most of the approaches just said the images were tampered and forgery has taken place, but did not specify which forgery it was.

Table 1 Comparison between different image forgery methods

Authors	Method used	Accuracy (%)	Drawbacks
Jijina et al. [1]	Structural similarity index (SSIM) and convolutional neural network (CNN)	82.46	Produces false positive and false negative
Alkawaz et al. [2]	Expectation maximization algorithm (EM rule)	–	Does not specify which type of forgery has taken place
Rhee [3]	Median filter residual (MFR)	98	Cannot detect spliced region
Purvisha and Kumar [4]	Deep learning using convolutional neural network (CNN)	–	No efficient performance
Ahmed et al. [5]	Spatial feature domain	98.44	Other types of image forgeries were not distinguished
Kanwal et al. [6]	Fast Fourier transform (FFT) and enhanced local ternary pattern (ELTP)	88.62	Uses complex operations like FFT, could not localize the spliced region and was not performed for other types of forgeries
Pham et al. [7]	Coefficient-wise and block-wise Markov features	96.90	Was done for spliced images only and localization was not performed
Siddiqi et al. [8]	Discrete wavelet transform (DWT) and dominant rotated local binary pattern (DRLBP)	98.95	Could not detect spliced region and was not performed for other types of forgeries
Sreelakshmy and Anver [9]	Adaptive overlapped segmentation algorithm and speeded up robust features (SURF)	–	Was not performed for other types of forgeries and SURF could not being able to locate the exact borderlines of the identical regions
Abdalla et al. [10]	Generative adversarial networks (GAN) and convolutional neural networks (CNN)	93–97	Not performed for other forgery types
Khan et al. [11]	Block-based and non-block-based technique	74–90	Could not be done for splicing and geometry-based forgeries
Alahmadi et al. [12]	Local binary pattern (LBP) and discrete cosine transform (DCT)	97	Not performed for other forgery types
Dixit et al. [13]	Statistical image features with discrete wavelet transform (DWT)	98	Copy-scale-move and copy-rotate-move duplicated image portions can be closely examined
Meena and Tyagi [14]	Noiseprint and ResNet-50	97.24	Could not localize spliced region
Moussa [15]	Block matching and KD-tree	94–96	Can make the algorithm parallel for extra large data set

Existing systems were able to detect either Copy-Move forgery or Spliced forgery but not both together. As mentioned earlier, these two types of forgeries are the most widely used. But most of the systems detect only one forgery type and a combination of both techniques is not widely used. The systems could detect whether forgery had occurred, but they could not pinpoint where it had happened. Localization or masking was not there in most of the approaches.

4 Conclusion

Detecting digital image forgery is becoming extremely relevant. Images in courtrooms, in medical reports and the military are vital in making important decisions. Through social media, false images or forged images might lead to negative consequences or even endanger a person's life. A lot of image forgery detection techniques have been introduced nowadays. Since passive techniques don't require any prior information, they are becoming more popular.

As discussed in the paper there are various approaches to image forgery detection. All these methods are able to detect different kinds of forgery. But their efficiency varies. Some methods fail to discover the real forged region, while others have extremely long processing times. This means that the techniques mainly have drawbacks like accuracy of detection, high time complexities, and are not efficient against attacks like rotation, scaling, brightness adjustments, etc. Another significant disadvantage is that the approach developed for copy-move forgeries is incompatible with image splicing or resampling, and vice versa.

An efficient way to detect image forgery would be a system that uses highly advanced deep learning and image processing tools having a large data set. The system should be able to detect atleast two or more passive forgery techniques and localize where the forgery has taken place. As a result of these constraints, there is potential for developing a more effective and faster image forgery detection approach that can overcome them. An image forgery technique that can identify more than one type of image forgery and localize the forged area is not available. In spite of having various techniques to identify image altering, the demand for an advanced method to detect digital image forgery is becoming more essential every day.

References

1. Jijina MT, Koshy L, Warriar GS (2020) Detection of recoloring and copy-move forgery in digital images. In: 2020 fifth international conference on research in computational intelligence and communication networks (ICRCICN), pp 49–53. <https://doi.org/10.1109/ICRCICN50933.2020.9296173>
2. Alkawaz MH, Veeran MT, Bachok R (2020) Digital image forgery detection based on expectation maximization algorithm. In: 2020 16th IEEE international colloquium on signal processing & its applications (CSPA), pp 102–105. <https://doi.org/10.1109/CSPA48992.2020.9068731>

3. Rhee KH (2020) Detection of spliced image forensics using texture analysis of median filter residual. *IEEE Access* 8:103374–103384. <https://doi.org/10.1109/ACCESS.2020.2999308>
4. Purvisha P, Kumar D (2019) Smart healthcare forgery detection using deep learning. https://ijariie.com/Smart_Healthcare_Forgery_Detection_Using_Deep_Learning_ijariie10489. Accessed 9 Apr 2022
5. Ahmed IT, Hammad BT, Jamil N (2021) Image copy-move forgery detection algorithms based on spatial feature domain. In: 2021 IEEE 17th international colloquium on signal processing & its applications (CSPA), pp 92–96. <https://doi.org/10.1109/CSPA52141.2021.9377272>
6. Kanwal N, Girdhar A, Kaur L, Bhullar JS (2019) Detection of digital image forgery using fast Fourier transform and local features. In: 2019 international conference on automation, computational and technology management (ICACTM), pp 262–267. <https://doi.org/10.1109/ICACTM.2019.8776709>
7. Pham NT, Lee J, Kwon G, Park C (2019) Efficient image splicing detection algorithm based on Markov features. *Multimed Tools Appl* 78(9). Article ID 12405
8. Siddiqi MH, Asghar K, Draz U, Ali A, Alruwaili M, Alhwaiti Y, Alanazi S, Kamruzzaman MM (2021) Image splicing-based forgery detection using discrete wavelet transform and edge weighted local binary patterns. *Secur Commun Netw*. <https://www.hindawi.com/journals/scn/2021/4270776/>. Accessed 11 Apr 2022
9. Sreelakshmy IJ, Anver J (2016) An improved method for copy-move forgery detection in digital forensic. In: 2016 online international conference on green engineering and technologies (IC-GET), pp 1–4. <https://doi.org/10.1109/GET.2016.7916684>
10. Abdalla Y, Iqbal T, Shehata M (2019) Copy-move forgery detection and localization using a generative adversarial network and convolutional neural-network. *Information* 10:286. <https://doi.org/10.3390/info10090286>
11. Khan UA, Kaloi MA, Shaikh ZA, Arain AA (2018) A hybrid technique for copy-move image forgery detection. In: 2018 3rd international conference on computer and communication systems (ICCCS), pp 212–216. <https://doi.org/10.1109/CCOMS.2018.8463337>
12. Alahmadi AA, Hussain M, Aboalsamh H, Muhammad G, Bebis G (2013) Splicing image forgery detection based on DCT and local binary pattern. In: 2013 IEEE global conference on signal and information processing, pp 253–256. <https://doi.org/10.1109/GlobalSIP.2013.6736863>
13. Dixit R, Naskar R, Sahoo A (2017) Copy-move forgery detection exploiting statistical image features. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET), pp 2277–2281. <https://doi.org/10.1109/WiSPNET.2017.8300165>
14. Meena KB, Tyagi V (2021) A deep learning based method for image splicing detection. *J Phys Conf Ser* 1714:012038. <https://doi.org/10.1088/1742-6596/1714/1/012038>
15. Moussa AM (2015) A fast and accurate algorithm for copy-move forgery detection. In: 2015 tenth international conference on computer engineering & systems (ICCES), pp 281–285. <https://doi.org/10.1109/ICCES.2015.7393060>

Two Stage Deep Learning Channel Estimation Scheme for Massive MIMO Systems



B. Pradheep T. Rajan and N. Muthukumaran

Abstract In the proposed method, we are concentrating on channel estimation for the massive MIMO systems using Deep Learning. In the existing model, we build up the channel estimation scheme for the case where the number of transmit antennas is larger than the pilot length. In the proposed plan, we are having two phase estimation processes: (i) DL-based pilot-Assisted channel estimation (DLPACE) and (ii) DL-based data-Assisted iterative channel estimation (DLDAICE). In the initial phase, the channel estimator and pilot itself are combined and intended by using Neural Network with Deep Learning (NND) and Neural Network with Two layer (NNT). In the second phase, we can use another NND in iterative mode to make further improvement in the precision of channel estimation. From the result, it reveals that the proposed scheme has delivered improved output than the conventional scheme.

Keywords NNT · NND · Channel estimation · Massive MIMO system · Deep learning

1 Introduction

Multiple input and multiple-output (MIMO) having massive measure antenna collections that's speculated to referred to as big MIMO is one of the maximum encouraging strategies in-order to boom the information fee and to maintain the excessive conversation reliability for destiny the Wi-Fi structures [1, 2]. In this form of big MIMO machine, a massive measure antenna collection is located at the bottom station (BS) to distribute a significant antenna improvement, and it is especially relies upon at the accuracy of the channel estimation. Thus, finding the precise channel estimation could be significant thing to make certain such dedications of the big MIMO strategies. The channel estimation troubles for the big MIMO structures which is typically primarily based totally at the linear minimal imply rectangular error (LMMSE) method, [3–5]. The fashionable statement which includes [3–5] is that the pilot period

B. P. T. Rajan (✉) · N. Muthukumaran
Department of ECE, Francis Xavier Engineering College, Tirunelveli 627003, India
e-mail: pradheeptrajan@gmail.com

L_s that's identical to that or more than the range of transmit antennas N_t , that is mostly a large quantity within the downlink of big MIMO. Neglecting this statement of $L_s \geq N_t$, overall act of the channel estimator is drastically weakened, as tested in [6] for the conventional channel estimator. For the big MIMO machine nevertheless, it's miles tough to defend the knowledge of $L_s \geq N_t$ for three major reasons. The primary motive is to confirm that the $L_s \geq N_t$, huge quantity of time useful resource ought to be used to check the transmission, which ends up in reduced useful resource for the transmission of the information, which ends up in very small spectral efficiency [7–9]. Secondly, the complexity in the computational wished for channel estimation will boom as increasing in s . And Finally making sure $L_s \geq N_t$ won't be capable of occur, the motive in the back of the non-incidence is the L_s cannot be large than the channel coherence interval, that's commonly out of control [10, 11]. It could be very crucial to broaden a channel estimation scheme within the big MIMO machine for $L_s < N_t$ but it's miles a totally tough trouble for growing an powerful channel estimator for $L_s < N_t$ due to the fact of the subsequent reasons. Let us anticipate $L_s \geq N_t$, S is the pilot matrix that's certainly calculated where the row vectors have been orthogonal, that is required to eradicate the interference with the various pilot sequences that is communicated from one of a kind antenna [12–14]. At $L_s < N_t$, it's miles not likely to layout the pilot matrix S within the above referred to way. (1) This approach that the trouble regarded within the channel estimator must be collectively solved with the trouble of manipulative the pilot S , which seems to be very tough method due to the fact the excessive appearing shape of S isn't always precise while $L_s < N_t$. Additional trouble is that, while $L_s < N_t$, here may be no assurance which the LMMSE is an top notch channel estimator [15, 16]. Next in this approach that within the channel estimator layout, one ought to now no longer certainly limit the channel estimator maintained as linear, if you want to mean that the trouble handy is a nonlinear optimization trouble. Overall, the impartial issues of the pilot layout and the estimation of channel are certainly very tough, now no longer to say their joint optimization issues, wherein the truth, it ought to be resolved on the end. In above, we've mentioned about the demanding situations to be keep away from and to set up a totally powerful mechanism of estimating the channel for the massive MIMO system while at $L_s < N_t$, we commonly yield a one of a kind method through advancing the deep learning primarily constructed totally on channel estimation scheme with two degree. At the primary degree of channel estimation, we're focusing on lowering the imply rectangular error (MSE) through assemble each the (NNT) Neural Network with Two Layer and (NND) Neural Network with Deep Learning [17, 18]. At 2D degree, in addition development within the channel estimation overall performance may be made through adopting an repeated channel estimation method through building every other NND as an extra channel estimator, wherein information detection and channel estimation are achieved again and again to refine the overall performance for channel estimation. Thus the consequences demonstrates that the given proposed channel method will significantly exceeds the modern current method.

2 Channel Estimation Stages

This section represents the explanation for the two-degree channel estimation technique, and this is contained on this paper pilot-Assisted channel estimation (PACE) technique and data-Assisted iterative channel estimation (DAICE) technique. We best recognition at the downlink component withinside the big MIMO system, wherein the transmitter segment is about with N_t antennas and N_r antennas receiver. The channel which used between the transmitter section and the receiver section is mentioned by $H \in CN_r \times N_t$. In this paper, the channel which is not reciprocal due to the fact we keep in mind the frequency department multiplexing (FDD), and additionally, the channel is dependable upon the of \tilde{L} time slots with coherence block due to the block fading channel model.

2.1 Pilot-Assisted Channel Estimation (PACE)

At the preliminary, the coherence fading block which is present inner to L_s time slots and the pilot $S \in CN_r \times L_s$ which begins off evolved from sender to the receiver phase and the pilot sign that is acquired is noted as $Y_s \in CN_r \times L_s$ is given by

$$Y_s = HS + Z_s \quad (1)$$

where $Z_s \in CN_r \times L_s$ denote the noise matrix. The MSE values of channel estimation may be minimized and the channel H is envisioned via way of means of the usage of the information of pilot S and the obtained pilot sign Y_s . The illustration of trouble in channel estimation in a handy shape, we want to say the matrix in vector shape which is $Y_s = \text{Vec}(Y_s) \in C^{N_r L_s * 1}$. Updated received signal Eq. (1) mentioned as,

$$Y_s = (S^T \otimes IN_r)h + Z_s \quad (2)$$

By using the obtained sign Y_s , the information of the channel estimator with S pilot is denoted via way of means of $F(\cdot)$ is used to guess the channel h . The problem in this approach of channel estimator which will be joint problem of the channel estimator $F(\cdot)$ and pilot S which can be written as,

$$\min_{s, F(\cdot)} E \left[\left\| h - \hat{h}_s \right\|^2 \right] \quad (3)$$

2.2 Data-Assisted Iterative Channel Estimation (DAICE)

This is the next phase channel estimation process in which the pilot S is transmitted, after the time interval L_d , the data $X \in \mathbb{C}^{N_t} \times L_d$ are transmitted to the receiving side, wherein $L_d = \bar{L} - L_s$. The obtained signal $Y_x \in \mathbb{C}^{N_r} \times L_d$ is given by

$$Y_x = HX + Z_x \quad (4)$$

in which $Z_x \in \mathbb{C}^{N_r} \times L_d$ represented noise matrix. From the previous stage using the PACE, the \hat{h}_s was obtained. From the first stage of X , we can estimated and it is termed as \hat{X} . Now in the second stage, we just take \hat{X} as the pilot signal and then we are going to estimate channel h and it is denoted as \hat{h}_X . The same procedure is repeated in additional to improve the performance of data detection and channel estimation until K th iteration.

3 Neural Network Channel Estimation Using Deep Learning

Here we will progress a neural network using DL-based channel estimation scheme. This channel estimation scheme using Deep Learning structure consists of two stages which are represented as DLPACE and DLDAICE.

3.1 DL-Based First Stage Pilot Assisted Channel Estimation (DLPACE):

To enterprise a pilot S , assume a NNT as a pilot designer, the following consideration are taken into account: The signal component $(ST \otimes IN_r)h$ in (2) that will remove the noise source Z_S can be modeled exactly in the NNT output only when h is the input of the NNT, and the weight matrix of the NNT is given by $(ST \otimes IN_r)$ with activation functions which is unity and all zero biases. After the whole accepted received signal Y_S in (2) can be mentioned by the summation of the output NNT and the noise function Z_S . In the first stage scheme, a NND acts as channel estimator which is nonlinear $F(\cdot)$ and is represented as NND1. The vector polarization and weight matrix are present in the m -th hidden layer of NND1 are represented by W_m and a_m for $m = 1, 2, \dots, M_1$, respectively.

Let us have in mind that there are M_1 hidden layers, then the predicted channel \hat{h}_s may be preciously mentioned as

$$\hat{h}_s = W_0 \varphi_{M_1}(W_{M_1} \varphi_{M_1-1}(\dots \varphi_1(W_1 Y_S + a_1) \dots) + a_{M_1}) + a_0,$$

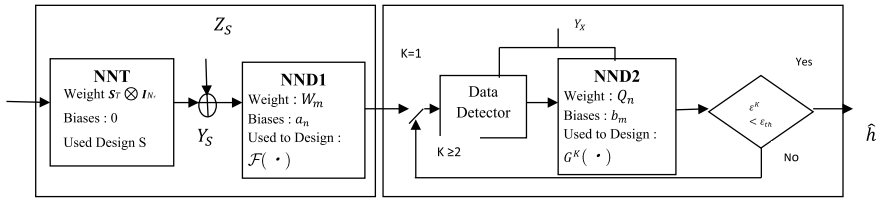


Fig. 1 Stage 1—DL-based pilot-Assisted channel estimation and Stage 2—DL-based data-Assisted iterative channel estimation

in which $\varphi_m(\cdot)$ indicates the m th hidden layer activation capabilities on the given nodes. Here we practice the rectified linear unit (ReLU) because the characteristics activation on the nodes of all hidden layers. On the finishing touch of NNT and the NND1 training, the pilot S in (P1) may be acquired with the aid of using clearly analyzing the burden matrix ($ST \otimes IN_r$) of the NNT.

3.2 DL-Based Second Stage Data-Assisted Iterative Channel Estimation (DLDAICE)

As in addition to enhance the performance of the channel estimation, we are in need to add another NND, that is signified with the aid of using NND2, as a channel estimator withinside the second stage. Likewise mentioned in (4), $\hat{h}_X(K)$, the channel which is anticipated with the aid of using the usage of the acquired information sign Y_X and the estimation of information $\hat{X}(K)$, that is anticipated with the aid of using the usage of the LMMSE detector. Then the mixed vector of Y_X and $\text{vec}(\hat{X}(K))$ are the entered onto the NND2 and the result of the NND2 which is anticipated channel $\hat{h}_X(K)$ as provided in Fig. 1. The unfairness vector and the weight matrix withinside the m th hidden layer of the NND2 are denoted with the aid of using $Q_M(K)$ and $b_M(K)$ for $m = 1, 2, \dots, M_2$, respectively. There are M_2 hidden layers, the anticipated channel $\hat{h}_X(K)$ may be scientifically represented as,

$$\hat{h}_X^{(K)} = Q_0^{(K)} \varphi_{M_2} \left(Q_{M_2}^{(K)} \varphi_{M_2-1} \left(\dots \varphi_1 \left(Q_1^{(K)} \left[Y_X^T, \text{Vec}(\hat{X}^{(K)})^T \right]^T + b_1^{(K)} \right) \dots \right) + b_{M_2}^{(K)} \right) + b_0^{(K)}$$

3.3 Proposed Scheme Training Method

In the preceding stage, knowledge of NNT and NND1 for combined layout of the pilot S and the channel estimator $F(\cdot)$ has been made in which the channel estimation of MSE is decreased as withinside the P1 (Problem). Then the characteristic loss is about to be the pattern suggest of squared mistakes as given as,

$$J_s = \frac{1}{|H|} \Sigma \left[\left\| h - \hat{h}_s \right\|^2 \right] \quad (5)$$

where H indicates the channel samples sets created for exercise purpose. Stochastic gradient descent (SGD) method is applicable to renew all the biases of the NND1 and weights which is given as $\theta_s \leftarrow \theta_s - \alpha \nabla \theta_s J_s$.

In the following stage, NND2 is educated for manipulating the channel estimator $G^{(K)}(\cdot)$ such that the channel estimation by MSE may be decreased as mentioned in the P2 (problem). i.e., the loss feature is ready to imply the pattern of squared mistakes on the k th iteration is mentioned below:

$$J_x^{(K)} = \frac{1}{|H|} \Sigma \left[\left\| h - \hat{h}_x^{(K)} \right\|^2 \right] \quad (6)$$

The SGD approach is used to replace all of the weights and biases of the NND2 as follows $\theta_x \leftarrow \theta_x - \alpha \nabla \theta_x J_x$ in which θ_x represents the biases and set of weights of NND2. The second-degree technique is typically repeated one and continually converges due to the fact the SGD approach in no way enlarges the price feature of (7). It is for the reason that the given set of rules converges inside 5 iterations. Computational complexity is a whole lot much less issue in offline gaining knowledge of due to the fact time isn't strictly restricted. When the advanced system is skilled, the pilot S, DL-primarily based totally pilot-Assisted channel estimator $F(\cdot)$, and DL-primarily based totally data-Assisted channel estimator $G^{(K)}(\cdot)$ are directly received through NNT, weight matrix, weights and biases of NND1, and people of NND2, respectively. In the case of on-line gaining knowledge of the channel estimation in case of computational complexity, and it may be given as follow; $O((M_2 - 1)\eta\lambda_2^2)$, in which η shows the range of repetitions and λ_i shows the range of nodes in every NND-I hidden layers.

4 Simulation Results

The simulating results for the massive MIMO structure may be bear in mind with the pilot strength constraint to be for the reason that $E_s = L_s T P_s$ and additionally with $N_r = 16$, $N_t = 64$, and $\bar{L} = 64$. In which T citing the time slot period and P_s citing the transmit power. At the simulating time can set $T = 1$ and $P_s = 1$. In this current method, set $\alpha = 0.001$, $\lambda_1 = 2N_r L_s$, $M_1 = M_2 = 5$, and $\lambda_2 = 2(N_t + N_r) L_d$. The signal to noise ratio (SNR) that represented as $\gamma = \frac{1}{\sigma^2}$, in which $\sigma^2 = \sigma_S^2 = \sigma_x^2$. There are over 25 epochs with 110 noise samples and 110 channel samples and additionally for overall performance assessment makes uses of 104 samples apart from training samples. In this regard, we are evaluating the proposed method with the existing scheme to the massive MIMO channel.

From the Fig. 2, we are able to look at the MSE of channel estimation vs SNR is proven for $L_s = 19$. By using the check channel samples on this proposed estimation

of channel scheme, can able to receive the channel estimation of MSE. From this output of the complete SNR range, we are able to visualize that the proposed approach produces an improved overall performance than the traditional scheme. This proves that the proposed scheme, the usage of the pilot, and channel estimator was well implemented; additionally, the overall performance of the channel estimation is step by step more suitable with the aid of the DLCE.

In Fig. 3, we can see that the lower limit of the capacity is plotted against the SNR with $L_s = 20$. The lower limit of capacity and this lower limit is very close to the exact capacity mentioned in [6]. We can clearly observed that the current estimation of channel method achieve better in terms of C_{Low} compared to the existing scheme for the whole SNR range.

Fig. 2 MSE comparison for the channel estimation schemes

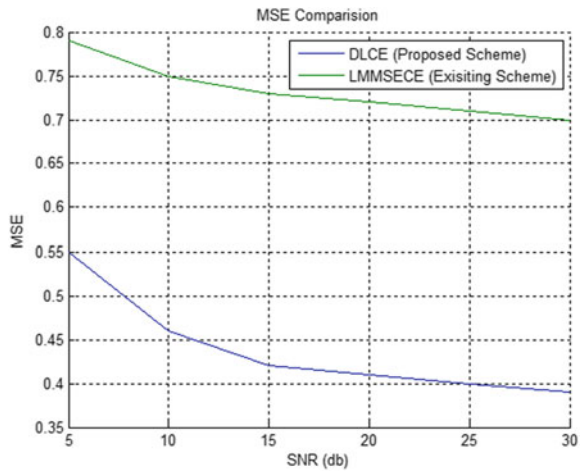
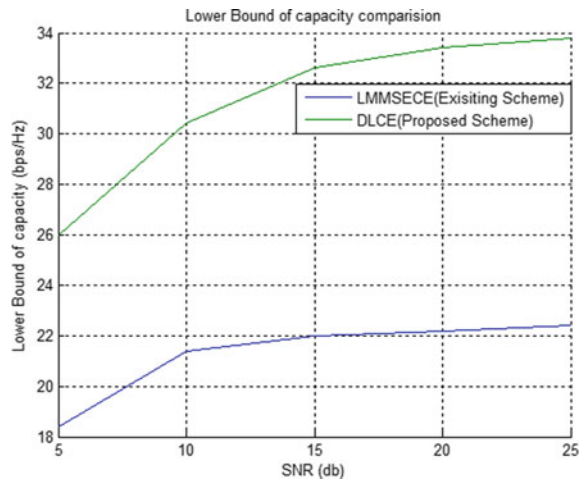


Fig. 3 Comparison of lower bound capacity with SNR



From Fig. 4, MSE is measured as opposed to L_s , in very huge range of L_s values (for e.g., $L_s > 50$). When as compared to the proposed estimation of channel method, the Mean Square Error value of the traditional estimation of channel method is lower. However, while we examine the L_s and L_d , if the L_s data could be very huge, then the data transmission time L_d turns to be very small, which will result to smaller capacity which is represented in Fig. 5.

From Fig. 5, it will plotted between C_{Low} and L_s . When $L_s = 19$, can able to visualize the expansion of lower bound of capacity. Since for each of the channel estimation method, it could be observed that the performance of data rate degrades one of the two condition if L_s is too short (because of an excessive amount of error in channel estimation), else if L_s is too long (because of too few remaining resources

Fig. 4 Comparison of MSE with pilot length

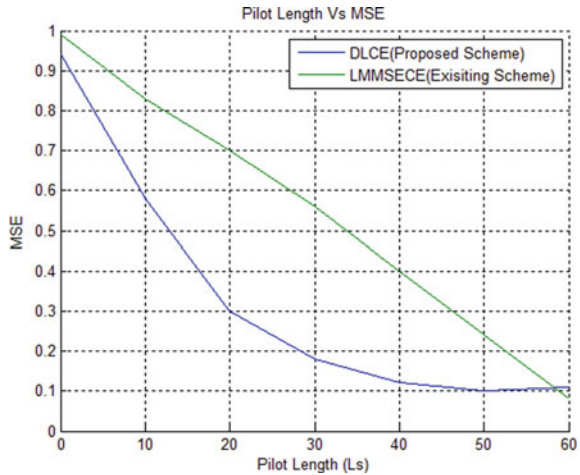
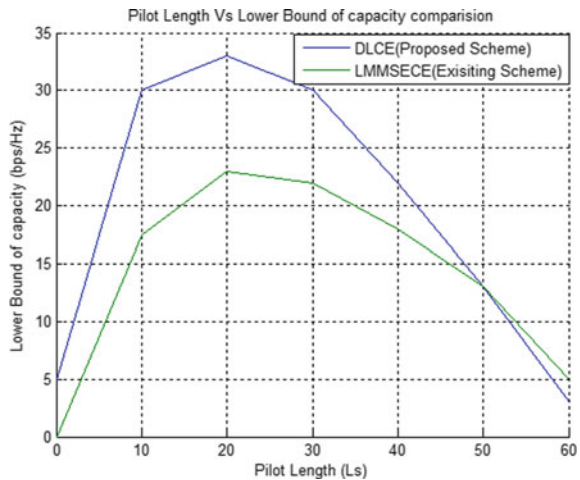


Fig. 5 Lower bound capacity with pilot length comparison



available for data transmission). The complexity of computational for the existing scheme, which is able to be slightly beyond the complexity of the computational for the proposed scheme once the parameters from the simulations are considered.

5 Conclusion and Future Scope

From all, the parameter is considered with the simulation outputs and we can have a clear cut conclusion that the proposed method which is channel estimation-based Deep Learning scheme which is represented for the massive MIMO system which provides an improved performance when compared to the traditional channel estimation scheme. In the future scope, we can add hybrid precoding design with this channel estimation scheme to ease the complexity for the MM wave massive MIMO system.





References

1. Marzetta TL (2010) Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans Wirel Commun* 9(11):3590–3600
2. Xie H et al (2016) A full-space spectrum-sharing strategy for massive MIMO cognitive Radio systems. *IEEE J Sel Areas Commun* 34(10):2537–2549
3. Ma J, Ping L (2014) Data-aided channel estimation in large antenna systems. *IEEE Trans Signal Process* 62(12):3111–3124
4. Li Y et al (2017) Channel estimation and performance analysis of one-bit massive MIMO systems. *IEEE Trans Signal Process* 65(15):4075–4089
5. Huang C et al (2019) Deep learning for UL/DL channel calibration in generic massive MIMO systems. [arXiv:1903.02875](https://arxiv.org/abs/1903.02875)
6. Hassibi B, Hochwald BM (2003) How much training is needed in multiple-antenna wireless links? *IEEE Trans Inf Theor* 49(4):951–963
7. Kay SM (1993) *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Englewood Cliffs, NJ, USA
8. Devaraj BA, Aruna T, Muthukumar N, Roobert A (2022) Adaptive cluster-based heuristic approach in cognitive radio networks for 5G applications. *Trans Emerg Telecommun Technol* 33(1):1–13
9. Suresh Chinnathampy M, Aruna T, Muthukumar N (2021) Design and fabrication of micro strip patch antenna for cognitive radio applications. *Wirel Personal Commun* 121(3):1577–1592
10. Buzzi S, Lops M, Sardellitti S (2004) Performance of iterative data detection and channel estimation for single-antenna and multiple antennas wireless communications. *IEEE Trans Veh Technol* 53(4):1085–1104
11. Hornik K et al (1989) Multilayer feed forward networks are universal approximators. *Neural Netw* 2(5):359–366
12. LeCun Y et al (2015) Deep learning. *Nature* 521(7553):436–444
13. Wen C-K et al (2018) Deep learning for massive MIMO CSI feedback. *IEEE Wirel Commun Lett* 7(5):748–751
14. Rajan BPT, Madhan KK (2013) A survey on channel estimation schemes for MIMO systems. *Int J Adv Res Electron Commun Eng* 2(1)

15. Rajan BPT, Madhan KK (2013) Kalman channel estimation by soft decision feedback for MIMO system. In: International conference on engineering and technology, pp 243
16. Pondhiyya JPJ, Rajan BPT (2013) A profound survey on various detection schemes for ultra wideband systems. *Int J Adv Res Electron Commun Eng (IJARECE)* 2(4)
17. Rajan BPT (2020) Brain tumor classification from cross modality MRI images using CNN. *Int J Adv Sci Technol* 29(6)
18. Cristy EE, Rajavel SE, Rajan BPT (2016) Energy efficient collaborative spectrum sensing in cognitive radio networks. *Glob Res Develop J Eng* 2(1)

Gamification as a Tool in Psychomotor Development: Case of Study Elementary School



Carlos Fabián Martínez Vásquez , Genesis Dayana Pinto Almeida , Evelyn Paulina Rovalino Ortega , and José David Sillagana Torres 

Abstract The following research work deals with the Integratec game and psychomotor development, with the purpose of helping in the progress of skills that have been affected during the pandemic, those skills can be evidenced in the educational, social and family field. The study is focused on verifying the implementation of games in the educational environment to build up motor, affective and cognitive areas in students, and this implements the Integratec platform to interact and improve the learning abilities of schoolchildren. The type of methodology is experimental and exploratory and analyzing causes and effects, and the approach is qualitative and quantitative; qualitative, because qualities, behaviors and conducts were analyzed; quantitative, because data were obtained, tabulated, processed and interpreted. As for the collection of information, two techniques were applied, the first one, the survey and the second one, the structured observation together with their instruments, questionnaire of questions and observation form, directed to teachers and students of the third year of General Basic Education. The population that was worked on belongs to the “CEC” Educational Unit with 3 teachers and 30 students of the third year of General Basic Education. The Chi-square parameter was used to test the hypothesis. According to the experimentation, it is evident that the teachers apply games that allow the development of cognitive, affective and motor areas of the students; however, some students have difficulties at the time of carrying out activities and games implemented by their teachers.

Keywords Games · Psychomotor development · Integratec platform

C. F. M. Vásquez
MecanicS.A./Ambato, Ambato, Ecuador

G. D. P. Almeida (✉)
High School “Nuevo Mundo”, Ambato, Ecuador
e-mail: pintodayana72@gmail.com

E. P. R. Ortega
High School “Madre Gertrudis”, Ambato, Ecuador

J. D. S. Torres
High School “Suizo”, Ambato, Ecuador

1 Introduction

Zamora [1] In your research “Use of the game and the strengthening of motor skills,” highlights that the game focused on learning allow improving the way students learn, with the purpose of strengthening skills, such as problem solving, improving communication, good organization, teamwork, creativity and innovation. In addition, currently, educational demands are becoming stronger, since educational institutions require teachers who master technological tools with great efficiency and effectiveness. For this reason, it is important that teachers are trained and investigate new technological tools that are used for learning and psychomotor development. Since, due to the “COVID 19” pandemic, psychomotor development is affected in many children and adolescents, due to the fact that confinement during the health emergency prevented them from practicing activities and games that help in the development of fine and gross motor skills.

The problem for which it was decided to investigate this topic arises because in Ecuadorian education, it has been shown that the majority of Ecuadorian teachers have difficulties in the use and implementation of technological tools that are related to gamification and help psychomotor development, it is for that reason, it has implemented the Integratec digital page to help in the development of emotional, cognitive and psychomotor aspects of students. It is like this that, the teachers must direct the game with the contents to teach, it does not have to be a totally free game, since the teachers must understand and analyze in depth the activities to be implemented in the classroom for learning to be useful, in order to have a good performance in attitudes and strengthening the skills of schoolchildren [2].

In addition, children and young people can easily access an endless number of online games, which in many cases include violent content and inappropriate language, that with carries to physical and verbal violence in the individual, which can later be reflected in the family, school or social environment. For this reason, the teachers in the educational field must investigate, select and adapt games and activities that generate meaningful learning and do not generate undesirable behaviors in their students [4].

Gamification as a tool in psychomotor development is defined as the technique that favors meeting the proposed objectives through the implementation of innovative games that can be played in any space, in order to get students to observe, explore, participate and acquire significant learning through the use of their body parts “fine and gross motor skills” developing skills that will be useful in the environment in which they find themselves. The game is important in the educational field for the development of skills of children and young people, since they are the ones who play the most during their growth, which means that through educational processes, the subjects are becoming aware of the real [4].

According to Pinto [2] who mentions that: The Integratec game is found on the digital page strikingly, whose software has a set of instructions redirected to be practical in a virtual or face-to-face way, thus integrating documents, videos, games of brain gymnastics, mental, memory, riddles and even stories. The page provides

access to other pages that allow interaction through games for the development of cognitive and motor skills. Thus, the Integratec game must be expanded at the school level and be discussed among teachers about its advantages and disadvantages, in order to maximize its potential and be used to teach most content.

2 Methodology

The ADDIE methodology was used for the development of the research work. “Analysis, Design, Development, Implementation and Evaluation (Fig. 1).”

Memory games: In the main menu, there is a button called memory game, directing to a web page “Arbol ABC,” where there are memory activities that allow interaction between teacher and student, as well as an interactive video to implement it virtually or in person (Fig. 2).

Mental games: Inside this option, there are activities that stimulate the different types of memory, either to short, medium or long term (Fig. 3).

Stories: Inside this section, there are activities that help the development of imagination and creativity, in order to get into real events or a fictional world to generate good attitudes, values and ethical concepts in students (Fig. 4).

Riddles: This section presents contents that allow students to analyze, reflect and give an answer that may or may not be correct (Fig. 5).

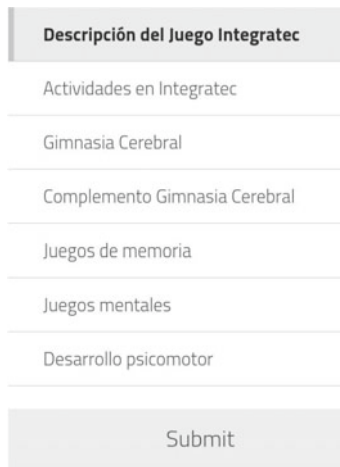


Fig. 1 Integratec platform menu



Fig. 2 Memory games

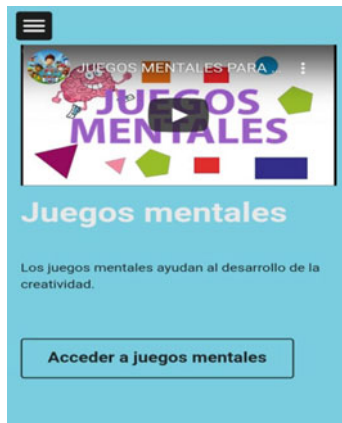


Fig. 3 Mental games

Brain gymnastics: Inside this option, there are documents and videos that provide indications to implement this activity in class, since it allows the brain to be activated, and so the student learns in an entertaining way (Fig. 6).

Videos: The Integratec platform consists of videos compiled from the YouTube platform showing brain gymnastics activities, stories, riddles, mental and memory games that allow the integral development of individuals (Fig. 7).

Documents: The Integratec platform contains files, such as; stories, brain gymnastics and riddles that can be downloaded or viewed from the web browser, with the purpose of carrying out activities that motivate children and young people in the classroom (Fig. 8).



Fig. 4 Stories

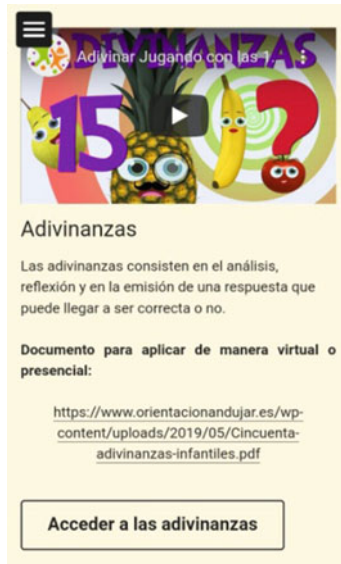


Fig. 5 Accessing the riddles

Evaluation: The application of the observation form was considered, with the purpose of verifying the Integratec platform was valid or rejected within the study population, with the objective of investigating the impact of the contents it provides in the educational field and how individuals perceive such usefulness (Fig. 9).



Fig. 6 Brain gymnastics

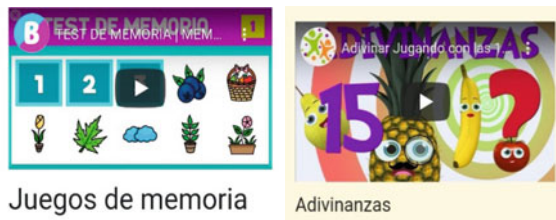


Fig. 7 Videos

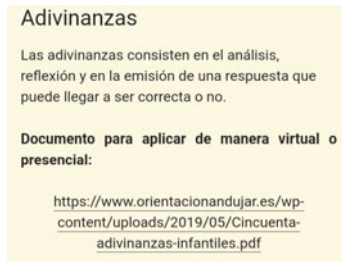


Fig. 8 Documents

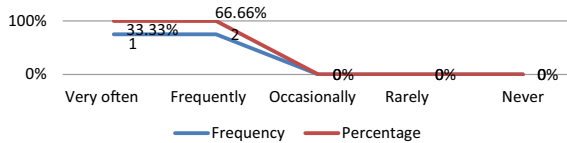


Fig. 9 Implementation of games. By Martínez C. (2022)

3 Results

1. How often do you implement games in class?

See Table 1.

Analysis and Discussion

Out of a total of 3 teachers surveyed, which represents 100%, 2 (67%) affirm that they implement games in class very frequently, while 1 (33%) indicates that they implement them frequently. This makes it possible to establish that very frequently and frequently the teachers of the third year of EGB of the “CEC” Educational Unit implement games in class, which is a positive aspect, since through games the children can learn, discover and strengthen new skills.

2. How often do you use innovative strategies to encourage play and strengthen psychomotor development?

See Table 2 and Fig. 10.

Analysis and Discussion

Of a total of 3 teachers surveyed, 67% says that they frequently use innovative strategies to promote the play and strengthen psychomotor development, while 33% indicate that they use them very frequently. This allows us to establish that teachers very frequently and frequently use innovative strategies to promote the play and strengthen psychomotor development, so demonstrating their concern for promote the development of their students in terms of their learning and development of different types of skills.

Table 1 Implementation of games frequency percentage

	Frequency	Percentage (%)
Very often	2	67
Frequently	1	33
Occasionally	0	0
Rarely	0	0
Never	0	0
Total	3	100

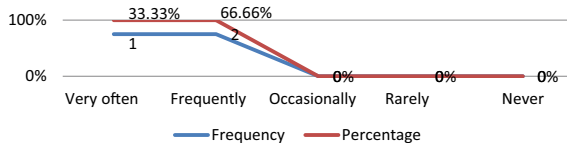


Fig. 10 Strategies to promote the play. By Martínez (2022)

Table 2 Strategies to promote the play

	Frequency	Percentage (%)
Very often	1	33,33
Frequently	2	66,66
Occasionally	0	0
Rarely	0	0
Never	0	0
Total	3	100

3. **Gives his best during the game** (Table 3 and Fig. 11).

Analysis and Discussion

Of a total of 30 children observed, representing 100%, 14 students (46%) almost always give their best during the game; 8 (27%) always give their best; 5 (17%) sometimes; 3 (10%) rarely; and 0% (0%) never. This makes it possible to establish that most of the children contribute their best during the game implemented by the teacher, while a minority sometimes and rarely contributes their best, which means that most of the students like to participate in activities implemented by the teacher.

4. **Easily integrated into the games implemented by the teacher** (Table 4 and Fig. 12).

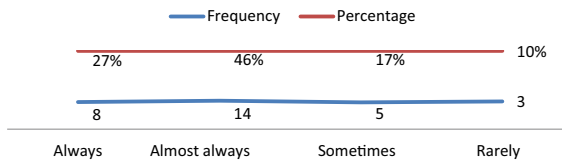


Fig. 11 Gives the best in the game. *By Martínez C. (2022)*

Table 3 Gives the best in the game

	Frequency	Percentage (%)
Always	8	27
Almost always	14	46
Sometimes	5	17
Rarely	3	10
Never	0	0
Total	30	100

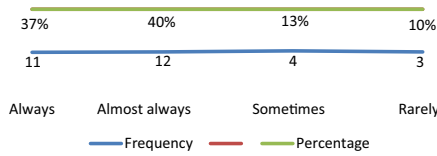


Fig. 12 Easily integrated to the games. *By Martínez C. (2022)*

Table 4 Easily integrated to the games

	Frequency	Percentage (%)
Always	11	37
Almost always	12	40
Sometimes	4	13
Rarely	3	10
Never	0	0
Total	30	100

Analysis and Discussion

From a total of 30 boys and girls observed, representing 100%, it is obtained that 12 students, representing 40%, almost always integrate easily to the games implemented by the teacher; 11, representing 37%, always integrate easily; 4, representing 13%, sometimes; 3, representing 10%, rarely; and 0%, never. This allows us to establish that most of the students integrate easily in the games implemented by the teacher; while there are few students who have difficulty in integrating. This may be due to the fact that the children did not pay adequate attention, or there was a distractor, causing them not to understand the activity implemented by the teacher.

Verification of Hypotheses

For this research, it was essential to verify the hypothesis by means of the Chi-square statistic, based on two questions representative of the research study; the questions are based on the object of study and the results shown by the experiment. Item 2 “gives the best of himself during the game” identified with the Independent Variable (Integratec Game) and Item 6 “executes activities in a coordinated manner and with adequate control of strength and muscle tone such as throwing, catching and kicking balls” represented with the Dependent Variable (Psychomotor Development) were taken into consideration.

Data collection and statistical calculation
 Frequency of observed and expected values

Contingency table. Gives the best of himself during the game. * Executes activities in a coordinated manner and with adequate control of strength and muscle tone such as throwing, catching and kicking balls

(continued)

(continued)

			Executes activities in a coordinated manner and with adequate control of strength and muscle tone such as throwing, catching and kicking balls				Total
			Rarely	Sometimes	Usually	Always	
Gives the best of himself during the game	Rarely	Recount	2	1	0	0	3
		Frequency spread	0.2	0.4	0.6	1.8	3.0
	Sometimes	Recount	0	3	2	0	5
		Frequency spread	0.3	0.7	1.0	3.0	5.0
	Usually	Recount	0	0	4	10	14
		Frequency spread	0.9	1.9	2.8	8.4	14.0
	Always	Recount	0	0	0	8	8
		Expected frequency	0.5	1.1	1.6	4.8	8.0
Total		Recount	2	4	6	18	30
		Expected frequency	2.0	4.0	6.0	18.0	30.0

4 Conclusions

The play is an essential activity that should be implemented in the classroom, since due to several factors; schoolchildren do not develop their cognitive, motor and affective skills satisfactorily, making it impossible in many cases to acquire knowledge.

In general, the students who have not developed basic skills have difficulties when participating in activities and games carried out by the teacher. Most of the students in the two parallel years of the third year of General Basic Education have developed the psychomotor aspect; however, there are students who still need to strengthen these skills, which are importance for the child to continue advancing in his or her integral development.

The technology in the contemporary world has made it possible to develop contents in an entertaining way, so teachers in the educational field should direct games and activities that make it possible to strengthen abilities and skills, so that schoolchildren awaken their curiosity, liveliness and become self-confident.

References

1. Zamora C (2020) La intervención del docente en una sesión de psicomotricidad
2. Pinto V (2020) El uso del juego Integratec y el fortalecimiento de habilidades motrices. Obtenido de <https://sites.google.com/view/juego-integratec/p%C3%A1gina-principal>
3. Agila Pitizaca JE, Guamán Guacho NP (2020) El juego en el desarrollo psicomotor en niños y niñas de 4 a 5 años en la Unidad Educativa “Pedro Luis Calero” en el periodo lectivo 2019–2020
3. Lopez P (2020) Los juegos en línea y su relación con la interacción social entre pares. Un análisis de caso en niños de cuarto año de educación básica. dspace

The Implementation of Octave-S in the Risk Management Measurement of Information System in the Department of Communication, Informatics, and Statistics



Oggsa Sukardi, Ford Lumban Gaol, and Tokuro Matsuo

Abstract An information system is an important part in an organization to support business processes and to achieve its vision and mission. The information system nowadays has been one of the assets that ought to be protected by an organization. The Department of Communication, Informatics, and Statistics has the information system supporting its various operational activities. In order to protect the information system, OCTAVE-S method as the management method is used. It is expected that the OCTAVE-S method can manage the risks to the information system as the asset by specifying the possible risks and developing countermeasures and mitigation plans for each of the risks.

Keywords Information system · Risk management · OCTAVE-S

1 Introduction

The development of technology in state of institutions is increasing and thus, the necessity to acquire information is also increasing in terms of its accuracy and time efficiency [1]. To integrate the information system into the service activities, the support of proper Information Technology (IT) assets is needed [2]. Each of those assets has a role in supporting the stability of IT services in the office of The Department of Communication, Informatics, and Statistics. If one of the assets encounters an unexpected problem, the whole given IT services will be affected, and therefore,

O. Sukardi

Master of Information System Management, Binus Graduate Programs, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: oggsa.sukardi@binus.ac.id

F. L. Gaol (✉)

Computer Science Department, Binus Graduate Program—Doctor of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo

Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aait.ac.jp

the risk management of IT assets is needed. However, the conducted risk management is still passive nowadays [3].

One of the methods to conduct the risk assessment is the OCTAVE method [4]. It is expected that with this research, the IT provider in the office of The Department of Communication, Informatics, and Statistics will be more aware of the risk management importance for the IT assets and will be able to give supports to arrange strategies in order to prevent the possible threats in the information system and its infrastructure supporting The Department of Communication, Informatics, and Statistics [5].

2 Research Method

This research is conducted through several stages [6, 7]. The first stage is by conducting a preliminary study, which is finding articles, books, and other sources related to the research topic as well as learning the information about a company.

Based on the above diagram about the framework, it can be explained that the risk management processes in the information system at The Department of Communication, Informatics, and Statistics use the OCTAVE-S method with its three main phases, which are:

1. Building threat-based asset profiles
2. Identifying infrastructure vulnerabilities
3. Developing security strategies and plans.

3 Discussion

3.1 *Building Threat-Based Asset Profiles (Model of OCTAVE-S Phase 1)*

This stage aims to know the measure of the risk impacts that will be undergone by The Department of Communication, Informatics, and Statistics. To know the measure of the risk impacts, the data collection is done through the process of filling questionnaires together with the written questions to the staff of The Department of Communication, Informatics, and Statistics. The results of the questionnaires and the written questions will be examined in order to produce a formulation of the risk impacts that will be undergone by The Department of Communication, Informatics, and Statistics [8].

In the guideline of OCTAVE-S model from Alberts et al. (2005), three stages of risk impacts can be undergone by an organization. They are Low Impact which is the lowest stage of risk impacts, Medium Impact which is the intermediate stage

of risk impacts, and High Impact which is the highest stage of risk impacts in an organization [5].

In formulating the risk impacts on the aspect of productivity, the level formulation of risk impacts should refer to staff’s extra working hours with the intention to cope with the existing risks in The Department of Communication, Informatics, and Statistics [9]. Risk Impacts on the Aspect of Reputation is given in Table 1. The data about the staff’s working hours in The Department of Communication, Informatics, and Statistics are taken based on the results of the answers from the questions and questionnaires filled by The Department of Communication, Informatics, and Statistics. Risk impacts on the aspect of productivity is given in Table 2.

The next step is to identify information assets of The Department of Communication, Informatics, and Statistics. The assets of the information system are arranged in a form of Table that delineates the role of each information system owned by The Department of Communication, Informatics, and Statistics in running its service role. The information system assets of The Department of Communication, Information and Statistics is given in Table 3.

Besides the information system asset, employees as assets are also included in the OCTAVE-S process which will become one of resources to do the analysis of the existing risks. In The Department of Communication, Informatics, and Statistics, the IT staff work together to build and support the development of the information system as well as to keep the system running in a stable pace.

The next step is to do security evaluation in The Department of Communication, Informatics, and Statistics for which OCTAVE-S has 15 activities employed to identify security practices of the information system. In determining security practices of an organization, the security practice value is symbolized by color cards of which

Table 1 Risk impacts on the aspect of reputation

Reputation of Services		<i>Low impact</i>	<i>Medium impact</i>	<i>High impact</i>
	Reputation	The reputation is barely affected and needs no effort to recover the reputation	The reputation decreases and is badly impacted so that an action and cost is needed to recover the reputation	The reputation becomes considerably bad so that it is almost impossible to recover

Table 2 Risk impacts on the aspect of productivity

Productivity			
<i>Impact type</i>	<i>Low impact</i>	<i>Medium impact</i>	<i>High impact</i>
Staff’s working hours	Staff’s working hours are added 1–3 h per day in order to cope with risk impacts	Staff’s working hours are added 3–6 h per day in order to cope with risk impacts	Staff’s working hours are added 6 h per day in order to cope with risk impacts

Table 3 The information system assets of The Department of Communication, Informatics, and Statistics

Information system and applications				
No	System	Information	Application and services	Other assests
1	<i>Jaga Riau</i>	APBD (Regional Income and Expenditure Budgets) of Riau Province and the data of schools in Riau Province	<ul style="list-style-type: none"> • PHP 7 • MySQL 	
2	<i>e-office</i>	The data of letters and administrations related to DISKOMINFOTIK of Riau Province	<ul style="list-style-type: none"> • PHP • XML • SQL Server 	
3	<i>e-Absen</i>	The data of the staff's attendances in DISKOMINFOTIK of Riau Province	<ul style="list-style-type: none"> • PHP • SQL Server 	e-office
4	<i>Streaming Riau</i>	Live streaming of the existing activities in the Regional Government of Riau Province	<ul style="list-style-type: none"> • PHP • SQL Server 	

each of the cards represents the implementation level of the security practices based on qualities in The Department of Communication, Informatics, and Statistics. Based on the evaluation results of 15 security activities, it can be concluded through the following graphic:

Based on the Fig. 1, it can be concluded that most of the security practices with OCTAVE-S model conducted in The Department of Communication, Informatics, and Statistics are in the category of dissatisfaction (42 activities are in the red status) **meaning that most of the security activities have not been yet implemented well** by The Department of Communication, Informatics, and Statistics. On the other hand, the yellow status has 28 activities while the green status has 3 activities. This security evaluation data has been obtained from some respondents that are associated and have knowledge of security practices, and the data obtained is calculated using the Likert Scale method.

The next step is to determine the critical assets. Based on the guideline of OCTAVE-S model and questionnaires filled by The Department of Communication, Informatics, and Statistics, the security needs for important assets in The Department of Communication, Informatics, and Statistics are divided into three categories, which are Confidentiality, Integrity, and dan Availability. The following identification result of security needs is for critical assets of The Department of Communication, Informatics, and Statistics. Identification of the needs for e-office critical assets are given in Table 4.

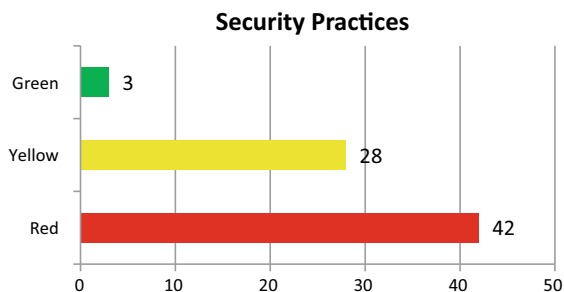


Fig. 1 Graphic of security practices in The Department of Communication, Informatics, and Statistics

Table 4 Identification of the needs for e-office critical assets

Critical assets	Selection basis	Description
<i>e-office</i>	The reason is that this system makes the staff easily perform the delivery of reports on service activities, financial statements, and letters of administrative service as well as makes the staff who serve out of town easily access information and documents so that they do not have to be in the office to do so. This kind of system is the most accessed asset by the party of The Department of Communication, Informatics, and Statistics	All the employees in Riau Province
Related assets	Security requirements	Security requirements focus
Information: Information of reports on service activities Information of letters of service	Confidentially: Information of service activities may only be accessed by the employees who have the authorization to their account	<i>Confidentially</i> <i>Availability</i>
Financial: Information of APBD (Regional Income and Expenditure Budgets)	Integrity: The stored data of reports and information have to be clear and suitable for the used necessities	
Services and Application: Database: My SQL Tools: Notepad + + , Sublime Server: Xampp	Availability: The access to this system has to be available wherever the employees are as well as easy to access by the employees so that every letter can be quickly verified	

Table 5 System of interest, access points, and system access by people

Component class	Access path	Hardware
Server	Internal network	Internal PC
	External network	Laptop
		External PC

3.2 Identifying Infrastructure Vulnerabilities (Model of OCTAVE-S Phase 2)

This stage is a process of identifying how an actor can access the important assets in The Department of Communication, Informatics, and Statistics. Based on the results of questionnaires and analyses referring to the guideline of OCTAVE-S model from Alberts, C., et al. (2005), it can be concluded that the system which is in the closest relation with the important assets of information in The Department of Communication, Informatics, and Statistics is e-office information system [10].

The following table explains about the components related to the critical assets of The Department of Communication, Informatics, and Statistics. These data are based on the results of an interview with The Department of Communication, Informatics, and Statistics [11].

Based on the Table 5, the e-office information system can be accessed by both internal and external networks. In the internal network, hardware that can access the assets is PC connected to the internal network of The Department of Communication, Informatics, and Statistics. Meanwhile in the external network, outsiders can access the assets by using an Internet network. Server, which is the class component, is used as the location of e-office information system stored and utilized as the database of this system. The server has public IP in order that the e-office system can be accessible for other services by using an Internet network [12].

3.3 Develop Security Strategies and Plans (Model OCTAVE-S Fase 3)

Based on the evaluation of threats and risks in The Department of Communication, Informatics, and Statistics, the next step is to determine mitigation plans of the risks. The risk mitigation plans are established from the result of the discussion with IT staff in The Department of Communication, Informatics, and Statistics and the one with experts in relation to the existing important assets. The further mitigation plans are as follows [13].

1. Awareness of Security and Training

In this aspect, the needed mitigation is to conduct a training related to the information security and important asset for all the staff in The Department of Communication,

Informatics, and Statistics so that the staff's awareness of security and responsibility can be well increased. Moreover, the training should be conducted periodically in order to continuously increase the knowledge of information security along with the development of information technology.

2. Security Strategies

In this aspect, the needed mitigation is to arrange the strategies for the organization by incorporating elements of the security information. The elements of the security information should be linked to the vision and mission of The Department of Communication, Informatics, and Statistics so that the arranged strategies will refer to the vision and mission consistently. The strategies are periodically re-reviewed so that the result of the review will create a new strategy which is more integrated with the security.

3. Security Policies and Regulations

In this aspect, the mitigation in The Department of Communication, Informatics, and Statistics should have clearly documented regulations and Standard Operating Procedure (SOP) in executing security actions or accesses to information. The arranged regulations can refer to the existing state constitution and be communicated with the head of department so that the regulations can be well socialized to the staff of The Department of Communication, Informatics, and Statistics.

4. Collaborative Security Management

In this aspect, The Department of Communication, Informatics, and Statistics should have policies and regulations of the use of vendor services in doing the efforts for the information security. The role of security and vendor's responsibilities ought to be determined and documented in line with the security policies of The Department of Communication, Informatics, and Statistics.

5. Management Plans for Disasters

For the mitigation activities in this aspect, The Department of Communication, Informatics, and Statistics should conduct an analysis of the operational activities related to the existing information system and important data so that the result of the analysis can be documented and monitored.

6. Security Monitoring and Auditing

In this aspect, The Department of Communication, Informatics, and Statistics should conduct the process of auditing periodically in order to ensure whether the security strategies have run in line with the established regulations as well as the vision and mission of The Department of Communication, Informatics, and Statistics. The anomalous findings are analyzed and given an action to prevent the occurrence of threats which can disturb the business processes of The Department of Communication, Informatics, and Statistics.

The activities of risk mitigation will run well if The Department of Communication, Informatics, and Statistics follows the following things [10]:

1. The mitigation process is conducted consistently.
2. The support of the head of department is available.
3. The support of IT staff who participate in the mitigation process is available.
4. The policies or standards in conducting specific things which have a high-risk value are realized.
5. The training for the staff who have weak security awareness is held.

4 Conclusion and Suggestion

4.1 Conclusion

Based on the result of the research conducted in The Department of Communication, Informatics, and Statistics, it is conclusive that:

1. The Department of Communication, Informatics, and Statistics has not yet implemented the risk management of IT assets specifically the information system.
2. After the measurement with OCTAVE-S has been conducted, the possible risks in The Department of Communication, Informatics, and Statistics are in the aspects of Security Awareness and Training, Security Strategy, Security Policies and Regulations, Collaborative Security Management, Contingency Planning/ Disaster Recovery, Physical Access Control, Monitoring and Auditing IT Security, Authentication and Authorization, Vulnerability Management, Encryption, Architecture and Design of Security, Incident Management.
3. Mitigation on the IT assets specifically the information system in The Department of Communication, Informatics, and Statistics is done by the following ways:
 - (a) Conducting periodic security training followed by the all staff as well as reviewing the security knowledge.
 - (b) Conducting documentation of the existing policies and regulations incorporating the elements of security and the existing vision and mission and then, socializing the regulations to the all staff and existing vendors.
 - (c) Consistently performing activities of auditing and security monitoring in order to ensure whether the security strategies have been well conducted in line with the established policies and regulations as well as to take an action when there are anomalous activities.
 - (d) Making and conducting socialization to formal procedures related to authorization and authentication toward rights of the access to the existing important information so that the use of information can be well handled and documented.

4.2 Suggestion

Based on the explanation of the conclusion, the proposed suggestions are as follows:

1. To protect the IT assets in The Department of Communication, Informatics, and Statistics, companies can implement the risk management of the IT assets so that the impacts of the risks can be minimized.
2. Companies can follow the mitigation process identified with the OCTAVE-S method by arranging formal procedures, conducting documentation of every management activity of IT assets, and providing training for staff and the third parties.
3. To decrease risk possibilities measured with the OCTAVE-S method in this research, The Department of Communication, Informatics, and Statistics can measure the risks with the OCTAVE-S method routinely based on the schedule determined by companies.

References

1. Undang-Undang Republik Indonesia No. 25 Tahun 2009. Pelayanan Publik
2. Darwish SZ, Gomes AM, Ahmed U (2021) Risk management strategies and impact on sustainability: the disruptive effect of Covid 19. *Acad Strateg Manag J* 20(2)
3. Dwiarto P (2020) Factors affecting the area of risk management disclosures. *Int J Econ Bus Account Res (IJEBAR)* 4(4)
4. Korshunov G, Varzhapetyan A, Petrushevskaya A, Polyakov S (2021) Assessment of target and usefulness of complex cyber-physical systems. *Int J Risk Assess Manag* 24(1)
5. Alberts C, Dorofee A, Stevens J, Woody C (2003) Introduction to the OCTAVE approach. Carnegie Mellon University, Pittsburgh, PA
6. Nugroho A (2002) Analisis dan Perancangan Sistem Informasi dengan Metodologi Berorientasi Objek, Informatika, Bandung
7. Anton SG, Nucu AEA (2020) Enterprise risk management: a literature review and agenda for future research. *J Risk Fin Manag* 13:281. <https://doi.org/10.3390/jrfm13110281>
8. Panda P (2009) The OCTAVE® approach to information security risk assessment. *ISACA J* 4:37–41. <http://www.isaca.org/Journal/Past-Issues/2009/Volume-4/Pages/The-OCTAVE-Approach-to-Information-Security-Risk-Assessment1.aspx>
9. Djojosoedarso S (2003) Prinsip-Prinsip Manajemen Risiko dan Asuransi, Edisi Revisi. Jakarta: Salemba Empat
10. Alberts C, Dorofee A, Stevens J, Woody C (2005) OCTAVE-S Implementation Guide, Version 1.0. Carnegie Mellon University, Amerika Serikat
11. Peltier TR (2000) Facilitated risk analysis process (FRAP). Auerbach Publications
12. Simister SJ (2004) Qualitative and quantitative risk management. *The Wiley guide to managing projects*, vol 3. Elsevier, Norway
13. Hildenbrand J (2016, March) Inside the different Android Versions. <http://www.androidcentral.com>

Data Protection Using Scrambling Technique



Pothuka Madhuri and E. Prabhu

Abstract Nowadays, securing information is one of the major concerns. It deals with preventing the data from unauthorized access, modification, cloning, attacking, and tampering. As a result, many cryptography algorithms such as AES, DES, RSA, and blowfish have been developed to overcome all threats and to improve security. But then, fast growth in computers may threaten these cryptography algorithms. This paper proposes two levels of security wherein data passes through two phases: one is encryption, and the second one is through scrambling. At the encryption phase, data is encrypted based on the array of elements that are considered. At the scrambling phase, the data is scrambled based on the proposed scrambling algorithm. This paper considers three performance metrics: time, space, and security.

Keywords Encryption · Decryption · Scrambler · De-scrambler · Prime number · Hardware Trojan · Verilog hardware description language · QuestaSim 10.4e

1 Introduction

Cryptography [1] is the technique of information security that deals with modifying data in such a manner that it is comprehensible to someone who is unaware of a secret. Modern cryptography is mostly based on mathematical computations. For modern computers, it is simple to perform those computations with a key, but then it is difficult enough to guess in the absence of the same key. Which means the true security of the data is determined by the encryption key employed. For example, if the algorithm uses a length of 4 bit key, the possible key combinations are 16. As a result, 16 different combinations must be explored in order to expose the information via a brute force approach.

Cryptography is categorized into two types: symmetric key algorithms and asymmetric key algorithms. Symmetric key algorithms [1] are concerned with a single key. The same key is used for both encryption and decryption. The key is sent to

P. Madhuri · E. Prabhu (✉)

Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: e_prabhu@cb.amrita.edu

those who require access to the data. Security is entirely dependent on the key used in this case. Any unintentional exposure of the key to non-authorized people results in the compromise of the entire system. The only shortcoming of the symmetric key technique is that it requires a secure key exchange, making it ideal only for safeguarding data for an individual or a small number of individuals. A few of the symmetric algorithms are: Data Encryption Standard (DES), Advanced Encryption Standard (AES), blowfish, and twofish [2].

Asymmetric key algorithms encrypt and decrypt data using a pair of keys. The keys are classified as either public or private. The private key is kept secret by the owner and is used to decrypt the encrypted data with the public key, which is freely available to everyone in the public domain. The data encrypted with the private key may be decrypted with the public key. To verify this, only the user of the private key is able to encrypt data. Because asymmetric key algorithms do not rely on a single key, they are more secure than symmetric algorithms. The key problem with asymmetric algorithms is that they are computationally expensive, which makes them difficult to use often within a process [3–6]. The three key criteria that separate a good algorithm from a bad algorithm are time, space, and security. The application in which the algorithm is utilized is critical in deciding which algorithm to employ. In this paper, we offer a method in which data passes through two phases: the first encrypts the data, and the second scrambles it.

The motivation behind implementing the scrambling technique is, all the above algorithms mentioned are using the concept of key for encryption and decryption of the data. Whereas, the scrambling is a different technique to secure the information where no key is used. As per the proposed algorithm the data changes based on the scrambler and based on the user defined prime numbers considered for encryption process. So, to reverse engineer and get the actual data information the attacker needs to go through huge number of permutations and it is even hard to go through all the iterations and come to the conclusion about the actual data.

The entirety of the paper is structured as follows: Sect. 3 introduces the proposed methodology, Sect. 4 outlines the attack on proposed scheme, Sect. 5 discuss the performance metrics, Sect. 6 explains the results and discussions, and Sect. 7 covers the conclusion.

2 Literature Survey

The encryption is a most used technology for protecting the sensitive information. The analysis of the different types of encryption algorithms are done as a part of literature survey. Let's discuss about the AES [2] algorithm first. To encrypt sensitive data, AES is used in hardware and software around the globe. Asymmetric Key Encryption is the name given to the encryption technique when distinct keys are used to both encrypt and decrypt the data. Even if the keys are distinct, they are mathematically related, making it possible to recover the plain text by decrypting the cipher text. In this system, each user is required to have a set of two dissimilar keys a, private key and a

public key. When one key is used for encryption, the other may be used to decode the cipher text and restore the original plain text. These keys are mathematically related. It demands that the private key be kept a closely guarded secret and the public key placed in a public repository. Consequently, this encryption method is also known as public key encryption. Although the user's public and private keys are related, it is computationally impossible to distinguish one from the other. This is a benefit of the plan. When Host1 has to transfer information to Host2, he gets the recipient's public key from the repository, encrypts the information, and sends it. The plain text is extracted by Host2 using his private key. The amount of processing power needed by a computer system to operate an asymmetric algorithm is greater. Coming to the DES [2], the popularity of the DES has been discovered to be somewhat declining as a result of the discovery that DES is susceptible to very strong assaults. Since DES is a block cipher, it encrypts data in blocks of 64 bits each. As a result, DES receives 64 bits of plain text as input and outputs 64 bits of cipher text. With a few minor variations, the same method and key are utilized for encryption and decryption. And about RSA [2]. The concept of RSA is based on the fact that big integers are hard to factor. The public key is made up of two numbers, one of which is the product of two enormous prime numbers. The same two prime numbers are also used to create the private key. Therefore, the private key is compromised if someone is able to factorize the huge integer. As a result, the key size completely determines how strong an encryption is, and doubling or tripling the key size significantly boosts encryption strength. RSA keys may normally be 1024 or 2048 bits long, however experts think that keys with 1024 bits could be cracked soon. But as of right now, it appears to be an impossible feat [7, 8]. Obfuscation technique and Hardware Trojan detection techniques with different methods are discussed in [9–11].

3 Proposed Methodology

According to Fig. 1, the proposed methodology is made up of four blocks: encryption, scrambler, de-scrambler, and decryption. Let's start with the block diagram before digging into the functionality of each block. First, the input data is fed into the encryption block. Where the data is encrypted in the format based on the proposed scheme outlined in Sect. 3.1. Then the encrypted data is fed to the scrambler, where the data is scrambled based on the proposed algorithm discussed in Sect. 3.2. Then to the de-scrambler and the decryption block and detailed explanation about the de-scrambler and decryption blocks functionality is discussed in Sects. 3.3 and 3.4.

3.1 Encryption

Encryption is the process of changing information or data into another format in order to prevent unauthorized access to the data. In the proposed encryption scheme,

Fig. 1 Block diagram of proposed methodology

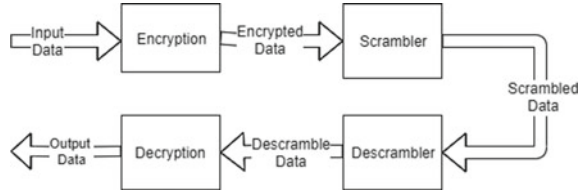


Table 1 Prime numbers consideration for encryption

	a(0)	a(1)	a(2)	a(3)	a(4)	a(5)	a(6)	a(7)
Prime Numbers	2	5	3	11	7	23	19	31
Input Data	0	1	0	0	1	1	0	1
Result	$5 \times 7 \times 23 \times 31 = (24,955)_{10} = (110_0001_0111_1011)_2$							

the data is encrypted based on the array of prime number elements selected by the user. Let’s consider an example. Suppose the input data is an 8-bit binary number as follows: $(01,001,101)_2$. Create an array with eight unique prime numbers: e.g., $a(8) = 2,5,3,11,7,23,19,31$ arranged in a specific order. Above the 8-bit binary number, a prime number must be assigned to each bit, as shown in Table 1. Then multiply the prime numbers that are positioned bit 1 that results in $(110000101111011)_2$. The resultant data is the encrypted data. If you compare the encrypted data and actual data, the encrypted data is longer than the actual data. For different data inputs, the encrypted data varies, and the length of the results will be different.

So, here, to standardize the block size, we consider the maximum possible value, i.e., the result of the multiplication of all array elements. $Max = (2 \times 5 \times 3 \times 11 \times 7 \times 23 \times 19 \times 31) = (31,293,570)_{10}$. If we convert the decimal to binary, it results in 25 bit data $(1110111011000000010000010)_2$. If the array elements are considered as Table 1, then the block size should be 25 bits long. If the array of the elements varies, the block size will also vary. Now the result of encrypted data becomes $(0000000000110000101111011)_2$ instead of $(110000101111011)_2$.

The proposed encryption scheme converts the input data based on the array of prime numbers considered and, by performing the multiplication of each block of prime numbers, results in the encrypted data.

3.2 Scrambler

As the encrypted data is input into the scrambler block, The 25-bit encrypted data is scrambled using the proposed algorithm. The proposed methodology for the scrambler and de-scrambler are performed as follows: The encrypted data obtained from the encryption block is act as an input to the scrambler as depicted in the Fig. 1 and

the encrypted data can either be even or odd. There are certain slight modifications that has to be understood between them.

When encrypted data is EVEN: Firstly, it is necessary to index all bits starting from the LSB. Differentiate the data by position of bits and classify them into two groups: even position bits into one group and odd position bits into another. Consider the group that contains odd bits of the data, reverse these bits and then finally concatenate them with the even bits. The reversal operation performed earlier is called Scrambling, which is responsible by the ‘Scrambler’. The data obtained at this point is called as the ‘scrambled output’. Final step is to de-scramble the output obtained by the scrambler. In de-scrambler the bits are re-ordered first from the MSB and next from the LSB. Repeat this to all the elements to complete the process.

Similarly when the encrypted data is ODD: Repeat the same scrambling process that is opted when encrypted data is even, now get the scrambled data. De-scrambler varies slightly when the considered encrypted data is odd that is described in the Sect. 3.3. This is repeated for all the elements.

3.3 De-scrambler

The scrambled data is input to the de-scrambler. Even if it is an odd branch or even a branch, till scrambling step, the procedure is the same.

The algorithm varies in the de-scrambler step. Here in the de-scrambler, the shuffling of the scrambled bits happens. It depends on the encrypted data. If the encrypted data is odd, the shuffling will start from the LSB first. If it is even, the shuffling starts from the MSB. As an example suppose the scrambled data that need to process is of 8-bit length and let’s consider the scrambled data as: $arr = 10110100$.

If the encrypted data is EVEN the shuffling of the scrambled data starts from “MSB” i.e., $arr[8], arr[0], arr[7], arr[1], arr[6], arr[2], arr[5], arr[3], arr[4]$ which resulted the de-scrambled output as: 1000110.

If the encrypted data is ODD the shuffling of scrambled data starts from “LSB” i.e., $arr[0], arr[8], arr[1], arr[7], arr[2], arr[6], arr[3], arr[5]$ which resulted in the de-scrambled output as: 01001101.

3.4 Decryption

The de-scrambled data is input to the decryption block. Conversion of encryption data to normal basic data requires the decryption process to be performed. So, to retrieve the original data, we divide the de-scrambled output, i.e., $(24,955)_{10}$ by all the prime numbers that are considered while encryption. From Table 2, we realize that there are four ones in the locations $a(1), a(4), a(5), a(7)$. These array bit positions are filled with ones and the remaining bit positions are filled with zeros. In encryption, data

Table 2 Decryption mechanism

	a(0)	a(1)	a(2)	a(3)	a(4)	a(5)	a(6)	a(7)
Encrypted data	(24,955)10							
Divisible prime numbers	2	5	3	11	7	23	19	31
Input data	0	1	0	0	1	1	0	1

is encrypted by performing multiplication, and in decryption, data is decrypted by performing division.

4 Trojan Attack on the Proposed Methodology

Hardware Trojan [3–6] attacks on integrated circuits have emerged as a key security problem. These attacks include untrustworthy personnel, design tools, or components modifying an IC during design or manufacture in an untrustworthy design house or foundry. These assaults take the form of malicious alterations to electronic hardware at various phases of the integrated circuit’s life cycle. Such changes can cause an IC to behave in an unexpected way or enable covert channels or back doors through which sensitive information can be exposed. With varied activation mechanisms referred to as triggers which triggers the Trojans and impacts referred to as payloads as the impact on the Trojan trigger will effect the payload, the number of potential Trojans for an IC of modest complexity can be inordinately enormous. An adversary might launch such an attack with the goal of causing operational failure or leaking sensitive information from inside the chip. As illustrated in the Fig. 2, the Hardware Trojan structure is made up of trigger logic and payload logic. To activate the Trojan, the trigger logic is employed as a trigger. When the Trojan is triggered, it modifies the real signal and has an influence on the payload data. Such malicious additions behave as “spies or terrorists on a chip,” and can be incredibly strong, potentially causing disastrous repercussions in a variety of applications. These harmful circuits have been termed “hardware Trojans” in contrast to software Trojans, which attack a computer’s operating system (OS).

The name is derived from a mythical episode attributed to the ancient Greeks during the Trojan War, in which the Trojan army was given a wooden horse, which they took into their city walls without understanding the enemy (Greek) soldiers were hidden inside the hollow horse. The seemingly trustworthy horse morphed into a strong and wicked weapon that had a significant impact on the outcome of the Trojan war. A hardware Trojan mainly have two key characteristics, similar to its mythological analogy as described:

1. It should posses to have a malicious intent.
2. It should escape detection throughout the post-manufacturing test/validation phase.

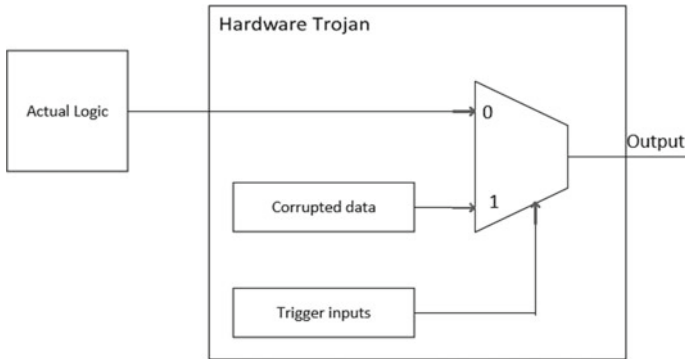


Fig. 2 Hardware Trojan

The current economic situation exacerbates the vulnerability to Trojan attacks. The trigger mechanisms in the hardware Trojan are divided into two types: analog and digital. In this research, the focus of study is on digital Trojans. Again, digital Trojans are divided into two types: combinational and sequential. Combinational Trojans deals with the fact that if a change in the inputs are detected, the output is activated. Sequential Trojans, on the other hand, are triggered when a certain sequence occurs.

The FSM [12, 13] based sequential attack is performed by inserting the sequential Hardware Trojan [14, 15]. In the design phase of the IC life cycle, the hardware Trojan is inserted. The Finite State Machine (FSM) based Trojan attack is proposed as shown in Fig. 3. As the FSM processes the serial data only. It is not possible to fed the parallel data from the scrambler. So, Parallel in Serial Out (PISO) is used to convert the parallel data to serial and fed the serial data to the FSM. The PISO is a shift register which stores the data, and the data is delayed by number of stages times the clock period and shifts for every clock cycle. It enables the parallel data loading into all stages prior to any actual shifting takes place. The serial data from PISO is fed to the FSM. And FSM is designed in such a way that if a particular sequence is detected, then the Trojan trigger will occur. The FSM output acts as a select line to the multiplexer, where the multiplexer is supposed to send the corrupted data from the memory or actual scrambler data to the de-scrambler block based on the FSM output depicted in Fig. 3.

5 Performance

The performance section deals with the performance metrics that are considered. In this paper, time, space, and area are targeted as performance metrics.

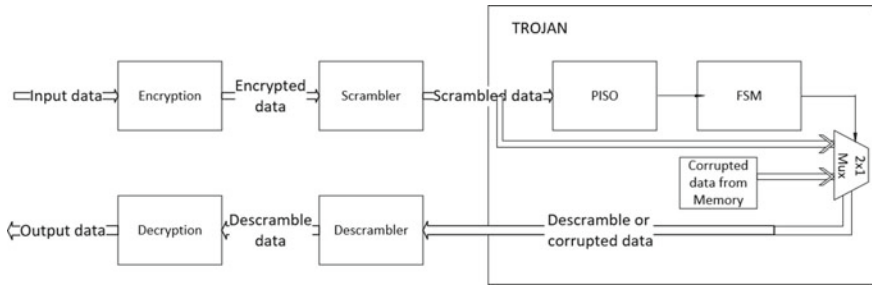


Fig. 3 Hardware Trojan insertion on proposed methodology

5.1 Time

The time criteria are determined by the prime numbers that the user considers during the encryption and decryption procedure. When higher digit prime numbers are used, the computation of multiplication takes time during the encryption process, and division takes time during the decryption process. So, if time is an essential element for the user, he or she must pick prime numbers carefully.

5.2 Space

The space criteria depend on how many prime numbers to choose and the values that have been chosen. Suppose ‘x’ ranges from 1 to 8. If $x = 1$, the minimum prime number to be used is 2, which is represented using $\max = 2$, i.e., two binary bits are required to represent the decimal number 2. The overhead proportion is calculated using \max/x , which is $2/1$, resulting in a value of 2. If $x = 2$, the two minimum prime numbers selected are 2 and 3. As a result, the maximum outcome will be 6. To represent the decimal digit 6 in binary, 3 bits are required. The overhead proportion is computed as $3/2$, yielding 1.5. The overhead proportion is shown in the Table 3 for the first eight prime numbers.

Assume the user requires a threshold of overhead of 5. He or she must determine the overhead fraction of the prime numbers using the technique outlined in the Table

Table 3 Overhead analysis of first eight prime numbers [1]

	2	3	5	7	11	13	19	19
x	1	2	3	4	5	6	7	8
Max result	2	6	30	210	2310	30,030	51,510	9,699,690
Max	2	3	5	8	12	15	19	24
Overhead	2	1.5	1.67	2	2.4	2.5	2.71	3

3. If the threshold value is constrained to 5, the calculations are completed by treating the threshold as 5. If the threshold value is restricted to 5, the observations are free to pick any prime integer between 1 and 26. By considering the 26th prime number the max result is 129 bits. Overhead will be $\max/x = 129/6$ which results to 4.96. It is entirely depending on the user that the amount of overhead to be chosen. To have a higher security the user is free to choose any value of x . So, how can the overhead be specified? The answer is that it totally depends on the application in which this technique is used.

For example, if x is 8, the prime numbers evaluated for encryption are 2,3,5,7,11,13, and the eighth prime number is 1511. To select the block size in the encryption procedure, we multiply all the prime integers. The result is 771,280,610 (7.71E+08), which will need a maximum of 30 binary bits to represent. The resulting overhead is $\max/x = 30/8 = 3.75$. Even if the overhead is smaller than the threshold value of 5, the user must make time trade-offs. If he or she chooses the larger prime number, encryption and decryption will take time. The conclusion is that even if we consider the highest prime number as a random 8th number, the overhead fraction will still be less than 4, and how can we gain from this? This is addressed under the section on security.

5.3 Security

Assume the attacker has an extremely fast computer and is able to locate all of the possibilities. It should be noted that nothing can be done unless we have all of the prime number values and their orders. The challenge is estimating how many prime numbers will be used, their values, and their order. Assume we choose eight unique integers that maps to “ r ” in Eq. 1 from a set of 26 prime numbers, relates to “ n ” in the Eq. 1 with no duplication and that the order is important. The permutations for processing the numbers is represented as “ P ”. Equation 1 describes the permutations formula.

$$P = n!/(n - r)! \tag{1}$$

Assuming, $n = 20$ and $r = 8$, then $p = 9.7E+18$. Fortunately, attacker does not know the right order of all the bits and the prime numbers considered.

6 Results and Discussions

The simulations are carried out using Verilog [15] as a Hardware Description Language on QuestaSim 10.4e simulator. To exercise the design, different input patterns are being generated and the simulation results are obtained. Figure 4 describes the results on the proposed methodology as shown in the Fig. 1. The generated

clock frequency is 50 GHz, and the input data is driven to the top module where all the internal components connections will be present. After the reset being de-asserted this data is fed to encryption block. The encrypted data is transferred to the scrambler block after the input data is encrypted using the encryption method as shown in Table 1. This scrambler block scrambles the encrypted data, making it more difficult for attackers to decrypt it. The scrambler block is followed by the de-scrambler block, which uses the de-scrambling algorithm to decode the scrambled input as depicted in Fig. 2. The de-scrambled output is transferred to the decryption block, which decrypt's it using the decryption algorithm illustrated in Table 2, which is the opposite of the encryption algorithm. The simulations are extended for the 16-bit input data as well by considering the 16 different prime numbers, i.e., 2,5,3,11,7,26,19,31,41,47,71,89,97,53,73, and the results are captured in Fig. 5.

The simulations are been carried out by inserting the FSM-based Trojan depicted in the Fig. 3. Here, two state Mealy FSM is been considered. So, two bits out of scrambled data are fed to the PISO block. FSM is designed in such a way that, it will be trigger only when it will detect the 2'b10 sequence from PISO. If this sequence is not detected the results are normal as in Fig. 4. As shown in the Fig. 6 once the FSM detects the 2'b10 sequence, the FSM is triggered and which results in the transfer of undesired data to the de-scrambler, from memory through multiplexer.

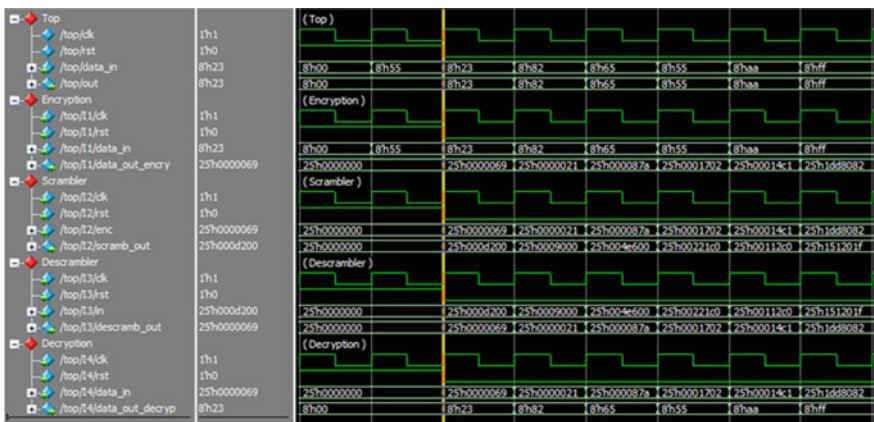


Fig. 4 Simulation waveform of the proposed methodology without Trojan attack for 8-bit input data

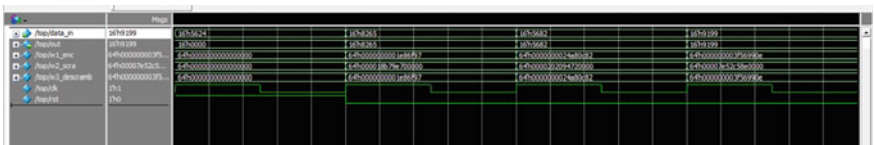


Fig. 5 Simulation waveform of the proposed methodology without Trojan attack for 16-bit input data

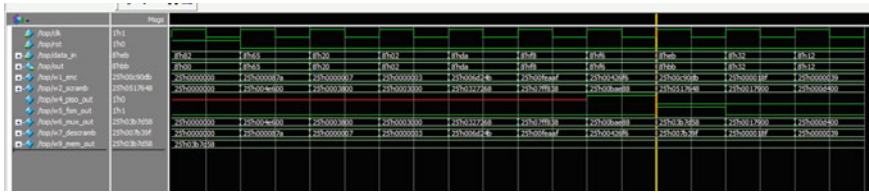


Fig. 6 Simulation waveform of the proposed methodology with Trojan attack

The multiplexer has two inputs, one from the needed scrambled output and the other from the attacker’s undesirable data from memory. As, the FSM is triggered the undesirable data is fed to de-scrambler. As a result of this, the system outputs unexpected data.

7 Conclusion

The present-day data is being subjected to different security issues, which promotes more research on the different ways the data could be encrypted. Hence, an efficient two-level encryption methodology, including the scrambler and de-scrambler technique has been formulated, ensuring double security. Firstly, a simple scrambling and de-scrambler methodology has been discussed. This has been implemented in the practical scenario, where a Trojan has been inserted. The discussed technique has been used to safeguard the actual data from the adversary. The scrambled data must always be same as the de-scrambled data, unless the Trojan is inserted. Once the data is attacked by the Trojan, the scrambled data and the de-scrambled data no longer remain same. This therefore makes the adversary retrieve the wrong data. The scrambling and de-scrambler significantly increases the difficulty in detecting the original data, hence allowing better encryption.

References

1. Haidar I, Haidar A, Haraty R (2018) Scrambled prime key encryption, pp 94–99. <https://doi.org/10.1145/3281375.3281376>
2. Singh G, Kinger S (2013) A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. *Int J Comput Appl* 67:33–38. <https://doi.org/10.5120/11507-7224>
3. Chakraborty R, Narasimhan S, Bhunia S (2009) Hard-ware Trojan: threats and emerging solutions, pp 166–171. <https://doi.org/10.1109/HLDVDT.2009.5340158>
4. Chakraborty R, Saha I, Palchaudhuri A, Naik G (2013) Hardware Trojan insertion by direct modification of FPGA configuration bitstream. *Design Test IEEE* 30:45–54. <https://doi.org/10.1109/MDT.2013.2247460>
5. Bhunia S, Hsiao MS, Banga M, Narasimhan S (2014) Hardware Trojan attacks: threat analysis and countermeasures. *Proc IEEE* 102(8):1229–1247. <https://doi.org/10.1109/JPROC.2014.2334493>

6. Jin Y, Kupp N, Makris Y (2009) Experiences in hardware Trojan design and implementation, pp 50–57. <https://doi.org/10.1109/HST.2009.5224971>
7. Joe CV, Raj JS (2021) Deniable authentication encryption for privacy protection using blockchain. *J Artif Intell Capsule Netw* 3(3): 259–271
8. Shakya S (2021) collaboration of smart city services with appropriate resource management and privacy protection. *J Ubiquit Comput Commun Technol (UCCT)* 3(01):43–51
9. Hemachand M, Prabhu E (2020) Secured netlist generation using obfuscation technique. *J Crit Rev* 7(4):878–881
10. Gowtham M, Nikhil J, Eswar MS, Ramesh SR (2021) Hardware Trojan detection using supervised machine learning. In: 2021 6th international conference on communication and electronics systems (ICCES), pp 1451–1456
11. Tripathi SL, Saxena S, Sinha SK, Patel GS (2022) Digital VLSI design and simulation with Verilog. Wiley Online Library, pp 1
12. Deepthi S (2021) Hardware Trojan detection using ring oscillator. In: 2021 6th international conference on communication and electronics systems (IC-CES), pp 362–368. <https://doi.org/10.1109/ICCES51350.2021.9488935>
13. Wang T-H, Edsall T (1998) Practical FSM analysis for Verilog. In: Proceedings international Verilog HDL conference and VHDL international users forum, pp 52–58. <https://doi.org/10.1109/IVC.1998.660680>
14. Krishna AV, Prabhu E (2022) Hardware Trojan detection at behavioral level using inline assertions and verification using UVM. In: Smys S, Balas VE, Palanisamy R (eds) Inventive computation and information technologies. Lecture notes in networks and systems, vol 336. Springer, Singapore
15. Thomas DE, Moorby PR (2007) The Verilog® hardware description language. Springer US. <https://books.google.co.in/books?id=cZIMBwAAQBAJ>

Impact of Industrial Revolution 4.0 on Vietnamese Ethnic Minority Human Resources



Hien Thu Thi Ta, Trung Tran, and Phuong Thuy Thi Nguyen

Abstract Industry 4.0 is upon us with the dynamic developments in science and technology that are creating significant changes in global production. As a result, Vietnam is experiencing profound changes, requiring its labor force to meet the demands of a modern, digital world. To meet that requirement, Vietnam must solve new problems, including training and retraining its workforce to build a contingent of cadres and civil servants to develop economic opportunities in ethnic minority areas. This study examines 464 questionnaires from managers in nine provinces representing disparate ethnic minority regions of Vietnam. Our results indicate a positive effect of Industry 4.0 awareness on physical and mental strength and managerial acumen. Furthermore, it positively impacts the development of ethnic minority human resources. Based on the influence level of factors, the authors propose several solutions to promote positive impacts and mitigate the adverse effects of Industry 4.0 for developing Vietnamese ethnic minority human resources.

Keywords Industrial revolution 4.0 · Influencing factors · Human resources · Ethnic minority managers · Human resource development · Ethnic minorities

1 Introduction

Humanity is entering a new era—the digital age ushered in by the Fourth Industrial Revolution (Industry 4.0). The advancements in science and technology, especially digital technology, connected networks, and metadata, are the tools and means to help connect globally and promote profound change in all fields in the world, from

H. T. T. Ta · T. Tran (✉)

VNU University of Education, Vietnam National University, Hanoi, Vietnam

e-mail: trungt1978@gmail.com

T. Tran

Vietnam Academy for Ethnic Minorities, Hanoi, Vietnam

P. T. T. Nguyen

Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

e-mail: phuongntt@ntt.edu.vn

production, business, culture, society, defense, and security services of all countries. However, Industry 4.0 is posing new challenges for countries especially developing countries like Vietnam, which require high-quality human resources to adapt and keep pace with a rapidly evolving society.

Currently, Vietnam ranks 15th among the most populous countries globally, ninth in Asia, and third in Southeast Asia [1]. Vietnam's population is primarily rural (approximately 68%), especially the ethnic minority workforce, accounting for a significant proportion, often living in remote or mountainous areas with challenging terrain and undeveloped socio-economic conditions. The breakneck development of science and technology during Industry 4.0 profoundly impacts the people in rural Vietnamese communities. The problem lies in the solutions to improve public awareness and develop socio-economic channels for ethnic minorities in mountainous areas to meet the integration demands of the latest technological revolution. One key initiative to achieve this goal is establishing ethnic minority human resources because the human resource component is considered the most critical and vital factor for determining growth and socio-economic development. At the same time, it is a decisive factor in the exploitation, use, protection, and regeneration of other resources. Ethnic human resource development is one of the three strategic breakthroughs institutionalized in Resolution 52/NQ-CP by the Government of Vietnam to promote and develop ethnic minority human resources during 2016–2020 until 2030 [2]. To meet that goal, Vietnam must build cadres of civil servants with sufficient quantity and quality assurance. Raising awareness among officials and civil servants about the new Industrial Revolution 4.0; training and retraining to equip cadres and civil servants with professional knowledge to meet Industry 4.0 requirements and international integration, and developing strategies to attract talent to work in state administrative agencies are practical and effective measures. Therefore, this study examines human resources development in the new era of Industry 4.0, focusing on ethnic minority human resources. We also present the model and analyze its influence from the perspective of ethnic minority area managers within the framework of the new Industrial Revolution regarding physical, mental, spiritual, and cognitive factors and activities. Finally, we propose several solutions to promote positive impacts and mitigate adverse effects of Industry 4.0 in ethnic minority human resource development in Vietnam.

2 Literature Review

2.1 *Human Resource Development in the Era of the Industrial Revolution 4.0*

The Fourth Industrial Revolution is a digital revolution with digital elements, artificial intelligence (AI), the Internet of things, and big data technology [3, 4]. It poses many urgent problems requiring innovative solutions from Vietnam's human resource

operation, which needs to train its labor force to improve efficiency. Resolution of the 13th National Party Congress set out 12 national development initiatives for the period 2021–2030, including (1) prioritizing resources for rural infrastructure development in mountainous, ethnic minority areas; (2) accelerating the national digital transformation and developing the digital economy based on science and technology innovation; (3) creating breakthroughs in fundamental and comprehensive education and training; and (4) developing high-quality human resources and attracting and appreciating talents. Previously, Decision No. 579/QĐ-TTĐ, dated April 10, 2011, of the Prime Minister of Vietnam on the Strategy for Human Resource Development for the period 2011–2020, focused on comprehensive human resource development, including the following: physical factors, knowledge, skills, behaviors, and political and social consciousness-raising as required for comprehensive human development and sustainable national development. This strategy contains a solution to renew the perception and use of human resource development.

Industry 4.0, with its new technologies affecting value chains, supply chains, and many new and emerging industries, creates radical changes in the workforce and management trends [5]. Therefore, it is critical to raise awareness of human resource management and improve the quality of human resources. Hue [5] proposed models, trends, and human resource management requirements in the digital era, which include attracting, training, developing, and maintaining human resources. The knowledge creation model by Nonaka et al. proposed several solutions to change human resource development management to meet the needs of the Industrial Revolution 4.0 [6]. Solutions focused on (1) building a flexible organizational structure, (2) digital transformation and establishing a human resource apparatus adapted to high technology, (3) setting up knowledge warehouses and information sharing channels, (4) creating training policies and fostering the direction of personal initiative, or learning through experience, and (5) building a culture of sharing.

Human resources are “*the population and quality of people, including physical and mental health, health and intelligence, capacity and qualities.*” That is, “*total human potential, including physical, mental, health, intellectual, qualities, ethics, capacity, the experience of a country, territory, locality or a company capable of mobilizing in the process of economic development society*” [7]. When discussing human resources, one must consider quantity, structure, quality, and cultural characteristics, such as traditions, customs, and national consciousness. From an individual perspective, it is necessary to consider quality, capacity, and physical factors [8].

Human resource development during the Industry 4.0 must consider aspects related to a worker’s capacity to respond to job requirements and the labor force. According to Fabian Hecklauer et al., core competencies are needed to respond to economic, social, environmental, technological, political, and legal challenges [9]. This group of competencies presents in Table 1.

Table 1 Derivation of core competencies for identified challenges [9]

Fields	Core competencies to deal with challenges
Economic challenges	<i>Ongoing globalization</i> Intercultural skills, language skills, time flexibility, networking skills, and process understanding
	<i>Increasing need for innovation</i> Entrepreneurial thinking, creativity, problem-solving, work under pressure, state-of-the-art knowledge, technical skills, research skills, and process understanding
	<i>Demand for higher service-orientation</i> Conflict resolution, communication skills, ability to compromise, and networking skills
	<i>Growing need for cooperative and collaborative work</i> Ability to be compromising and cooperative, ability to work in a team, communication skills, and networking skills
Societal challenges	<i>Demographic change and changing social values</i> Ability to transfer knowledge, accept work-task rotation and work-related change (ambiguity tolerance), time and place flexibility, and leadership skills
	<i>Increasing virtual work</i> Time and place flexibility, technology skills, media skills, and understanding IT security
	<i>Growing complexity of processes</i> Technical skills, process understanding, motivation to learn, ambiguity tolerance, decision making, problem-solving, and analytical skills
Technical challenges	<i>Exponential growth of technology and data usage</i> Technical skills, analytical skills, efficiency in working with data, coding skills, understanding IT security, and compliance
	<i>Growing collaborative work on platforms</i> Ability to work in teams, virtual communication skills, media skills, understanding of IT security, and ability to be cooperative
Environmental challenges	<i>Climate change and resource scarcity</i> Sustainable mindset, motivation to protect the environment, and creativity to develop new sustainable solutions
Political and legal challenges	<i>Standardization</i> Technical skills, coding skills, and process understanding
	<i>Data security and personal privacy</i> Understanding of IT security, compliance

2.2 Developing Human Resources for Ethnic Minorities in the Industrial Revolution 4.0

Vietnam has 54 ethnic groups, including 53 ethnic minorities distributed throughout the Northeast, Northwest, North Central, Central, South Central, Central Highlands, Southwest, and Southeast. Over the years, the party and state set many educational

policies to invest in education and training in ethnic minority and mountainous areas, creating positive change for human resource development. In the Fourth Industrial Revolution, an investment in all aspects is critical to developing ethnic human resources.

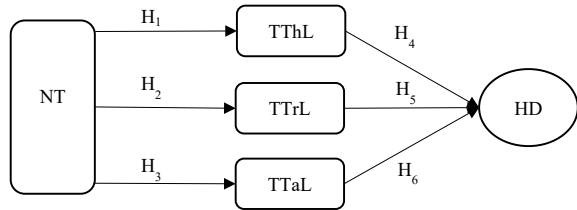
The market economy is considered the first breakthrough in industrialization and modernization. It directly affects the development of ethnic minority human resources. The market economy promotes goods and product production with exchangeable value while driving the quality of ethnic minority human resources. Several methods develop this group besides training and fostering. There are specific production examples where workers are encouraged to receive scientific education and professional training for technology-related jobs. According to Hue [5], the human factor is essential in human resource management models. He points to the human-centric models at Harvard and Michigan, where human resources relate to awareness, training, motivation, promotion, and results.

The ethnic minority workforce is capable of participating in the country's socio-economic development, especially in mountainous and ethnic minority regions. Human resources become evident through quantity (number of employees, size, distribution, and structure) and quality (quality of physical and mental health, related to moral qualities, beliefs, and lifestyle; quality of intellect related to education level and professional skills and capacity) of the ethnic minority workforce [10]. In other words, ethnic minority human resources are the total quantity, quality, and structure of ethnic minorities with the criteria of physical and mental strength, intelligence, and labor reserve of ethnic minorities. Ethnic minority human resources create the capacity of ethnic minority workers, formed in the creative work process for development and social progress [11].

To develop human resources and human management, it is necessary to determine each worker's position in the collective social system by defining their function, power, and organizational role. Furthermore, training and fostering people to perform well in their professional roles is essential. By creating optimal conditions for every individual in the organization to succeed and fulfill their functions, obligations, and powers, the system will work efficiently for all [12]. Several adaptations and integration considerations are required in human resource development. They include physical and physiological adaptation, professional and technical operations, stress, working time, and adapting to co-workers in terms of psychology, temperament, personality, tendency, value orientation, interest, conception, and habit. All these factors play a role in creating a productive psychological atmosphere in the work environment.

Therefore, it can be said that developing human resources in the Industrial Revolution 4.0 is the essential process of creating a physically and mentally healthy workforce with a high level of education, career training, and a lifestyle and behavior suitable to Vietnam's socio-economic development requirements. When training and fostering ethnic minority human resources, all three factors (physical, mental, and intellectual) must be considered. The three factors are intertwined and interact with each other to create a quality workforce. Physical capacity is the foundation of human

Fig. 1 Proposed model of the relationship between factors



resources. Intelligence determines the quality, and mental energy transforms physical and mental properties into production practice. Therefore, to develop the ethnic minority human resources, the quality of labor, expressed through physical, mental, and intellectual vitality, must remain the focus.

In this study, we identified the factors affecting the development of ethnic minority human resources in the context of the new Industrial Revolution. From our research, while selecting several legal documents and proposals issued by the Vietnamese Government, we realized the necessity of clarifying the relationship between the managers' cognitive factors and ethnic minority human resource development [2, 13]. Next, we selected cognitive factors related to physical, mental, and intelligence that affect ethnic minority human resources development. Specifically, we propose a research model with the factorial relationship, as shown in Fig. 1.

Based on this research model, this study proposes the following hypotheses:

H1: Cadre managers' awareness of Industry 4.0 (NT) has a positive influence on physical strength awareness (TThL).

H2: Cadre managers' awareness of Industry 4.0 (NT) has a positive impact on intellectual capacity awareness (TTrL).

H3: Cadre managers' awareness of Industry 4.0 (NT) has a positive effect on mental awareness (TTaL).

H4: The physical strength awareness (TThL) factor positively affects the activities of human resource development of ethnic minorities (HD).

H5: The intellectual capacity awareness (TTrL) factor positively affects the activities of human resource development of ethnic minorities (HD).

H6: The mental strength awareness (TTaL) factor positively affects the activities of human resource development of ethnic minorities (HD).

3 Research Methodologies

Our scientific research project, funded by the Nafosted Foundation, titled "*Barriers to the development of human resources for ethnic minorities in Vietnam in the context of the Industrial Revolution 4.0*," had consent to survey 464 voluntary participants (ethnic minority management officials) in nine provinces and cities in all three Vietnamese regions. Specifically, 03 provinces Son La, Ha Giang, and Thai Nguyen (representing the Northern mountainous region); 04 provinces Nghe An, Quang

Nam, Dak Lak, and Lam Đong (representing the central coast region and the central highlands region); and 02 provinces Soc Trang and Tra Vinh (representing the South-west region). Regarding the breakdown of ethnic minority groups, Vietnam consists of 54 ethnic groups, of which 53 ethnic minorities account for 14.6% of the total population. However, we took a representative sample and focused on surveying two groups: ethnic minority groups with a high level of human resources development (Muong, San Diu, and Khmer) and ethnic minorities with challenges in human resource development (Xo Đang, Kho Mu, and Mong). Before the survey, we informed the participants that our objective was only for scientific research, not for profit purposes, and all personal information would be kept confidential. Survey participants were fully guided on the question content, rated on a Likert scale, and spent time completing the assessment.

Our objective was to collect data from 464 local managers on their perceptions of Industry 4.0 and their awareness of local managers' intellectual, physical, and mental strength. The four-part questionnaire consisted of 36 questions assessing the impact of the Industrial Revolution 4.0 on ethnic minorities. It was scored on a five-point Likert scale from one to five, corresponding to strongly agree, agree, confused, disagree, and strongly disagree. Each participating administrator received the same ticket and answered all general information and quantitative questions. We then synthesized the data for analysis based on results from the managers' responses via mathematical-statistical methods and data processing with SPSS 20 software.

Of the 464 questionnaires, we obtained 450 valid surveys. Tables 2 and 3 present our results, including the number of votes by ethnic groups and the job positions of participating officers in charge.

The process of designing and analyzing data: After collecting the surveys, we coded and entered data into the SPSS 20 software. The data was cleaned using the frequency command in SPSS 20 to remove invalid input values beyond the conventional values, and then the original questionnaire was rechecked for correct data

Table 2 Number of specific questionnaires by provincial ethnic minority group

Province	Ethnic							
	Muong	San Diu	Kho Mu	Mong	Kinh	Khmer	Xo Đang	Others
Ha Giang	0	0	0	12	15	0	0	23
Son La	5	0	0	10	15	0	0	20
Nghe An	0	0	0	1	25	0	0	25
Thai Nguyen	1	3	0	3	13	0	0	30
Quang Nam	0	0	0	0	15	0	34	0
Lam Đong	0	0	0	0	9	0	1	40
Dak Lak	3	0	0	5	24	0	7	11
Soc Trang	0	0	0	0	15	34	1	0
Tra Vinh	0	0	0	0	0	50	0	0
Total	9	3	0	31	131	84	43	149

Table 3 Job positions of officers

Job	Quantity	%
Village level	17	3.8
Commune level	86	19.1
District level	180	40
Provincial level	167	37.1
Total	450	100

adjustment. Next, we tested the reliability of Cronbach's alpha of the scales. If the Cronbach's alpha coefficient satisfied the condition greater than 0.6 and the correlation coefficient of the total variable was higher than 0.3, the scale was reliable. We continued to use exploratory factor analysis (EFA) to test the fit of the research model with the obtained data set, getting relevant factors to measure observed variables and discover factors to test the proposed hypotheses. Finally, we performed Pearson linear correlation identification and multivariate regression analysis to consider the impact of the model factors, thereby testing the proposed hypotheses with statistical significance at 5%.

4 Results

4.1 Descriptive Statistics

Regarding general information, while surveying the managers, we asked questions related to mobile phone use, computers, and televisions with Internet connections (serving the Fourth Industrial Revolution) to clarify the usage level at home and work.

According to Table 4, most of the surveyed officials in nine provinces used regularly and often used mobile phones with Internet connections. The rate reached 100% in Ha Giang, Nghe An, Ha Giang, Dak Lak, and Tra Vinh. In Quang Nam, the users account for 96% (47/49 people).

Regarding the factors assessing managerial perceptions of the Fourth Industrial Revolution, per Table 5, most officials had particular perceptions about it. However, the information skills and related technology use were still not well developed or frequent (mean values ranged from 2.51 to 2.58).

There was substantial feedback on the level of less participation, and the use of modern Internet communication means: (1) NT7: "Ethnic minorities can participate in online meetings online," (2) NT8 "Ethnic minority people know and use advanced machines," and (3) NT9 "Ethnic minority students completely learn online via the Internet and exploit learning materials via the Internet and social networks."

Table 4 Mobile phone use with an internet connection by provincial officials

Province	Rarely	Sometimes	Frequently	Very often
Ha Giang	0	0	46	4
Son La	5	0	30	15
Nghe An	0	0	51	0
Thai Nguyen	2	2	33	13
Quang Nam	0	47	2	0
Lam Dong	0	1	4	45
Dak Lak	0	0	25	25
Soc Trang	0	3	9	38
Tra Vinh	0	0	21	29
Total	7	53	221	169

Table 5 Assessment of awareness levels of officials regarding industry 4.0

Variable name	NT1	NT2	NT3	NT4	NT5	NT6	NT7	NT8	NT9
The average value	2.00	2.06	2.19	2.39	2.42	2.49	2.58	2.51	2.55
Standard deviation	0.706	0.735	0.855	0.891	0.888	0.866	0.831	0.858	0.924

4.2 Test the Reliability of Cronbach's Alpha Scale and Evaluate the Value of the Scale

To evaluate the reliability, validity, and scale value, we used Cronbach's alpha coefficient and EFA with the cleaned data set. The results showed three observed variables (TThL2, TThL3, and TThL6) in the factor group TThL, and the group of factors TTrL with one observed variable (TTrL6), having the corrected item-total correlation less than 0.3, so these four observed variables were excluded from the scale. After removing the four observed variables that did not meet the above conditions, the remaining scales of the factor group NT, TThL, TTrL, TTaL, and HD all had Cronbach's alpha coefficients greater than 0.6, and the corrected item-total correlation of all observed variables was greater than 0.3, thereby showing that the observed variables in the factors have high consistency and that the scale of the factors is reliable (Table 6).

With EFA analysis results, all KMO coefficients satisfied the condition of $0.5 \leq \text{KMO} \leq 1$, Bartlett's test value $p < 0.05$, and show that the factor analysis was suitable with the obtained data set (Table 6).

Table 6 Information table of factor rating scales

Factors	Observed variables	Cronbach's alpha	Corrected item-total correlation minimum	KMO	p-value	Variable type
NT	NT1, NT2, NT3, NT4, NT5, NT6, NT7, NT8, NT9	0.895	0.543	0.842	0.000	Independent/dependent
TThL	TThL1, TThL4, TThL5, TThL7, TThL8, TThL9	0.924	0.700	0.816	0.000	Independent/dependent
TTrL	TTrL1, TTrL2, TTrL3, TTrL4, TTrL5	0.925	0.694	0.861	0.000	Independent/dependent
TTaL	TTaL1, TTaL2, TTaL3	0.921	0.771	0.703	0.000	Independent
HD	HD1, HD2, HD3, HD4, HD5, HD6, HD7, HD8, HD9	0.935	0.692	0.869	0.000	Dependent

4.3 Test the Observed Variables Fit and Model Factors

To check the convergence of observed variables with the factor extracted in the theoretical model, we used EFA for models of hypothesis H1 (independent variable NT, dependent variable TThL); model of hypothesis H2 (independent variable NT, dependent variable TTrL); model of hypothesis H3 (independent variable NT, dependent variable TTaL); with the model of hypothesis H4, H5, and H6 for the group of independent variables TThL, TTrL, TTaL, and HD (dependent variable). As a result, we obtained a data table of independent variables converging in three factors, as shown in Table 7.

At the end of the EFA step, we obtained five factors, including NT, TThL, TTrL, TTaL, and HD that were the most suitable with the 32 best-observed variables. Therefore, to change how we measured observed variables to factor measurement to test the proposed hypotheses, we created representative factors and performed Pearson correlation evaluation and linear regression analysis count.

Table 7 Rotated component matrix for the model of hypothesis *H4*, *H5*, and *H6*

Observed variables	Component		
	1	2	3
TThL7	0.844	–	–
TThL8	0.812	–	–
TThL5	0.812	–	–
TThL1	0.792	–	–
TThL4	0.758	–	–
TThL9	0.725	–	–
TTrL2	–	0.858	–
TTrL3	–	0.851	–
TTrL1	–	0.835	–
TTrL4	–	0.817	–
TTrL5	–	0.710	–
TTaL2	–	–	0.921
TTaL1	–	–	0.900
TTaL3	–	–	0.837

4.4 Test Structural Models and Hypotheses

We performed the Pearson correlation test between the independent and dependent variables. The results showed that all variables have a close linear relationship (Pearson correlation coefficient r is greater than 0) (Table 8).

We analyzed four structural equation models to test six hypotheses of the theoretical model. The results of multivariable regression analysis (Table 9) of the four equations all have test values $F < 0.05$, so the regression models are consistent with the obtained data set. The adjusted R^2 values evaluated the fit of the equation, specifically with the four regression equations:

+ Eq. 1 (Hypothesis H1):

$$\text{TThL} = 0.544 * \text{NT} \quad (1)$$

Table 8 Correlation analysis results

	HD	TThL	TTrL	TTaL	NT
HD	1	–	–	–	–
TThL	0.769**	1	–	–	–
TTrL	0.625**	0.621**	1	–	–
TTaL	0.523**	0.476**	0.458**	1	–
NT	0.492**	0.544**	0.452**	0.169**	1

** . Correlation is significant at the 0.01 level (2-tailed)

Table 9 Results of multivariable regression analysis

Models	Hypothesis	F test	Adjusted R square	Standardized coefficients beta	p-value	VIF
1	H1	0.000	0.294	0.544	0.000	1.000
2	H2	0.000	0.203	0.452	0.000	1.000
3	H3	0.000	0.027	0.169	0.000	1.000
4	H4	0.000	0.644	0.570	0.000	1.762
	H5			0.197	0.000	1.725
	H6			0.161	0.000	1.369

with an adjusted R^2 value of 0.294 shows that the independent variable NT can only explain 29.4% of the change of the dependent variable ThL, and 70.6% is due to out-of-model variables and random error.

+ Eq. 2 (Hypothesis H2):

$$TTrL = 0.452 * NT \tag{2}$$

with an adjusted R^2 value of 0.203 shows that the independent variable NT can only explain 20.3% of the change of the dependent variable TrL, and 79.7% is due to out-of-model variables and random error.

+ Eq. 3 (Hypothesis H3):

$$TTaL = 0.169 * NT \tag{3}$$

with an adjusted R^2 value of 0.027 shows that the independent variable NT explains only a tiny percentage of 2.7% of the change of the dependent variable TaL, up to 97.3% is due to variables outside the model and random error.

+ Eq. 4 (Hypothesis H4, H5, and H6):

$$HD = 0.570 * TThL + 0.197 * TTrL + 0.161 * TTaL \tag{4}$$

with an adjusted R^2 value of 0.644, showing that the independent variables ThL, TrL, and TaL are well-explained: 64.4% of the change of the dependent variable HD, and 35.6% due to variables outside the model and random error.

The significance test value of the regression coefficient (p -value) is all < 0.05 , showing the significant impact of the independent variables on the dependent variable. At the same time, the VIF coefficients of the independent variables are all less than 10, so the data does not violate the assumption of multicollinearity. Furthermore, the normalized regression coefficients beta (Table 9) values all have positive values, showing that the independent variables positively affect the dependent variables. Therefore, the hypotheses H1, H2, H3, H4, H5, and H6 have all been accepted.

Regarding the level of impact, the beta coefficient helps us easily compare each factor's influence. Regression Eqs. 1, 2, and 3 show that the perception of managers

in the context of Industry 4.0 positively changes the perception of those cadres' physical, intellectual, and mental health. Especially with regression Eq. 4, when all factors of body and mind impacted ethnic minority human resource development activities, the physical had the most significant influence (beta is 0.57), and the psychological factors (beta is 0.161) had the lowest.

5 Discussion

Our research affirms that there must be digital people to build a digital society. To create a digital government, there must be administrators with high-quality science and technology training. The constant access to information and the opportunities and challenges presented by the Industrial Revolution have changed how people think about education, health, and spiritual life [14]. People communicate easily and connect conveniently through modern technology. In particular, the Industrial Revolution 4.0 is changing people's "traditional" identities such as privacy, sense of ownership, consumption methods, skills training, and networking. Biotechnology, AI, and many other fields profoundly improve life expectancy, health, and awareness of social and ethical matters [3, 4]. The research results elucidate the impact of Industry 4.0 on ethnic minority human resources. In particular, it is critical to developing scientific and technological opportunities for ethnic minorities in rural and remote mountainous areas, focusing on comprehensive physical, mental, behavioral, legal, and professional skills. Priority should be given to opportunities for ethnic minorities with limited human resources to gradually close the gap and meet Vietnam's development requirements and current and future labor market needs. We must build a contingency of qualified intellectuals, entrepreneurs, cadres of ethnic minorities, and direct laborers to promote socio-economic development, sustainable poverty reduction, and ensure national defense and security and ethnic minority areas and mountainous areas. In addition, we must implement the instructions in Resolution No. 52-NQ/TW dated September 27, 2019, of the Politburo on several guidelines and policies to participate actively in Industry 4.0. Resolution No. 50/NQ-CP, dated April 17, 2020, of the Government, promulgates the Government's action plan to implement Resolution No. 52-NQ/TW on several topics and policies to participate actively in Industry 4.0, etc., thereby contributing significantly by transforming the thinking and requiring adaptive changes to improve the capacity, technical level, and skills and physical strength of the human resources of ethnic minorities before new requirements. The research results also show that the ethnic minority cadres, although aware of the importance and benefits of the impacts of the Fourth Industrial Revolution, are almost all conditions, habits, and skills for participating in Industry 4.0 are extremely limited and infrequent (Tables 4 and 5).

The Industrial Revolution 4.0, whose core is digital transformation in all aspects of social life, will bring opportunities but creates many challenges for all levels of state administrative agencies in general and ethnic minority human resources in particular. The digital environment leads to reduced personnel, downsizing the apparatus, and

increasing requirements for qualifications and skills in e-government and digital government management for civil servants minorities. It forces the human resources group to improve their intelligence constantly. Education in the digital space will also create unlimited learning opportunities, such as STEM, STEAM, flipped classrooms, online learning, blended learning, and open mass online learning (MOOC) [15–17]. Research results from two ethnic groups found different consensus levels on the positive impacts of the Fourth Industrial Revolution on intellectual, physical, and mental factors. In particular, the ethnic group with difficulties in human resource development has more hesitation and limited reception than the ethnic group with a higher level of human resources development. These results align with a survey on the socio-economic status of 53 ethnic minorities in 2019 [18] which shows the proportion of ethnic minority groups trained at intermediate, college, and university levels or above (intermediate: 2.5%; college: 1.7%; and university and higher: 3.3%). Ethnic groups with difficulties in human resource development, such as Muong, San Diu, Kho Mu, Mong, So Đàng, and Khmer, have low rates (Table 10). In many cases, after training, people do not return to their locality, creating a shortage of qualified human resources in ethnic minority areas.

This result shows that education is a giant barrier for ethnic minority cadres participating in the digital transformation and the Fourth Industrial Revolution. At the same time, it raises urgent policy-related issues to focus resources on human resource training, attracting and retaining good people, encouraging and creating an environment for learning, training and improving ethnic group qualifications.

To increase intellectual capacity, one of the critical solutions of the “Strategy for human resource development in Vietnam for the period 2011–2020” is to use and evaluate human resources based on capacity and job results and not overemphasize formal qualifications in recruiting and evaluating human resources. In the fiercely competitive environment of global integration, the problem exists of developing quality human resources, especially information technology human resources. Science and technology training needs emphasis to develop human resources. Along with developing human resources is the issue of using and managing Vietnam’s current human resources. For students training under the recruitment system at universities, colleges, and professional intermediate schools throughout Vietnam, after completing the course, the provinces need to help graduates secure employment

Table 10 Survey findings on co the socio-economic status of 53 ethnic minorities in 2019 [18]

Ethnic	Intermediate (%)	College (%)	University and above (%)
Muong	3.0	1.9	3.8
San Diu	2.5	2.1	4.2
Kho mu	1.0	0.6	0.8
Mong	1.4	0.7	1.1
So Đàng	1.4	0.7	1.3
Khmer	0.7	0.8	1.9

in a timely and reasonable manner befitting their training and expertise. These are critical core functions in developing high-quality ethnic minority human resources in the provinces today. Finally, we must focus on retraining ethnic minority grassroots cadres to ensure they meet the professional and expert requirements of the new situation with a reasonable remuneration. This measure will encourage stability and long-term staff establishment, especially in remote, isolated, and disadvantaged areas, so that workers feel secure and dedicated to their livelihoods.

Besides cognitive and intellectual factors, physical strength and the impact of the Fourth Industrial Revolution on ethnic minority human resources are essential in crafting human resource development strategies. Science and technology advances have created new health care and nutrition concepts. Tremendous developments in the medical field with new methods of treating, diagnosing, and monitoring patient health conditions, including innovations in managing and organizing health systems, are currently being developed. Access to health care is also changing [19]. The term digital health care has become quite familiar. Through big data and AI, healthcare devices can be integrated into smartphones to help doctors and patients monitor disease and health conditions. Moreover, routine health checks and even surgery can be performed remotely. These measures enhance human health and increase life expectancy [20].

The construction and development of an information system for managing medical care, deployed on a digital map system, to assist people in searching for suitable medical facilities is convenient, easy, and part of a modern health care and disease prevention system based on digital technology. By comprehensively applying digital technology to medical examination and treatment facilities, administrative reform becomes possible and reduces hospital workload, improving the quality of medical care. Digital technology is transforming medical records, moving away from paper documentation, and increasing the shareability and transmission speed of critical patient information and lab results. It changes how hospital fees get paid and create “smart” hospitals, enabling the integration of information and data in a national health database in Vietnam [21]. Strengthening the training and retraining of civil servants, public employees, and healthcare sector workers in ethnic minority areas improves information technology skills and application, ensuring the safety and security of the operational infrastructure. Applying such skills to health officials and civil servants in ethnic minority areas who handle medical conditions as daily tasks requires special attention.

Recently, most ethnic minority communal areas have health stations, up to 99.5%, according to a 2015 survey. The percentage of communes with health stations meeting national standards for commune health in 2019 reached 83.5%, nearly twice as high as in 2015 (45.8%). Currently, more than 33.4 thousand leaders and healthcare workers are employed at such communal health stations. Ethnic leaders and staff account for 37.9%. Most (slightly more than 50%) leaders and staff at commune health stations have medical, doctor, or nursing qualifications. The number of midwives accounts for only 15.1% of the total number of leaders and employees of commune health stations. Unfortunately, although many programs and projects attract doctors to grassroots healthcare work, many localities, predominantly ethnic minority communes, still

have a shortage of medical staff. Compared to 2015, the percentage of villages in ethnic minority communes with village health workers decreased slightly, from 85.0% in 2015 to 83.5% in 2019. This rate is highly concentrated in remote areas such as the Northern Midlands and Mountains, the North Central Coast and the Central Coast, and the Central Highlands, which also decreased slightly. Therefore, the problem of maintaining and developing the village health network, an extension of the health sector to a wide range of remote and disadvantaged areas of the country, still needs resolution and more research to find a more appropriate and effective solution, especially one that meets the requirements on quantity and quality of human resources of the Industrial Revolution 4.0 [18].

To develop Vietnam's ethnic minority human resources, it is necessary to focus primarily on developing the "human capital" of ethnic minorities by strengthening reasonable policies to improve and raise the quality of manpower. According to this approach, from an economic perspective, human resource development can be considered as the accumulation of human capital and effective investment in the development of an economy. From a political perspective, human development is about preparing adults to participate in political processes, specifically as citizens of society. From a social and cultural perspective, human resource development helps people have a complete, affluent life and be less constrained by tradition. Human resources progress in many ways. The most obvious is through formal education, starting at the primary level and continuing with various forms of general education, then higher education with colleges, universities, and vocational training institutions. Next, human resources also develop through systematic and informal training programs in recruitment agencies, adult education programs, and members of cultural, religious, political, and social groups. The third way is self-development as individuals strive to acquire more knowledge, skills, and competencies by preparing their ideas—by attending courses that are not only in-person but can be done virtually, by cyberspace, technological devices, through online education platforms, and other means of digital technology [16, 22]. Motivation for self-development is directly related to personal ambition, social values, training, and career initiatives. Two additional processes of improving human resources development are (1) improved worker health through access to better healthcare programs to improve overall public health, and (2) better nutrition that increases functional ability. It is evident that improvements in health and nutrition are linked, and like formal education, can be both a cause and an effect for economic prosperity.

6 Conclusion

The Industrial Revolution 4.0 has impacted ethnic minority human resources in Vietnam, especially civil servants and public employees at all levels in ethnic minority and mountainous regions. The new requirements for high-quality human resources require effective policies and solutions that serve the ethnic minority populations of our country. Particular emphasis is needed to develop the quantity and quality of

human resources (such as physical and mental capacity, qualifications, knowledge, practical and professional skills, and moral qualities), which have a decisive impact on the Vietnamese socio-economic development process.

Acknowledgements This study was funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED), under the project entitled “Barriers to the development of ethnic minority human resources in Vietnam in the context of the industrial revolution 4.0,” grant number 02/2019/NCUD.

References

1. DanSo.Org (2020) Countries of the world by population
2. Government of Vietnam (2016) Resolution 52/NQ-CP on accelerating the development of human resource for ethnic minorities in the period of 2016–2020, orientation to 2030
3. Ranganathan G, Smys S (2019) Survey on robot process automation application in various industries. *J Electr Eng Autom* 1(2):113–122
4. Sivaganesan D (2020) Smart contract based industrial data preservation on block chain. *J Ubiquitous Comput Commun Technol* 2(01):39–47
5. Hue TT (2020) Human resource management in the digital era and the role of knowledge management. Hanoi National University Publication, House
6. Nonaka I, Toyama R, Hirata T (2008) Managing flow: a process theory of the knowledge-based firm. In: *Management flow a process theory knowledge-based firm*, pp 1–252. <https://doi.org/10.1057/9780230583702>.
7. Ha LS (2015) Female human resources in the period of industrialization, modernization and international integration: a number of theoretical and practical issues. Social Sciences Publishing House
8. Kha PV (2017) Model of human resource development and sustainable human resource development in educational science research. *Educ Sci* 1
9. Hecklau F, Galeitzke M, Flachs S, Kohl H (2016) Holistic approach for human resource management in industry 4.0. *Procedia CIRP* 54:1–6. <https://doi.org/10.1016/j.procir.2016.05.102>.
10. Baulch B, Chuyen TTK, Haughton D, Haughton J (2008) Ethnic minority development in Vietnam 43(7):1151–1176. <https://doi.org/10.1080/02673030701526278>.
11. Van Trung T (2015) Current policy on developing young human resources in the Northwest region of Vietnam. National Academy of Public Administration
12. Van Hoi H (2009) Culture in human resource management. *J Sci VNU* 25:92–98
13. Government of Vietnam (2020) Resolution 50/NQ-CP 2020 action plan to implement resolution 52-NQ/TW
14. Zervoudi EK (2020) Fourth industrial revolution: opportunities, challenges, and proposed policies. *Ind Robot New Paradigms*
15. Marr B (2019) 8 Things every school must do to prepare for the 4th industrial revolution
16. Thao TTP, Thai LD, Thanh HT, Tran T, Trinh LTT, Vuong QH (2019) Mobile learning for high-school mathematics as a path to better sustainability in a fast-changing society: an exploratory study from Vietnam. *Probl Perspect Manag* 17(2). <https://doi.org/10.31219/OSF.IO/NVQEA>.
17. Tran T, Ho MT, Pham TH, Nguyen MH, Nguyen KL, Vuong TT, Nguyen TH, Nguyen TD, Nguyen TL, Khuc Q, La VP (2020) How digital natives learn and thrive in the digital age: evidence from an emerging economy. *Sustainability* 12(9):3819. <https://doi.org/10.3390/su12093819>.

18. Ethnic Minorities Committee & General Statistics Office (2020) Results of the survey to collect information on the socio-economic status of 53 ethnic minorities in 2019. Vietnam Statistical Publishing House
19. e Castro e Melo JAGM, Faria Araújo NM (2020) Impact of the fourth industrial revolution on the health sector: a qualitative study. *Healthc Inform Res* 26(4):328–334. <https://doi.org/10.4258/hir.2020.26.4.328>.
20. Jung M (2019) Digital health care and the fourth industrial revolution. *Health Care Manag (Frederick)* 38(3):253–257. <https://doi.org/10.1097/HCM.0000000000000273>.
21. Ministry of Information and Communications (2020) Official dispatch 2390/BTTTT-THH 2020 implementing the national digital transformation program to 2025
22. Trung NT, Thao TP, Trung T (2019) Realistic mathematics education (RME) and didactical situations in mathematics (DSM) in the context of education reform in Vietnam. *IOP Publishing* vol 1340, no 1, p 012032

Fine-Tuning MobileNet for Breast Cancer Diagnosis



Huong Hoang Luong, Nghia Trong Le Phan, Toai Cong Dinh,
Thuan Minh Dang, Tin Tri Duong, Tong Duc Nguyen, and Hai Thanh Nguyen

Abstract Breast cancer can be considered one of the significant causes of death, especially among women worldwide. Therefore, it is essential to detect and diagnose breast cancer as early as possible to reduce the adverse effects on patients and protect women's health. This study proposes a model applying transfer learning and fine-tuning to classify and detect benign, malignant breast cancer, and normal breast. We train the proposed model with transfer learning from the pre-trained MobileNet model to identify breast cancers and optimize the prediction results. The dataset contains 780 ultrasound images categorized into three classes which are benign breast cancer (437 images), malignant breast cancer (210 images), and normal breast (133 images). The experimental results show that applying the transfer learning and fine-tuning technique from the MobileNet model achieves promising results, with the accuracy and $F1$ -score being 0.9651–0.9648, 0.9412–0.9417, and 0.9060–0.9085, respectively, with three scenarios.

Keywords Breast cancer classification · Transfer learning · Fine-tuning · Convolutional neural network

1 Introduction

Breast cancer comes out when damaged (characteristic) cancer cells are detected in the model, which usually develops from the tubes inside the breast. The cancer cells can then expand the entire organ and sing to different body parts. About 255,000 women and 2300 men every year in the United States are diagnosed with breast cancer, with approximately 42,000 women and 500 men dying. According to the survey, white women have a lower death rate than black women from breast cancer [1]. Breast cancer is the most common cause of cancer mortality among women aged

H. H. Luong · N. T. L. Phan · T. C. Dinh · T. M. Dang · T. T. Duong · T. D. Nguyen
FPT University, Can Tho 900000, Vietnam

H. T. Nguyen (✉)
Can Tho University, Can Tho, Vietnam
e-mail: nthai.cit@ctu.edu.vn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes
in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_60

841

40–69 in Singapore today. Breast cancer affects around 1 in every 16 Singaporean women at some point. In China, women have a higher risk of developing breast cancer than women in Malaysia and India [2].

Breast cancer is not an infectious illness since no known viral or bacterial infections have been linked to its development. When breast cancer is detected early, survival rates range from more than 90% in high-income countries to 40% in South Africa. Women are the gender with the highest risk of breast cancer. Breast cancer affects approximately 12% of women throughout their lives. Compared with men, the incidence in men is only about 1% [3]. Treatment of male patients is similar to that of female patients. The rate of breast cancer is higher when there is a family history of breast cancer. However, most of the patients' families do not have this history, so the rate of breast cancer in women is always high even if the family has no history of illness. Breast cancer risk is increased by advanced age, obesity, hazardous alcohol use, family history of breast cancer, reproductive history, cigarette use, and postmenopausal hormone therapy [4].

Breast cancer is a harmful disease. Therefore, early diagnosis to provide treatment is necessary. A doctor can detect breast cancer with a variety of techniques. Recently, using ultrasound images is considered one of the appropriate methods to specify abnormalities in the patient's breast [5]. Ultrasound imaging is a painless medical procedure that aids doctors in diagnosing and treating disorders. Ultrasound is a non-invasive, painless method of diagnosis. It takes real-time photographs of the inside of the body, using sound waves to illustrate the structure and movement of the body's internal organs and the flow of blood via the blood arteries. As a consequence, the dataset consists of ultrasound images providing the most accurate diagnostic results [6].

This research article has taken advantage of leading convolutional neural network architectures and applied transfer learning and fine-tuning technique to the pre-trained networks. The purpose is to optimize the model's accuracy and configure the hyperparameters effectively to classify and detect breast cancer. The major contributions are described as follows:

- We present three scenarios to evaluate prediction performance in breast cancer diagnosis. Predicting normal breast and benign breast cancer is conducted in the first scenario. In the second scenario, we distinguish normal breast and malignant breast cancer. We predict all breast ultrasound images in the third scenario, including benign, malignant, and normal breasts. The purpose of performing these three scenarios is to determine whether the ultrasound images of patients are considered normal breasts or those having breast cancer. If the results are considered breast cancers, the proposed model is classified as benign or malignant.
- We have deployed five well-known convolutional neural network architectures (MobileNet [7], VGG16 [8], VGG19 [8], InceptionV3 [9], Xception [10]) to evaluate the training and testing results and compare them to our proposed model. The results demonstrate the efficiency of classifying three breast ultrasound image categories: normal, malignant, and benign.

- We apply transfer learning to the pre-trained MobileNet architecture to keep all of the model's nodes that have been trained before.
- After applying transfer learning, we apply fine-tuning technique by adding hidden layers such as dense, batch normalization, dropout layers, and configure hyper-parameters with the appropriate values to prevent the model from overfitting and optimize the accuracy of the proposed model.
- The training, validating, and testing metrics of the proposed model are calculated and compared among the considered CNN architectures. We also use the $F1$ -score and the confusion matrix for making relevant comparisons between the three scenarios and all the compared models.
- From the obtained results, we discovered that our proposed model applying the transfer learning model from the pre-trained MobileNet architecture and fine-tuning that model has promising results compared with other original CNN architectures.
- The research results will benefit users to detect the disease early based on the ultrasound images of the breast so that the doctor can make better decisions in the conclusion of breast cancer.

This research article consists of five main sections. This section introduces some general information related to this research and points out the approach for solving the given problem. Next Sect. 2, we figure out some of the related research that we used for references. After that, the third section is the proposed method Sect. 3, which describes in detail all of the techniques applied in building the proposed model of the research. After the proposed method section, we discuss the experiments processes and the way we calculate metrics and make relevant comparisons of the proposed model with others in Sect. 4. Finally, in the last Sect. 5, we finalize our research article and revisit the required fields associated with the research.

2 Related Work

Breast cancer is the leading cause of death in women, so it is essential to detect the symptoms of the disease and treat it promptly. Many studies have been related to breast cancer prediction and detection using machine learning, specifically CNN. Convolutional neural network (CNN) is one of the effective models for disease prediction. Sun et al. [11] applied the method of detecting breast cancer and segmenting two kinds of cancer, luminal, and non-luminal, that is, using DCE-MR post-contrast sequence images based on CNN models combined with ensemble learning. The author studied enhanced-magnetic resonance imaging and molecular information of 266 breast cancer patients with luminal or non-luminal subtypes. They also collect information from damaged areas through the DCE-MR sequence imaging series. Then, three CNN models were trained with fivefold cross-validation. Ragab et al. [12] proposed a model applying a deep convolutional neural network (DCNN) named AlexNet to classify and connect to the support vector machine (SVM) at the

last fully-connected layer to obtain better accuracy when detecting breast cancer. The researchers also apply data augmentation to create new data from the original dataset. The precision of the proposed model was 87.2%, with the percentage of AUC being 94%. Other research articles also apply a convolutional neural network to boost the accuracy score and reduce the human mistakes in diagnosing breast cancer. Alanazi et al. [13] investigated the proposed system by applying many CNN architectures to classify breast cancer automatically. After that, the proposed system result compares the outcomes with the machine learning algorithms. It achieved an accuracy of 87%, higher than 9% from the machine learning algorithms.

Besides CNN, MobileNet is a CNN design that employs depth-wise separable convolution layers. Srinivasu et al. [14] applied MobileNetV2 in skin disease classification. Based on the HAM10000 dataset, the author uses MobileNetV2 and short-term memory (LSTM) models because the MobileNetV2 model achieves higher accuracy on lightweight computing devices. The author has compared this model with others such as FTNN and CNN. MobileNetV2 gives better results with 85% accuracy. In addition, this model gives quick results to help doctors identify the impact of the patient's disease. Furthermore, Taresh et al. [15] applied a COVID-19 detection network based on the MobileNet structure called KL-MOB on the big X-ray image dataset to detect COVID-19 in patients. The dataset includes 2128 COVID, 5,575 pneumonia, and 8,066 normal cases. Adding Kullback–Leibler (KL) divergence loss function improved KL-MOB's performance in detection greatly. Moreover, Vinayahalingam et al. [16] have done research on the classification of caries in third molars by using the MobileNetV2 model. The author trained this model on 400 images of lesions in the maxillary and mandibular third molars. The results show that classification accuracy reaches up to 87%. Glaucoma is one of the primary causes of blindness globally, affecting more than 60 million people. Therefore, early diagnosis, as well as early treatment, will reduce the risk of the disease. Olivas et al. [17] applied two machine learning models, such as MobileNetV2 and InceptionV3, which were retrained through transfer learning for diagnosing Glaucoma. The author used these two models to classify the set of optical coherence tomography (OCT) images into two types: glaucomatous and non-glaucomatous. This study investigated retinal nerve fiber thickness images obtained from cropped OCT. The evaluation results received are that the accuracy for the right and left eyes are 86% and 90%, respectively, in the mobile net model, and for the InceptionV3 model, both reach 90%.

Transfer learning (TL) applies the previously learned model to a new situation. Deep learning is a compelling prediction method, but it requires much data. Transfer learning will reduce the amount of training data required while improving the model's performance potential. Novakovsky et al. [18] applied TL. The model improved if the multi-task model is trained with biologically relevant transcription factors (TFs) in the pre-training stage. The author demonstrates the utility of transfer learning for TFs with 500 or more ChIP-seq peak areas. Additionally, Zhang et al. [19] have done research that used transfer learning for EEG decoding based on brain–computer interfaces. Electroencephalography (EEG) algorithms are mostly based on machine learning research, although machine learning might interfere with EEG processing, resulting in calculation errors. In EEG processing, the author combined machine

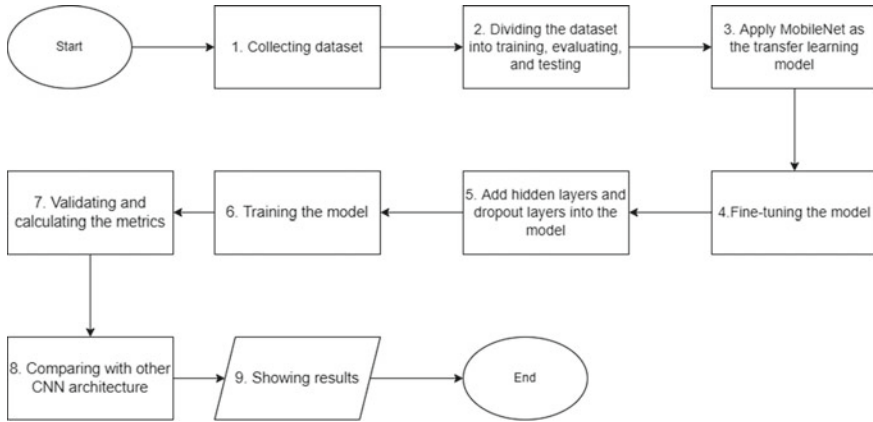


Fig. 1 Implementing procedure flowchart

learning with TL. The findings suggest that TL can increase decoding model performance across subjects/sessions and shorten the calibration time of brain–computer interface (BCI) systems. Both studies show the significant contribution of TL in research on machine learning.

3 Proposed Method

3.1 Implementing Process to Perform

The implementation process of our research will divide into nine main steps, which will be illustrated by the flowchart in Fig. 1 and explained in detail below.

1. Collecting dataset: The data was gathered by Al-Dhabyani et al. in 2018 with breast ultrasound images among women between 25 and 75 years old. The total number of female patients is 600. The photos are converted as PNG files. The pictures are divided into normal, benign, and malignant categories [20].
2. Dividing the dataset into training, evaluating, and testing: After collecting all the ultrasound images in the dataset, the images will be randomly chosen for the training, evaluating, and testing process. In detail, there are 70% images used for training, 15% images used for evaluating, and 15% images used for testing the model’s accuracy.
3. Applying MobileNet as the transfer learning model: We use MobileNet, a convolutional neural network intended for mobile, to categorize and predict what case of cancer the images in the dataset have because those images all have the same angle, but cancer cells are in various spots.

4. Fine-tuning the model: To optimize the proposed model and achieve the best result based on the given dataset, we change specific settings like the epoch, the batch size, or some hidden layers before starting to train and validate.
5. Adding hidden layers and dropout layers into the model: To avoid overfitting, we need to change the layer value to fit the pre-trained MobileNet model settings.
6. Training the model: Based on the three scenarios, we select the appropriate training sets from the ultrasound images dataset to train the model.
7. Validating and calculating the metrics: We check how many images have been correctly classified and calculate the metrics to compare the proposed model with other CNN architectures, importantly test accuracy and *F1*-score results.
8. Comparing with other CNN architectures: After training, validating, and testing the proposed model, we compared the results to original CNN architectures like VGG16, VGG19, InceptionV3, and MobileNet.
9. Showing results: After the data has been compared, tables and graphs will be displayed.

3.2 Proposed Model for Breast Cancer Prediction

This work has leveraged transfer learning techniques from the MobileNet and fine-tuning techniques to classify ultrasound images, including benign breast cancer, malignant breast cancer, and normal breast samples. The original architecture of the pre-trained MobileNet model is illustrated in Fig. 2.

Transfer learning Transfer learning applies knowledge learned from the previous model to the current problem. Transfer learning is used when our dataset has insufficient data to train a full-scale model from the beginning [21]. In our training process,

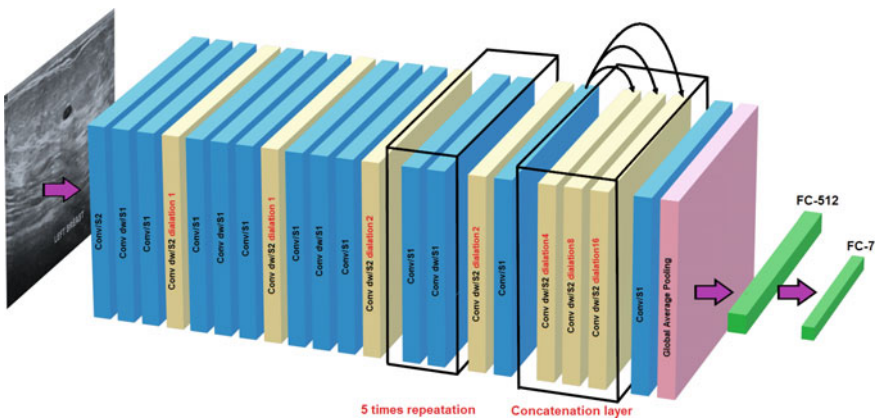


Fig. 2 MobileNet architecture to classify breast cancer

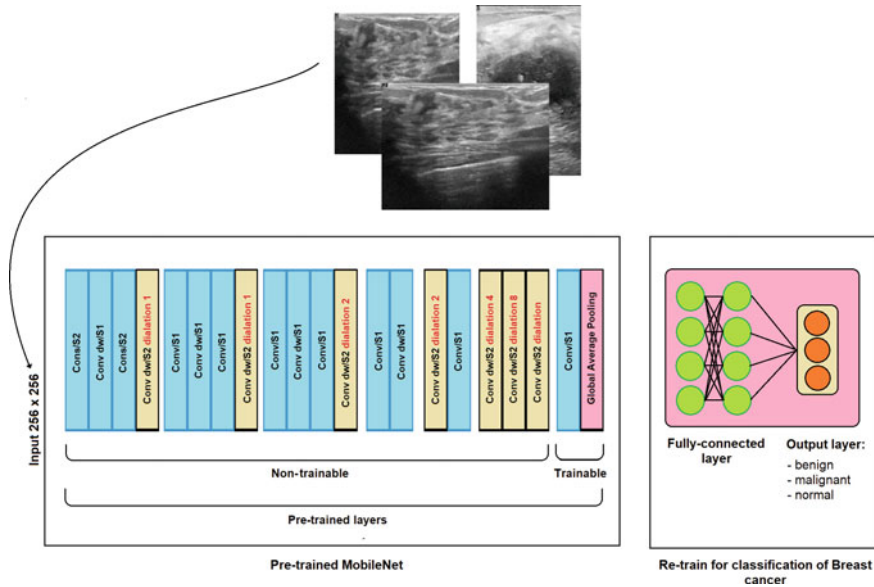


Fig. 3 Transfer learning model to classify breast cancer

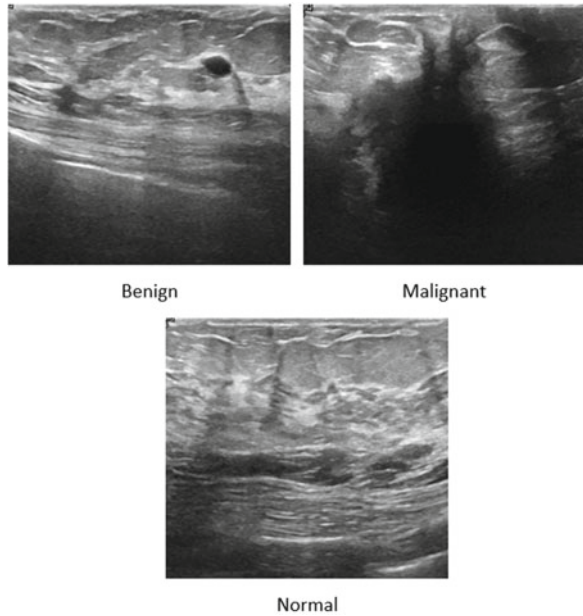
the previously trained MobileNet model parameters are reused. So the transfer learning will take advantage of the available layers of this model without having to retrain from scratch, helping to improve the model’s accuracy. For use in breast cancer classification problems, the MobileNet model will be added with a new fully-connected layer and an output layer with a softmax classification function [22].

Fine-tuning After applying the transfer learning, the result will be better if we continue fine-tuning. Fine-tuning uses a network model trained to perform a similar purpose for a given purpose. Typically, the model’s first layers are frozen. During the training process, the weight of these layers will not be modified [23]. These layers can extract low-level abstraction information; this ability was acquired during prior training. Freeze to take advantage of this ability and make training more effective. In the fine-tuning process, we continue to adjust the hyperparameters to help the model achieve the highest accuracy and also to avoid overfitting [24].

The hyperparameters used for the adjustment process are as follows: First, the number of training epoch ranges is tested with the values of 20, 25, and 30 to find a suitable epoch threshold. The number of batch sizes is tested with the value of 16–32–64. The hidden layer is tested with [256, 256, 64] and [512, 128], and the learning rate is tested with 0.001, 0.0001, and 0.00001.

Figure 3 illustrates the proposed model architecture applying pre-trained MobileNet for transfer learning and fine-tuning technique.

Fig. 4 Example of three categories of breast ultrasound images in the dataset



4 Experiment

4.1 Dataset

Breast ultrasound images from women aged 25 to 75 years old were obtained at the start of the study. This information was gathered in 2018 [20]. The total number of female patients is 600. The photos are saved as PNG files. The average size of ultrasound images is 500×500 . The ultrasound images are divided into normal, benign, and malignant categories. The data distribution follows 487 photographs for benign tumors, 210 images for malignant tumors, and 133 for normal tumors. Ultrasound scans are most commonly used to examine and detect breast cancer early. Furthermore, compared to other medical imaging procedures, it is incredibly safe. Figure 4 gives out an example of the three categories in the dataset.

4.2 Evaluating Method and Comparison

This research separates the dataset into three sets to evaluate the model: training, validation, and test sets to achieve a good result. To divide the dataset effectively, we randomly divide it to distribute data for these three sets. The most common golden ratios are 60–20–20, 70–15–15, and 80–10–10. Then, we evaluate through the model

Table 1 Number of images when splitting into sets in detail

The order of breast classifications	Types of set	Quantity (images)
1. Normal	Training	93
	Validating	20
	Testing	20
2. Malignant	Training	147
	Validating	32
	Testing	31
3. Benign	Training	305
	Validating	66
	Testing	66
Total		780

to choose the appropriate ratio, avoiding underfitting or overfitting because the dataset division does not have an even distribution for the given dataset. In this research, we apply the 70–15–15 ratio for splitting the dataset into training, validating, and testing sets. Table 1 illustrates the number of images when splitting the dataset into training, evaluating, and testing sets based on the ratio applied.

Then, we compare the findings with MobileNet, VGG16, Xception, and InceptionV3 using the accuracy metrics (accuracy—acc) and the *F1*-score. The research mentioned above was performed under the following three scenarios:

- Scenario 1 includes the following sequence: First, we get data from two categories which are normal and benign breast ultrasound images. Then, we perform training, evaluating, and testing on the proposed model. Finally, we compare the results between the proposed model and other CNN architectures. The purpose of conducting the first scenario is to check how our proposed model classifies two categories and the results of the experiments throughout test accuracy, *F1*-score, and the confusion matrix. We conduct this scenario because we predict and classify breast cancer, so we test the ability to classify benign and normal breasts to figure out and consider the given results. Then, we can evaluate the dataset as well as the proposed model.
- Scenario 2 includes the following sequence: First, we gather data from two breast ultrasound images: normal and malignant. Following that, we train, evaluate, and test the proposed model. Finally, we compare the proposed model’s results to those of other CNN architectures. The second scenario is being run to see how our proposed model classifies two categories and the results of the experiments in terms of test accuracy, *F1*-score, and confusion matrix. We conduct this scenario because we predict and categorize breast cancer, so we experiment with classifying malignant and normal breasts to determine and consider the given results. Then, we can evaluate the dataset and proposed model.
- Scenario 3 includes the following sequence: First, we collect data from three breast ultrasound images: normal, benign, and malignant. Following that, we train,

evaluate, and test the proposed model. Finally, we compare the proposed model’s results to various CNN architectures. The third scenario is being run to see how our proposed model classifies three categories and the results of the testing set in terms of test accuracy, *F1*-score, and confusion matrix. We conduct this scenario because we forecast and classify breast cancer. First, we test the ability to classify benign, malignant, and normal breasts to determine and examine the given ultrasound images dataset. Then, we can evaluate the dataset and proposed model.

We use Google Colab to perform these scenarios and evaluate the final results. All of the experiments and evaluations are performed and executed in the working session of Google Colab.

4.3 Scenario 1—Classify Benign Breast Cancer and Normal Breast

In this scenario, all models are run under the same hyperparameters: The number of epochs is 30, the learning rate is 0.0001, and the number of batch sizes is 64. We also propose adding a MobileNet model with additional parameters that can be modified. The hidden layer is adjusted to [256, 256, 64] for the proposed model only, and nine new hidden layers are added, including dense, batch normalization, and dropout layers. The outcomes of training and testing for scenario one are shown in Table 2.

This experiment shows that the transfer learning model with the proposed model gives the most remarkable results with the accuracy on the test set of 96.51% and the *F1*-score of 96.84%. Figure 5 illustrates the accuracy and loss during training of the proposed model in scenario 1. The confusion matrix of the first scenario is illustrated in Fig. 6.

From the illustration of graphs and confusion matrix in scenario 1, we can see that the training process of the proposed model achieves a promising result. Furthermore, the prediction results of two categories on the testing set achieved 98.0% accuracy on benign breast cancer and 90.0% accuracy on the normal breast. These results prove

Table 2 Results comparison when testing on the first scenario

Model	train_acc	val_acc	train_loss	val_loss	test_acc	test_F1
<i>Proposed model</i>	1.0000	0.9767	0.0399	0.0723	0.9651	0.9648
VGG16	1.0000	0.9651	0.0033	0.2037	0.8953	0.8898
VGG19	0.9950	0.9302	0.0228	0.1805	0.8837	0.8872
InceptionV3	1.0000	0.7907	0.0001	0.8452	0.8372	0.8012
MobileNet	1.0000	0.8023	0.0001	0.4404	0.7907	0.7176

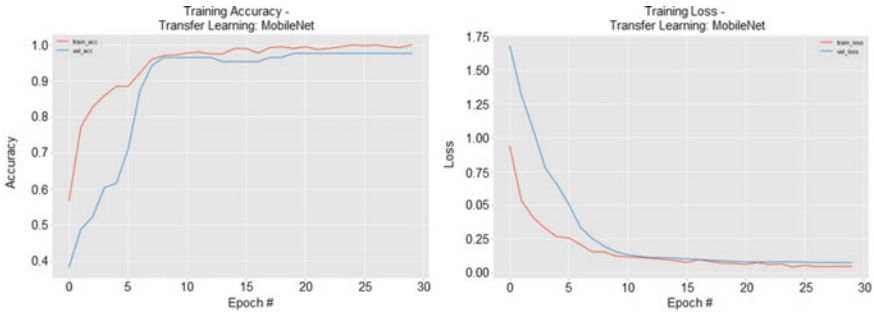
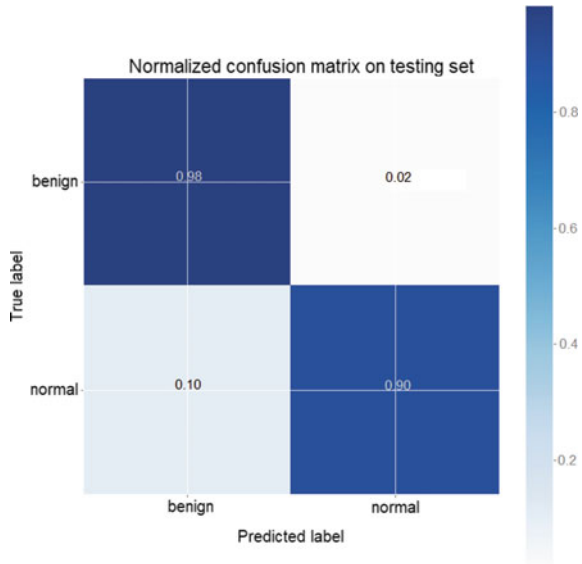


Fig. 5 Training accuracy and loss graph of the proposed model in scenario 1

Fig. 6 Confusion matrix illustration in scenario 1



that our proposed model can classify benign breast cancer better than normal breast; both results achieve a higher than 90%.

4.4 Scenario 2—Classify Malignant Breast Cancer and Normal Breast

All models are executed under the same circumstances as the first scenario. For example, Table 3 shows the results of training and testing for scenario 2.

This experiment shows that the transfer learning model with the proposed model still gives the best results, with an accuracy of 94.12% on the test set and an *F1*-score

Table 3 Results comparison when testing on the second scenario

Model	train_acc	val_acc	train_loss	val_loss	test_acc	test_F1
<i>Proposed model</i>	1.0000	1.0000	0.0130	0.0618	0.9412	0.9417
VGG16	1.0000	0.9808	0.0001	0.0675	0.9216	0.9216
VGG19	0.9750	0.9615	0.0757	0.2198	0.9216	0.9216
InceptionV3	1.0000	0.7308	0.0001	0.8480	0.8039	0.7906
MobileNet	1.0000	0.7885	0.0001	0.4168	0.7647	0.7334

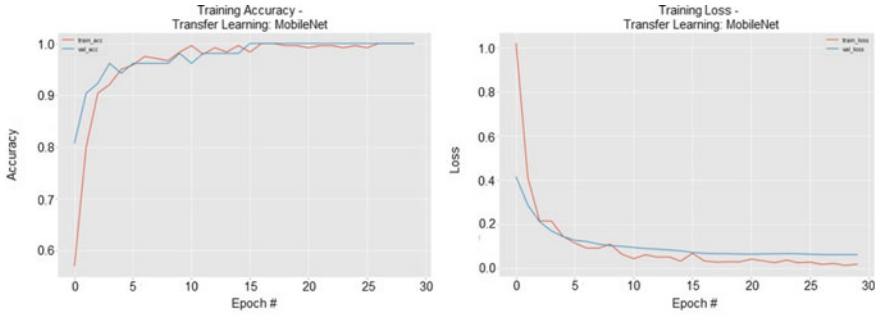
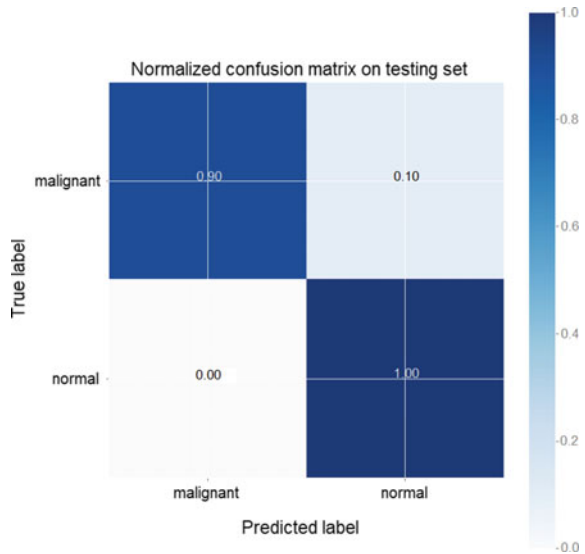


Fig. 7 Training accuracy and loss graph of the proposed model in scenario 2

Fig. 8 Confusion matrix illustration in scenario 2



of 94.17%. Figure 7 illustrates the accuracy and loss during training in scenario 2. The confusion matrix of the second scenario is illustrated in Fig. 8.

From the illustration of graphs and confusion matrix in scenario 2, we can see that the training process of the proposed model achieves a promising result. Moreover, the prediction results of two categories on the testing set achieved 90.0% accuracy on malignant breast cancer and 100.0% accuracy on the normal breast. These results prove that our proposed model can classify the normal breast better than malignant breast cancer, and both achieve a higher than 90%.

4.5 Scenario 3—Classify Benign Breast Cancer, Malignant Breast Cancer, and Normal Breast

All models are executed under identical conditions to the first and second scenarios in the third. Therefore, the outcomes of training and testing for scenario three are shown in Table 4.

This experiment shows that the transfer learning model with the proposed model gives the most significant results, with an accuracy of 90.6% on the test set and an *F1*-score of 90.58%. Figure 9 illustrates the accuracy and loss during training in scenario 3. The confusion matrix of the third scenario is illustrated in Fig. 10.

Table 4 Results comparison when testing on the third scenario

Model	train_acc	val_acc	train_loss	val_loss	test_acc	test_F1
<i>Proposed model</i>	0.9982	0.9322	0.0473	0.3237	0.9060	0.9058
VGG16	0.9963	0.8898	0.0108	0.5743	0.8291	0.8246
VGG19	0.9853	0.8814	0.0365	0.5874	0.8462	0.8428
InceptionV3	1.0000	0.7203	0.0002	0.8463	0.7265	0.7044
MobileNet	1.0000	0.7034	0.0001	1.4156	0.7094	0.6973

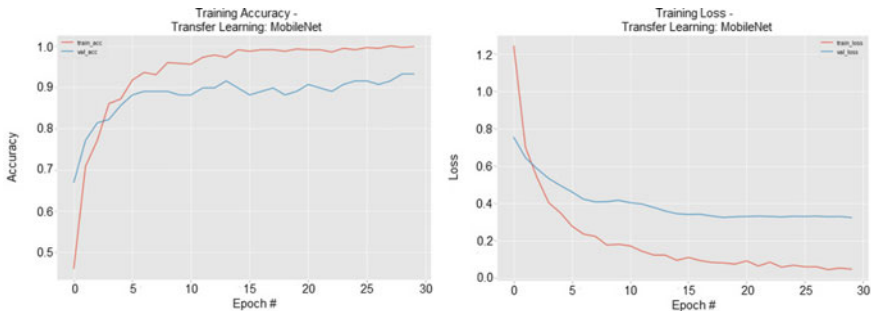
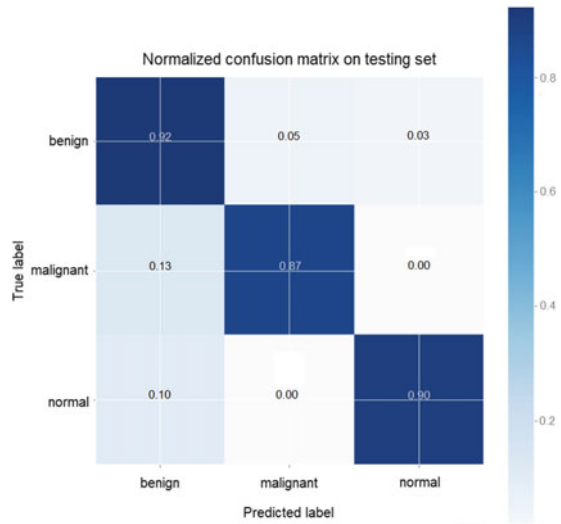


Fig. 9 Training accuracy and loss graph of the proposed model in scenario 3

Fig. 10 Confusion matrix illustration in scenario 3



From the illustration of graphs and confusion matrix in scenario 3, we can see that the training process of the proposed model achieves a promising result. Moreover, the prediction results of three categories on the testing set achieved 92.0% accuracy on benign breast cancer, 87.0% accuracy on malignant breast cancer, and 90% accuracy on the normal breast. These results prove that our proposed model can classify benign breast cancer better than normal breast and malignant breast cancer, and all of the results achieve higher than 87%.

4.6 Evaluating Experiment Result

All of the accuracy, validation accuracy, test accuracy, and $F1$ -score calculated of the proposed model in three different scenarios are shown in Fig. 11.

The experimental results show promising results when training, testing, and evaluating the proposed model in three scenarios suitable for breast cancer diagnosis. For example, in scenario 3, the proposed model achieves 90.6% accuracy when predicting three classifications, and the $F1$ -score is 0.9058 on the testing set. In addition, the accuracy score and the $F1$ -score achieve higher when eliminating one classification, with more than 96% accuracy in the first scenario and more than 94% accuracy in the second scenario.

In addition, the results are relatively positive compared to [12] with an accuracy of 87.2%, and in [13] with an accuracy of 87.0%.

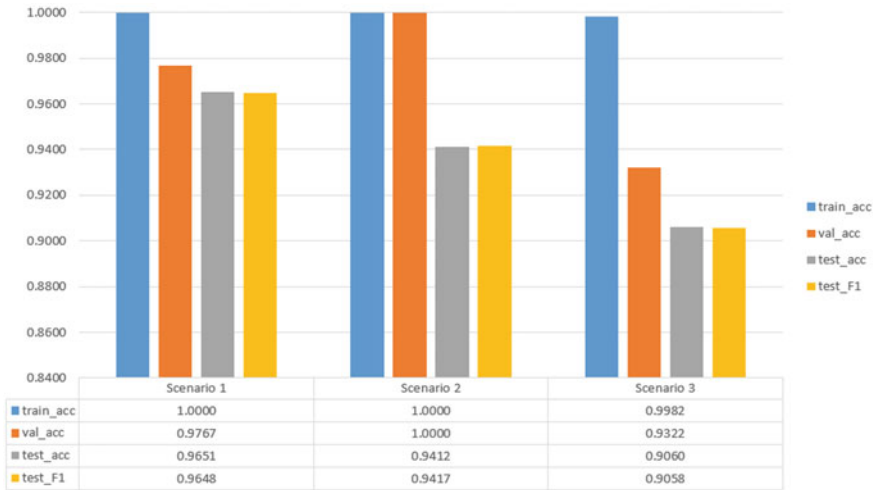


Fig. 11 Transfer learning model results on three different scenario experiments

5 Conclusion

Currently, there are many research articles in the medical field that have contributed significantly to the diagnosis of diseases that cause harmful consequences to patients’ health. This paper introduces a system for predicting and classifying breast cancer based on ultrasound images, including three benign, malignant, and normal breast classes. Applying transfer learning and fine-tuning to the pre-trained MobileNet model achieves high results with 90.6% accuracy. Moreover, the amount of time for training, evaluating, and testing these experiments is efficient based on the applied algorithms. We will continue to test various systems and collect more data on existing disease classes to improve model accuracy and apply visualization to explain predicting results.

References

1. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J, Cardoso F (2019) Breast cancer. *Nat Rev Dis Primers* 5(1):66
2. DeSantis C, Ma J, Bryan L, Jemal A (2014) Breast cancer statistics, 2013. *CA Cancer J Clin* 64(1):52–62
3. DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A (2015) International variation in female breast cancer incidence and mortality rates. *Cancer Epidemiol Biomarkers Prev* 24(10):1495–1506
4. Sun Y-S, Zhao Z, Yang Z-N, Xu F, Lu H-J, Zhu Z-Y, Shi W, Jiang J, Yao P-P, Zhu H-P (2017) Risk factors and preventions of breast cancer. *Int J Biol Sci* 13(11):1387–1397

5. Heck L, Herzen J (2020) Recent advances in X-ray imaging of breast tissue: from two- to three-dimensional imaging. *Phys Med* 79:69–79
6. Brattain LJ, Telfer BA, Dhyani M, Grajo JR, Samir AE (2018) Machine learning for medical ultrasound: status, methods, and future opportunities. *Abdom Radiol (NY)* 43(4):786–799
7. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) MobileNets: efficient convolutional neural networks for mobile vision applications
8. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition
9. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision
10. Chollet F (2016) Xception: deep learning with depthwise separable convolutions
11. Sun R, Meng Z, Hou X, Chen Y, Yang Y, Huang G, Nie S (2021) Prediction of breast cancer molecular subtypes using DCE-MRI based on CNNs combined with ensemble learning. *Phys Med Biol* 66(17):08
12. Ragab DA, Sharkas M, Marshall S, Ren J (2019) Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* 7:e6201
13. Alanazi SA, Kamruzzaman MM, Sarker MNI, Alruwaili M, Alhwaiti Y, Alshammari N, Siddiqi MH (2021) Boosting breast cancer detection using convolutional neural network. *J Healthcare Eng* 2021:5528622
14. Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ (2021) Classification of skin disease using deep learning neural networks with MobileNetV2 and LSTM. *Sensors (Basel)* 21(8)
15. Tareh MM, Zhu N, Ali TAA, Alghaili M, Hameed AS, Mutar ML (2021) KL-MOB: automated COVID-19 recognition using a novel approach based on image enhancement and a modified MobileNet CNN. *PeerJ Comput Sci* 7:e694
16. Vinayahalingam S, Kempers S, Limon L, Deibel D, Maal T, Hanisch M, Bergé S, Xi T (2021) Classification of caries in third molars on panoramic radiographs using deep learning. *Sci Rep* 11(1):12609
17. Olivas LG, Alférez GH, Castillo J (2021) Glaucoma detection in Latino population through OCT's RNFL thickness map using transfer learning. *Int Ophthalmol* 41(11):3727–3741
18. Novakovsky G, Saraswat M, Fornes O, Mostafavi S, Wasserman WW (2021) Biologically relevant transfer learning improves transcription factor binding prediction. *Genome Biol* 22(1):280
19. Zhang K, Xu G, Zheng X, Li H, Zhang S, Yu Y, Liang R (2020) Application of transfer learning in EEG decoding based on brain-computer interfaces: a review. *Sensors (Basel)* 20(21)
20. Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A (2019) Dataset of breast ultrasound images. *Data Brief* 28:104863
21. Falconí LG, Pérez M, Aguilar WG (2019) Transfer learning in breast mammogram abnormalities classification with MobileNet and NASNet. In: 2019 international conference on systems, signals and image processing (IWSSIP), pp 109–114
22. Kandel I, Castelli M (2020) Transfer learning with convolutional neural networks for diabetic retinopathy image classification. a review. *Appl Sci* 10(6)
23. Sutramiani NP, Suciati N, Siahaan D (2020) Transfer learning on Balinese character recognition of lontar manuscript using MobileNet. In: 2020 4th international conference on informatics and computational sciences (ICICoS), pp 1–5
24. Swati ZNK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, Lu J (2019) Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput Med Imaging Graph* 75:34–46

Identification of Leaf Disease Using Machine Learning Algorithm—CNN



P. V. Raja Suganya, A. R. Sathyabama, K. Abirami, and C. M. Nalayini

Abstract This study uses image processing techniques to analyze leaf properties for an agricultural field-based automated vision system. Autonomous leaf characteristics detection is an important part of agricultural research since it monitors large fields of crops and automatically detects the symptoms of leaf characteristic features on plant leaves. The proposed decision-making system makes use of image content characterization and supervised neural networks. For performing this type of decision analysis, image processing techniques include preprocessing, feature extraction, and classification stages. When an image is processed, it is scaled and a region of interest is selected if required. Color and texture features are extracted from an input dataset for performing network training and classification. This section includes color and texture characteristics such as energy, contrast, homogeneity, and correlation. The proposed system will classify the test images for analyzing and determining the leaf attributes. This approach can quickly and accurately classify objects by using a classifier NN, which learns from examples within a given category. The tangent sigmoid function is considered as the network's kernel function. According to a computer simulation, a network classifier is the best option used for attaining the highest training and classification accuracy.

Keywords Convolution neural network · Test database · Leaf disease

1 Introduction

The concern of effective plant disease protection is critical in terms of sustainable agriculture and climate change. According to the outcomes of the research studies, there is an evidence that the climate change influences pathogen development and

P. V. R. Suganya (✉)

Department of Artificial and Data Science, Velammal Engineering College, Chennai 600 066, India

e-mail: rajasuganya@velammal.edu.in

A. R. Sathyabama · K. Abirami · C. M. Nalayini

Department of Information Technology, Velammal Engineering College, Chennai 600 066, India

host resistance, resulting in physiological changes in host–pathogen interactions. The ease with which diseases can now be spread across continents and oceans exacerbates the issue. Diseases that were previously unknown and untreated could perhaps emerge in areas where there is no local expertise to combat them. Using pesticides without proper training can result in the development of long-term disease resistance by reducing the plant’s ability to defend itself significantly. Precision agriculture relies on timely and precise identification of plant diseases. Preventing wasteful loss of money and other resources is essential to achieve a healthy production by tackling long-term pathogen resistance development and minimizing the harmful consequences of climate change. In the rapidly changing world, the importance of early disease detection and prevention has been constantly increasing. Plant pathogens can be identified by using a variety of techniques. When there are no obvious signs of illness or when the effect is too late to be useful, a detailed testing is required. Since most of the diseases produce some type of visible manifestation, a skilled professional’s human eye examination is the primary method used for detecting plant illnesses in real time. One of the most important abilities for any plant pathologist is learning the ability to observe and identify specific plant disease symptoms. Symptoms of damaged plants can vary widely, and amateur gardeners and hobbyists may have a harder time identifying them than a plant pathologist. Amateur gardeners and skilled professionals could benefit from an automated system, which have been developed to assist in the identification of plant diseases based on their appearance and visual symptoms.

Consequently, convolution neural networks (CNNs) were developed and have produced excellent results when it comes to the classification of digital images. Innovative training methods and methodologies enable the system to be quickly and easily implemented in the real world. The developed model can distinguish between healthy leaves and leaves affected by 25 different types of plant diseases. This study analyzes the entire process of implementing a disease recognition model in real time, ranging from collecting images to creating a database by using CNN to identify the plant disease. The precision of the developed model ranges from 91 to 98.3% for various class tests with an average precision of 96.3%.

2 Literature Review

2.1 Diagnosis and Classification of Grape Leaf Diseases Using Neural Networks

Plant diseases have a large impact on crops, causing significant financial losses. Alternatively, early detection of plant diseases results in a significant improvement in the quality of final product. Pesticides are applied incorrectly when the disease and its severity are misdiagnosed. Image processing and artificial intelligence techniques will be used to diagnose the disease from the image of grape plant leaves [1]. The

proposed method takes an image of a grape leaf with a complex background as input. Anisotropic diffusion is used to reduce noise in the image after masking green pixels with thresholding. K -means clustering is then used to segment the diseases present in grape leaves. The diseased region is then identified by using the process of image segmentation. The best results were then obtained by training a feed forward back propagation neural network for performing classification [2–4].

2.2 Advances in Image Processing for Plant Disease Detection

In this research, a software solution for the automatic identification and categorization of plant leaf diseases is proposed and experimentally evaluated. Studying the visual patterns of a certain plant is referred to as plant trait/disease-based research. Crops continue to encounter a diverse set of diseases. Physical damage is one of the most common diseases caused by insects. Since some bird species are poisoned by insecticides, their effectiveness is not always demonstrated. It also harms natural food systems. Following the segmentation phase, the following two phases are added in sequence. They look for pixels that are predominantly green in color in the first stage. The pixels that are predominantly green are then masked using Otsu's approach, which computes specified threshold values. The next step was to remove all of the pixels with zero red, green, or blue values, as well as those that bordered the infected cluster (object). In the experiment, proposed methods for detecting plant diseases have been shown to be effective. The newly developed algorithm can detect and categorize the analyzed diseases with a precision in the range of 83–94%, and it can do so 20% faster than the approach provided in the study [5–7].

2.3 Detection and Classification of Plant Diseases by Image Processing

The main purpose of this research is to develop an image processing software for identifying and classifying plant leaf diseases. However, relying entirely on professionals' manual inspection to diagnose and classify diseases in rural areas and underdeveloped countries can be time consuming. As a result, the proposed image processing-based system is highly time efficient, self-contained, cost effective, and precise. In the first step of the solution, a color space transformation is done to the RGB leaf image followed by a color transformation structure. In the second stage, the images are segmented by using the K -means clustering technique. In the third phase, the textural properties of the segmented infected objects are calculated. In the fourth and final phase, the collected features are fed into a neural network, which is already trained [8–10].

2.4 Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques

This proposal describes a cutting-edge computing system to assist farmers in making better decisions based on various aspects of the agricultural development process. Crop disease detection and evaluation are critical for increased yields. Foliar is the most common fungal disease that affects cotton crops in India, and it can be found in all growing regions of the country. In this paper, new technological solutions based on mobile-captured cotton leaf spot symptoms are presented, and the diseases are then classified by using the proposed HPCCDD algorithm. In order to develop intelligent farming practices, the classifier is being trained to identify diseases in the groves, apply fungicides selectively, and so on. In order to identify diseases, image RGB feature ranging techniques (using ranging values) will be utilized to enhance the collected images initially. After that, the color image segmentation is used to find the desired areas (disease spots). These collected features are then used for classification purpose to identify the disease-affected areas by using different techniques such as sobel and canny filtering [11–13].

2.5 Plant Disease Identification Based on Deep Learning Algorithm in Smart Farming

In a complex environment, preventing plant disease begins with precisely detecting the disease. Due to the increasing practice of intelligent farming, sophisticated decision assistance, smart analysis, and smart planning are now available via digitalization and data-driven decision support. The development of a deep learning-based mathematical model for the detection and recognition of plant disease is proposed in this study. The region proposal network (RPN) is initially used to identify and locate the leaves in a range of difficult settings (Fig. 1).

Using the Chan–Vese (CV) method, images segmented based on RPN results contain different characteristics of disease symptoms. Transfer learning models are used to train these segmented leaves by using a dataset of sick leaves in the absence of any other context. Black rot, bacteria, and rust illnesses are also tested on the model. The accuracy of the proposed methodology is better than the existing method by 83.57%, which decreases the impact of illness on agricultural productivity and supports the long-term agricultural growth. Furthermore, deep learning algorithm has a big impact on intelligent agriculture, environmental protection, and agricultural productivity [14–15].

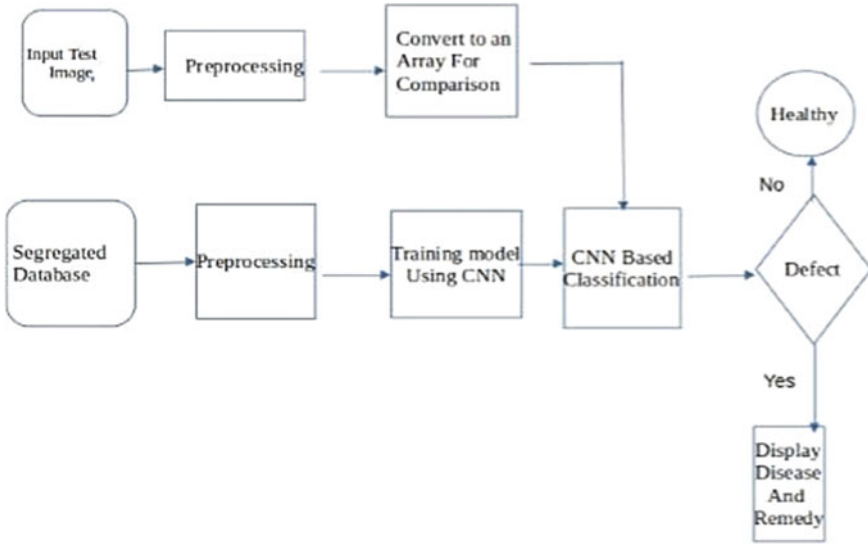


Fig. 1 Flowchart for disease detection

2.6 Detection and Classification of Plant Leaf Diseases by Using Deep Learning Algorithm

Insects and damaged plant leaves pose a serious threat to agriculture. Faster and more accurate predictions of crop leaf diseases could lead to the development of an early treatment method and a significant reduction in economic loss. Recent advances in deep learning have enabled researchers to significantly improve the performance and accuracy of object identification and recognition systems. The proposed study has used a deep learning-based technique to detect leaf illnesses by utilizing the images of plant leaves. The primary objective is to discover and build deep learning approaches that are better suited to the proposed work. This study has used the single shot multi-box detector (SSD), which is a region-based convolution neural network (faster R-CNN). The suggested system is capable of accurately detecting a wide range of diseases in plants and can handle scenarios that are both simple and complex.

3 Proposed System

3.1 Module 1

In the next stage, the input test image is acquired and preprocessed before being translated into an array form for comparison. The following procedures are used to preprocess the provided image:

- Import the required libraries. Here, the Keras library is used for developing a proposed model and training it.
- Loading the data.
- Visualize the data.
- Data preprocessing and data augmentation.
- Evaluating the result.

3.2 Module 2

The database is properly segregated and preprocessed and then renamed into proper folders. Obtaining a valid database is the first step in any image processing research. When possible, it's best to use a standard database; however, this is not always possible, so the images may be collected to build new database in these situations. In order to begin, the database must be purged and labeled. The images with the best resolution and angle are chosen from a large database. Prior to image selection, a thorough understanding has been gained on various leaves and their diseases. The plant village organization's repository is extensively used in this research. Further, a variety of plant images are examined and correlated. After a thorough investigation, the images are then categorized and labeled based on different disorders.

The model is properly trained by using CNN and then classification takes place. The CNN model is preprocessed and trained as follows: The database is preprocessed, including image shape, resizing, and array conversion. In the same way, the test image is processed. A database of approximately 500 different plant leaves is compiled, from which each image can be used as a software test image. The train database is used to train the model (CNN) to recognize the test image and the disease it contains. Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D are some of the layers in CNN. The algorithm can identify the sickness if the plant species is in the database, and the model has been properly trained. In order to make an accurate diagnosis, a comparison between a test image and the trained model must be made after training, and preprocessing has been completed successfully.

Convolution Neural Network:

Convolution: Image can be de-convolved in order to identify specific features in the image. When applied to an image, convolution can blur, sharpen, detect edges,

reduce noise, and perform a variety of other things that help the computer in learning about the image.

Pooling: When image is too big, it needs to be shrunk. The primary purpose of pooling is to minimize the size of an image without losing any of its details or patterns.

Flattening: For neural networks and classifiers, flattening turns a two-dimensional matrix into a vector of features.

Full Connection: Sending the image into a neural network is known as full connectivity.

The proposed convolution neural network (CNN) is built with the help of the Keras and TensorFlow frameworks.

The extraction of features from the entity set is a critical part of addressing any problem using machine learning. In image processing, the feature set essentially consists of each individual pixel that makes up the image. The features will be available based on the image's resolution and size. The image's color mode determines how many pixels are present in a megabyte.

In 8-bit (256 colors) picture, there are 1,048,576 or 1024×1024 pixels in one megabyte.

In 16-bit (65,536 colors) picture, one megabyte contains 524,288 (1024×512) pixels.

In 24-bit RGB (16.7 million colors) picture, one megabyte has approximately 349,920 (486×720) pixels.

In 32-bit CYMK (16.7 million colors) picture, one megabyte has 262,144 (512×512) pixels.

In a 48-bit picture, one megabyte has only 174,960 (486×360) pixels.

3.3 *Module 3*

A comparison between the test image and the trained model is performed. Additionally, a convolution basis is used to extract features. The classifiers will be trained to detect images with healthy or sick foliage by using these attributes.

3.4 *Module 4*

If there is a defect or disease in the plant, the software displays the disease along with the remedy. A GUI is used to display the type of disease, and the type of pesticide to be used is specified in the command prompt. For using GUI, tkinter package is used.

In 32-bit CYMK (16.7 million colors) picture, one megabyte has 262,144 (512×512) pixels.

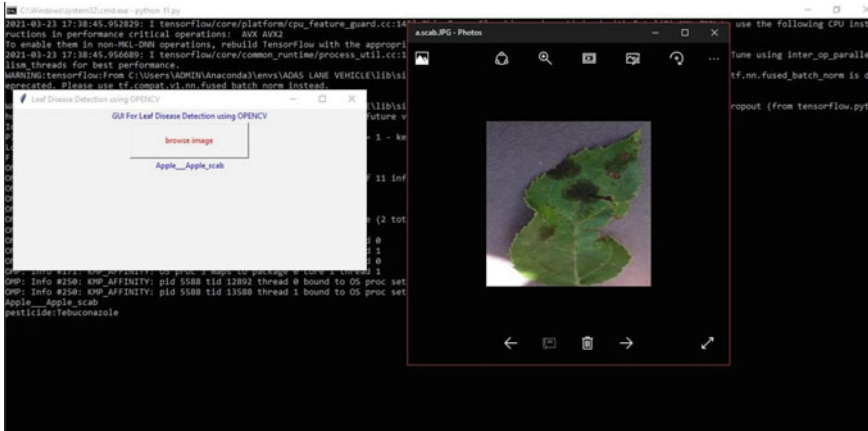


Fig. 2 Result for classification

In 48-bit picture, one megabyte has only 174,960 (486×360) pixels.

3.5 Module 5

If a defect or disease is observed in the plant, the software displays the disease along with a remedy. A GUI is used to display the type of disease and pesticide to be used in the command prompt. For using GUI, tinder package is used.

4 Results and Discussion

4.1 Screenshot of Final Result

The final result is shown in Fig. 2.

5 Conclusion and Future Enhancement

5.1 Conclusion

The proposed method has been devised for farmers and agricultural industry. From the results, it is evident that the proposed system has successfully detected the plant disease and recommended a treatment. It is also possible to improve the plant's

health by learning extensively about both the problem and corresponding solutions. The proposed system is coded in Python and has generated a 90% accuracy rate. The Goggles GPU has also been used to improve processing accuracy as well as speed.

5.2 Future Enhancements

The primary goal of future work will be to develop a server-side component by including a trained model as well as an application for smart mobile devices, which are capable of displaying diagnosed diseases in fruits, vegetables, and other plants. Regardless of experience level, this software will assist farmers in quickly identifying the plant diseases and making knowledgeable decisions about the use of chemical pesticides. This technology could be used to detect the entire plant rather than just individual leaves. When the number of features extracted is high, the results will be more accurate. Drones and unmanned vehicles can also be embedded with the system so that the agriculture fields can be monitored from air.

References

1. Kumar S, Srilekha A, Suman T, Rao BP (2021) Leaf disease detection and classification based on machine learning 3(2)
2. Guo Y, Zhang J, Yin C (2020) Plant disease identification based on deep learning algorithm in smart farming 1(2)
3. Akila M, Deepan A (2018) Detection and classification of plant leaf diseases by using deep learning algorithm 6(7)
4. Amoda N, Jadhav B, Pradnyakurle, Kunder S (2018) Detection and classification of plant diseases by image processing 2(6)
5. Ranjan M, Weginwar MR, Joshi N (2015) Detection and classification of leaf disease using ANN 3(3)
6. Sannaki S S, Rajpurohit VS, Nargund VB, Kulkarni P (2013) Diagnosis and classification of grape leaf diseases using neural networks 5(4)
7. Naikwadi S, Amoda N (2013) Advances in image processing for detection of plant diseases 2(11)
8. Wang H, Li G, Ma Z, Li X (2012) Image recognition of plant diseases based on PCA and neural network 3(1)
9. Revathi P, Hemalatha M (2012) Classification of cotten leaf spot diseases using image processing edge detection techniques 2(4)
10. Camargo A, Smith J (2008) An image-processing based algorithm to automatically identify plant disease visual symptoms 5(10)
11. Ampatzidis Y, De Bellis L, Luvisi A (2017) iPathology: robotic applications and management of plants and plant diseases 9(6)
12. Pantazi XE, Moshou D, Tamouridou AA (2019) Automated leaf disease detection in different crop species through image features analysis and one class classifiers 56(4)
13. Barbedo JGA (2018) Factors influencing the use of deep learning for plant disease recognition 172(7)
14. Geetharamani G, Pandian AJ (2019) Identification of plant leaf diseases using an in-layer deep convolutional neural network 76(8)

15. Konstantinos PF (2018) Deep learning models for plant disease detection and diagnosis 145(10)
16. Singh V, Misra AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques 4(1)
17. Mohanty SP, Hughes DP, Marcel S (2016) Using deep learning for image-based plant disease detection 7(2)
18. Amanda R, Kelsee B, Peter MC, Babuali A, James L, Hughes DP (2017) Deep learning for image-based cassava disease detection 8(4)
19. Ali H, Lali MI, Nawaz MZ, Sharif M, Saleem BA (2017) Symptom based automated detection of citrus diseases using color and textural descriptors 138(7)
20. Pujari D, Yakkundimath R, Byadgi AS (2013) Grading and classification of anthracnose fungal disease of fruits based on statistical texture features 52(9)
21. Iqbal Z, Khan MA, Sharif M, Shah JH, ur Rehman MH, Javed K (2018) An automated detection and classification of citrus plant diseases using image processing techniques: a review 53(2)
22. Turkoglu M, Hanbay D (2019) Leaf-based plant species recognition based on improved local binary pattern and extreme learning machine 527(5)
23. Zhang S, You Z, Wu X (2019) Plant disease leaf image segmentation based on super pixel clustering and EM algorithm 31(5)
24. Patil JK, Kumar R (2016) Analysis of content based image retrieval for plant leaf diseases using color, shape and texture features 10(2)
25. Yang N, Qian Y, EL-Mesery HS, Zhang R, Wang A, Tang J (2019) Rapid detection of rice disease using microscopy image identification based on the synergistic judgment of texture and shape features and decision tree-confusion matrix method 99(14)

BIO-VR: Design and Implementation of Virtual Reality-Based Simulated Biology Laboratory Using Google Cardboard with an Emphasis on Virtual Education



Towfik Ahmed, Afzal Un Nayeem Chowdhury, and Ziaul Hasan Mozumder

Abstract Virtual reality (VR) has become one of the most emerging technologies of this decade. Researchers can now exchange and study data like never before with the help of virtual reality-based tools. It's a key pipeline study in the field of human–computer interaction. In this proposed method, we created a virtual reality-based immersive simulated biology laboratory experience that allows students to learn different laboratory equipment. The immersive method for learning outcomes in the biology laboratory aimed to address the global problem where the underprivileged students lack the access of even the basic laboratory equipment. The proposed method was tested and surveyed among high school students coming from an underprivileged geography specifically in Bangladesh. The substantial finding of the study reflects how these individuals can be benefitted from this virtual reality-based biology laboratory. The results came out positive and along with that the ending calls for future works on immersive VR laboratories on other streams for high school students.

Keywords Human–computer interaction · Virtual biology laboratory · Virtual simulation · Virtual environment

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Towfik Ahmed, Afzal Un Nayeem Chowdhury, Ziaul Hasan Mozumder. The first draft of the manuscript was written by Towfik Ahmed, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

T. Ahmed (✉) · Z. H. Mozumder
Department of Computer Science and Engineering, Leading University, Sylhet, Bangladesh
e-mail: towfikahmed09@gmail.com

Z. H. Mozumder
e-mail: ziamozumder430@gmail.com

A. U. N. Chowdhury
Department of Computer Science and Engineering, Metropolitan University, Sylhet, Bangladesh
e-mail: nayeem.komol7@gmail.com

1 Introduction

The oversimplification of the human task through simulation and enhancement through precision has led the VR technology to gain ground in multifarious use case scenarios. Immersive virtual reality integrated applications have been developed which can replicate real-life situations and environments giving the users a real-life simulation that makes them enjoy, learn, and get a state-of-the-art virtual world experience right from their home whether it be online or offline. During the COVID outbreak, VR applications have come forward to solve many problems and give virtual reality platforms where users can come and have a shared space that will be able to give them a necessary simulated environment while social distancing right from their home. Due to the mass adoption of VR in recent times, it was able to penetrate the fields of education and learning [1], medical sciences, co-working space, travel, games and entertainment, etc. In the case of education, precisely in practical learning, the role of laboratory work is immense. High school laboratory works are very important for students to develop their skills and exhaustively understand a topic. According to Hodson [2], laboratory works include planning, collection of data, fact checking, sorting, and later coming to terms with the very own interpretation as well as a conclusion to the study. This makes the student more immersed in the process of learning and getting an understanding of scientific inquiry. Similarly, in learning biology, students must go through some rigorous practical work and do experiments as per their curriculum provided in their respective biology laboratories in their school to get into the actual scientific practice [3]. But in real-life scenarios in underdeveloped countries, practical biology learning is avoided due to the lack of scientific instruments, specimens, and overall infrastructure of schools, though all are included in their curriculum. The experiments in the biology laboratory are outright costly [4]. The virtual biology laboratory developed in this paper aims to solve this particular problem specifically for the underprivileged high school students who have biology laboratory work in their syllabus. The use of technology is widespread in education but all of these advances and focus are concentrated on higher education. Virtual reality laboratories are helping college students to get more out of their biology practical learning and through experiments; it has been proven to be more effective, and also the students that have used the learning simulation have expressed a positive attitude toward it [5]. All these can be accessed right from home using the web. Extensive research is being made to bring the use case of virtual reality for learning purposes, especially in the fields of science and technology [6]. Mostly in the case of practical learning, the use of technology can increase the learning output for students as well as increase the critical thinking capability through reasoning as well as making the learning engaging and entertaining [7]. Since the biology study materials require precise observation and examining intricate details of the specimen, the VR biology laboratory environment should facilitate enhanced image processing techniques in order to get sharp and better quality feed. The works of [8] proposed such an algorithm which eradicates the problem of low-quality image and provides a better user immersion experience.

Can such curriculum-based virtual biology laboratories also have a positive impact on the high school students of the disadvantaged geographical and economic location was the purpose for this paper. Also due to the COVID-19 outbreak, students lacked access to their biology laboratories but their theoretical classes were going on through online mediums. To minimize the gap between practical and theoretical scientific learning progress, the virtual biology laboratory could help. To understand if it was possible to bring better learning in practical works through virtual reality using an affordable and accessible medium like Google Cardboard and cheaper mobile devices among underexposed to virtual reality users was the main objective. The experiments and the survey proved to be positive and through further development works, it could bring a newer dimension to the learning process. The proposed method was developed and modified to address how low-cost VR headset has the potential to learn about biology laboratory in immersive way during pandemic maintaining social distance.

2 Related Works

2.1 VR in Education

It is found out in the paper [9] that it is necessary for students to be able to interact with real-life laboratory equipment and participate in laboratory experiments but that must be through a pedagogical structure to provide the best output. This curriculum-based work was done in [10] where students were given curriculum-based laboratory experiments, and they were carried out in traditional laboratory experimentation, computer-based experimentation, and VR-based experimentation. It was found that the students' learning efficacy increased during their virtual learning process. Similar works in the fields of journalism, paramedicine, and media production through mobile VR by the process of generating pedagogical content which is user generated [11]. Here through ubiquitous computing and through social interaction of the learners, a mobile VR platform enabled the learners to increase their efficacy in gaining practical knowledge of a subject. One of the key platforms which facilitates multi-user interaction in a VR simulated environment in learning is Second Life [12]. Second Life lets learners and teachers join an immersive virtual world where they could interact and have a common goal to achieve better efficiency in learning outcomes. This form of learning reflects the prevalent curriculum to ensure that pedagogy is kept intact, and students get a better environment to learn their lessons. A prototypical "virtual physics science laboratory" was developed that permits understudies to control the laboratory environment just as the actual objects in that laboratory facility [13], where they experimented measuring the time of the pendulum for various lengths furthermore, various sizes of gravity. Measure the normal pace of loss of energy of a ball brought about via air drag when dropped from various statures, and under various gravitational speed increases. Analyze the directions (particularly range and

greatest stature) of an item projected in two measurements with and without environmental drag. They found out this virtual physics laboratory licenses understudies to communicate with a convincing visual and acoustic “world”. There has been a shift of paradigm in the education sector due to the widespread and easy to access technologies but along with the rise of emerging technologies, there can be a lot of withdrawals from studying by the students if proper learning resources compatible with the tech is not provided [14]. The orthodox methods of learning are being replaced with a set of parameters and variables where the main theme is that all the learners are not equally treated when imparting knowledge rather the concept of flexible timing and self-paced learning curves are more prioritized where interactive audio–visual conceptual environment is more preferred for better output [15]. The young elementary students are found to be more excited and they tend to perform better with their reasoning skills if they learn something through enjoying the medium through which they are taught. The elementary students were given a science-based educational class through the VR technology and by gamification of the content to make it interactive. It is found out that their reasoning ability and problem-solving skills have increased to an extent [16, 17]. Geography was taught to students with the help of VR-Engage, a virtual reality simulation game based on levels to interact and compete and in turn have learnt a portion of geography based on prescribed curriculum. The key point was to make the game more likeable and usable instead of making it like regular hit games. It was a success to make the users get a grasp of the concepts of geography irrelevant to their gaming skills [18]. To promote curriculum effectiveness, an intuitive VR platform, virtual reality-based education expansion (VREX) came up with an O2O VR classroom where students get to learn abstract concepts of a subject where on the other hand, it is tough for the teacher to make the student understand with orthodox learning model. It gave positive results both online and offline and ultimately uplifts the total VR experience in any learning environment.

2.2 VR in Medical Science

A social virtual reality platform was proposed and implemented which facilitated patients to consult with medical professionals for their orthopedics needs. The proposed social VR clinic aimed to cut the volume of work a clinical staff goes through and along with that reduces the traveling hassle of patients [19]. The model was proven to be effective through taking responses from both end users, the medical consultants as well as the patients. A human–computer interface with high-end components both hardware and software creating VR simulated laparoscopic surgery environment for novice surgeons to practice as well as professional surgeons to carry on R&D. This VR simulator for minimally invasive surgery is a gateway for surgeons to become more efficient with their surgery procedures with least collateral [20]. Similar to this simulation to help in medical training, there are also various use cases of VR-based simulation environments which expedite medicine practitioners to pick

up first-person experiences and learn in an engaging interactive set which in real life would have been not feasible [21]. Health professionals were given a mixed platform where VR and AR was integrated through Microsoft's HoloLens in ensuring the curriculum of the faculty staff and students, and students could run smoothly in order to gain that real life like simulated experience to boost learning outcome [22]. The holographic mixed reality platform used head mounted display so that the overall outcome of the experiment has a few side effects as possible for the users availing the experience. Another problem for medical science students on anatomy is that they do not get adequate supply of wet specimen for the required amount of time to learn more as it is not possible to examine the live sample during surgery and so they must rely on their past knowledge alone [23]. To solve this problem, and to give a demonstration of the anatomy to the students from all angles an interactive learning content for a VR, AR and a tablet were created where three participant groups participated and they rated the experience as very good, and VR technology will actually help them in their anatomy academic learning [24]. VR has been found to assist surgeons in critical operation procedures which need visual accuracy and practice to get better output. A review has been made by analyzing several VR simulations and their aftermath results produced by the users of such simulations in flexible endoscopic skills and also laparoscopic surgeries [25]. The results suggest that the VR experience modeled with surgery courses and its framework's efficacy has proven to be a cut above [26]. VR as a medical science teaching medium has been gaining momentum. One of the most challenging topics in medical science is cardiac anatomy, and the teachers are having a hard time getting the base layers and retrospectives with the orthodox teaching methods. But a 3D modeled heart gave access to easier understanding; the students were found to learn with better accuracy [27]. A cutting edge proposed method was showcased in [28] where clinicians can provide PTSD patients with medically tested virtual immersive environments through VR technology based on the patient's individual requirements, and it was proven to help the therapy session to a large extent.

3 Methodology

The proposed method is created with an app named "VR-Biology". As a virtual reality headset, Google Cardboard was used. It is recommended to use Android 4.4 "KitKat" (API level 19 or above) for VR compatibility support. The method was created using the Unity3D game engine and the Blender 3D modeling software. The default platform was pc which was later on changed to Android in the build setting inside the game engine. Vulkan graphics API was disabled since it doesn't support for Google Cardboard VR for mobile. The user game object had main camera attached to it for providing stereoscopic view. The main camera inherits class interface from GvrRecticlepointer prefab to give a virtual look for the environment, which is done through sdk provided by the Google Cardboard by default pipeline inside the game engine (Figs. 1 and 2).

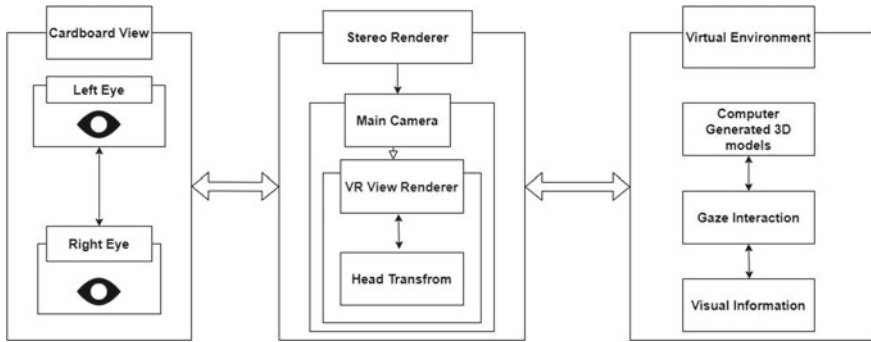
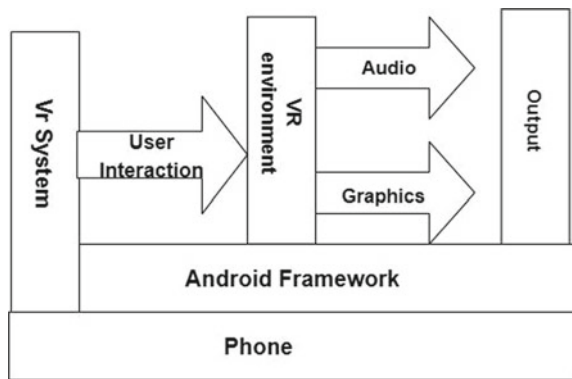


Fig. 1 Stereoscopic rendering workflow

Fig. 2 VR system architecture



User interacts with UI elements using gaze base interaction. Main camera continuously raycast invisible ray to detect if any interactable objects in the virtual environment through GvrRecticlepointer prefab. GvrPointPhysicsRaycaster continuously raycasts, a raycast which is an invisible ray that allows to select objects which is interactable if the object has a box collider attached to it. GvrRecticlepointer prefab is an image that changes its shape when it detects an interactable object. For interaction, each objects had a component called box collider attached to it. Interaction in VR is called gaze-based interaction. If any raycast collides with objects having box collider, the object is interactable (Fig. 3). Specific objects which were meant to be interactable such as microscopes, skeleton, 3D cells, these objects had box collider for interaction. The user had character controller component attached. The character controller’s radius and height were adjusted and calibrated for better users’ reachability. A csharp script was created “movement.” The script was designed in a way for the user to roam around when they look down and positional axis y is below at 45° angle (Fig. 4). The proposed method offers a biology laboratory in a virtual environment. The method offers user to roam around the laboratory and can interact with objects. One of the key features of this method is it offers two virtual

microscopes. Both microscopes have an information image above them. The virtual microscope provides 360 videos of two types of cell division Mitosis, Meiosis and also two types of educational video “White Blood Cell chasing Bacteria,” “Amoeba Under the Microscope” these are widely used in biology laboratory in high school in Bangladesh. User can select the microscope, and panel will show up. The panel consists four options and the user can select which type of video they want to watch. For four VR videos, there are four different scenes. User can decide which VR videos they want to watch. These videos are VR enabled. VR system architecture is shown in Fig. 2.

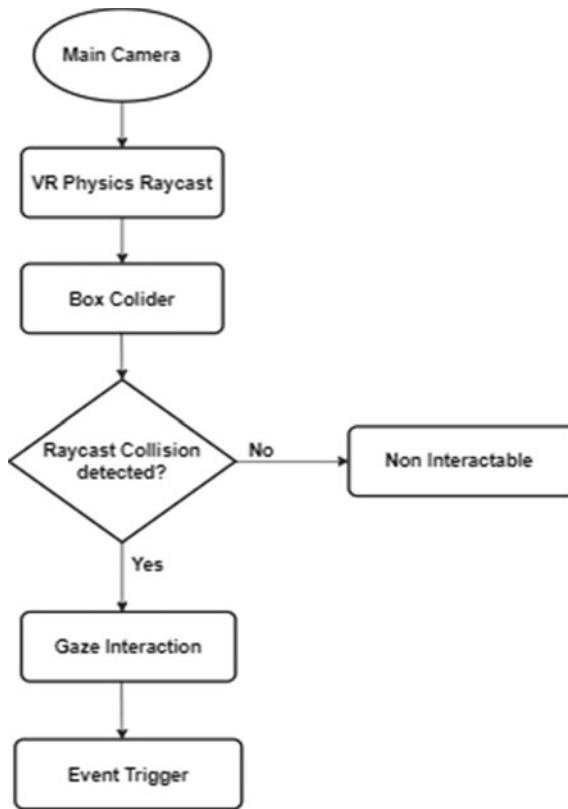


Fig. 3 Interaction workflow

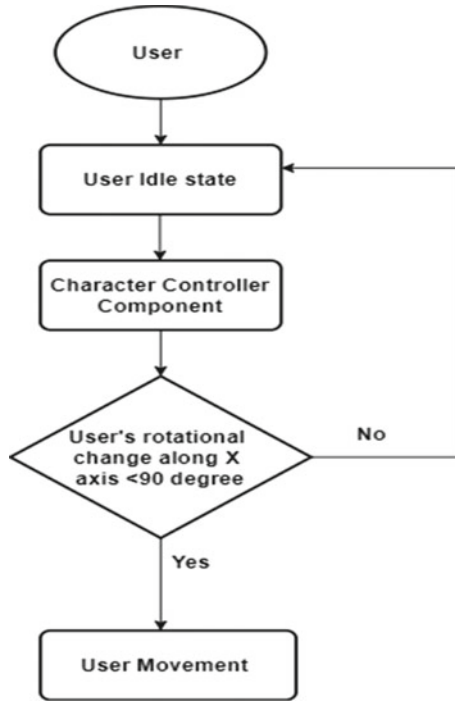


Fig. 4 User movement

Under the microscope, the two buttons were provided for zoom in and zoom out for the videos. Another microscope offers four types of 3D cell which are “Animal cell,” “Bacteria Cell,” “Fungal cell,” and “Plant cell”. The 3D models are also interactable as user can see the different names of the cell structures. The virtual laboratory also consists an interactable skeleton. Upon selecting any part of the skeleton, the name of the bone will appear. Add on features of the proposed model is shown in Fig. 5.

Features	Descriptions
Easy Interaction	The proposed method offers easy gaze-based interaction, which is easy for user to interact with virtual objects
Affordable	The vr headset used in the proposed method is affordable rather than HTC Vive, Oculus quest
User Friendly Navigation	The proposed method offers user friendly ui experience
Less side effects	During post assessment survey no side effects e.g. eye strain, headache was seen

Fig. 5 Add on features of the proposed model

3.1 Virtual Biology Laboratory

The virtual biology offers an immersive experience for the students. It has two microscopes (Fig. 6). Both of the microscopes have different functionalities. If the user interacts any of the microscopes, a panel will show up (Figs. 6 and 7). The panel consists of information about what the user wants to see which includes 360-degree video (Fig. 8). Another microscope offers which specimen the user wants to see. The 3D cells are also interactable. If the user selects any of the part of the cell, its corresponding name and its information will appear (Fig. 9). The virtual laboratory also contains a skeleton. The skeleton is also interactable. If the user interacts with any part of the skeleton, its name will show up (Figs. 10 and 11).

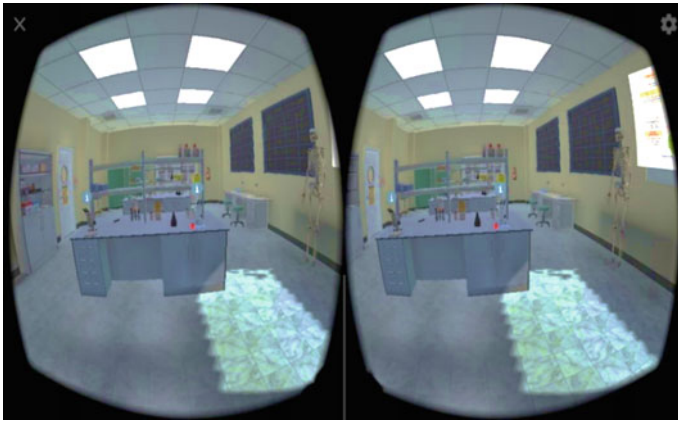


Fig. 6 Virtual view of the laboratory

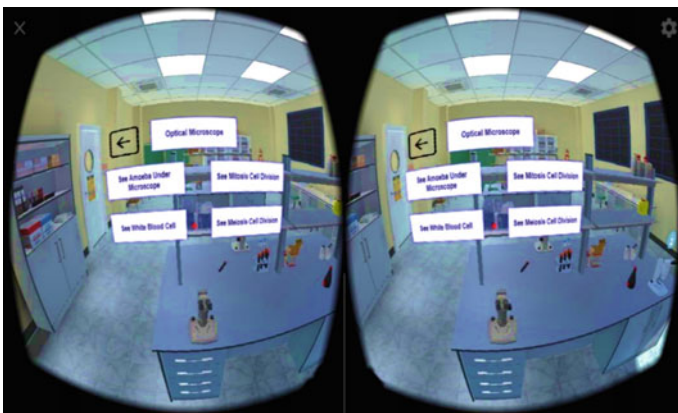


Fig. 7 Microscope panel

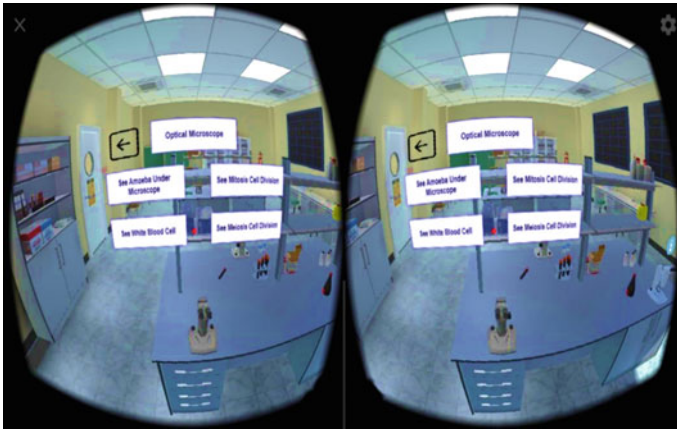


Fig. 8 Microscope panel for 3D cell specimen

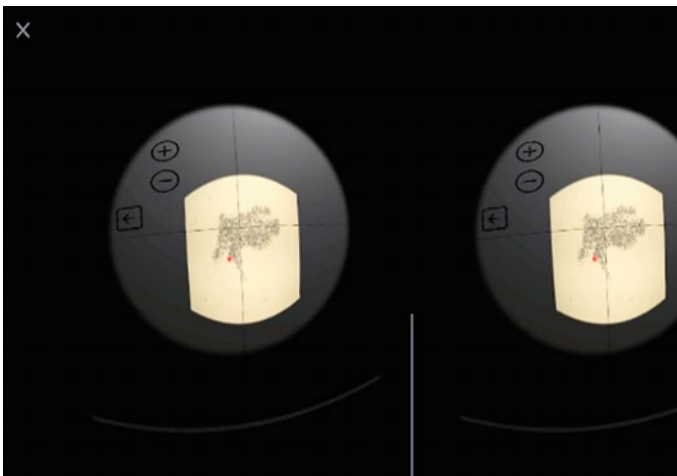


Fig. 9 360 video of amoeba under virtual microscope

3.2 Apparatus and Performance Analysis

For developing and experimental analysis, Unity 3D 2019.2.2f1 and Blender were used for the proposed method. In terms of the virtual reality headset, Google Cardboard was employed because it is inexpensive and widely available in Bangladesh. The method was extensively tested during the development process. The testing process was largely focused on how well the VR Laboratory operates in terms of a seamless user experience with reduced latency before testing the method among participants. To assure the method's performance, the VR Laboratory was constantly

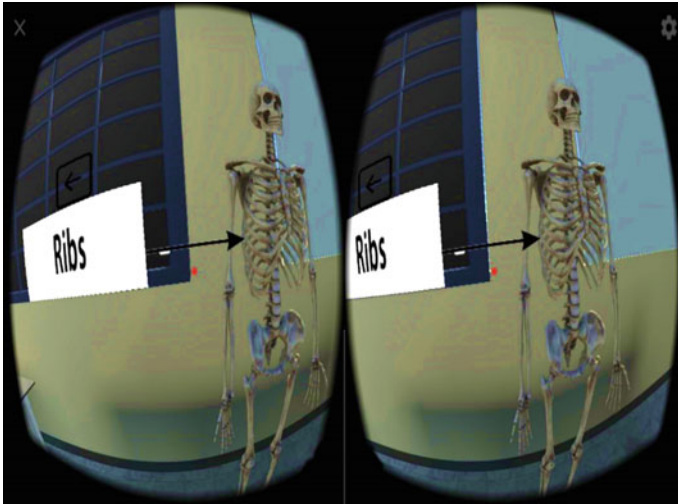


Fig. 10 Human skeleton with visual information

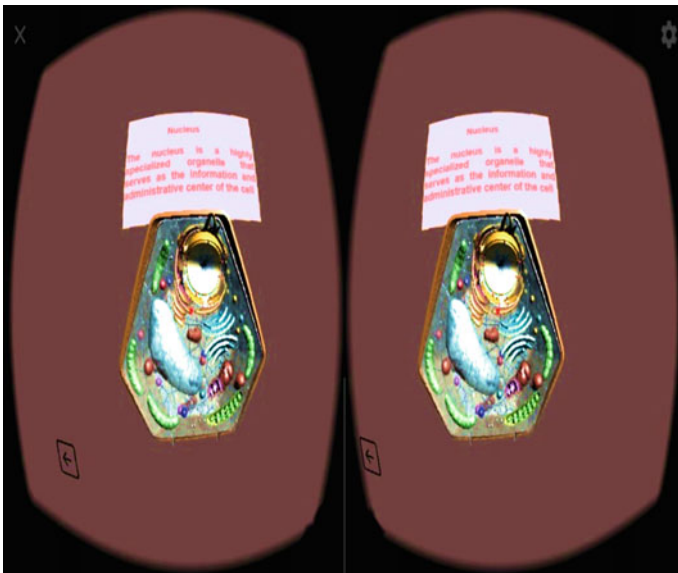


Fig. 11 Interactable 3D plant cell

watched, with an emphasis on improving frame rate. Any software should operate at a frame rate of 30fps which is considered as standard. A csharp script called "Fpscounter" was used to measure the frame rate. The proposed method's frame rate

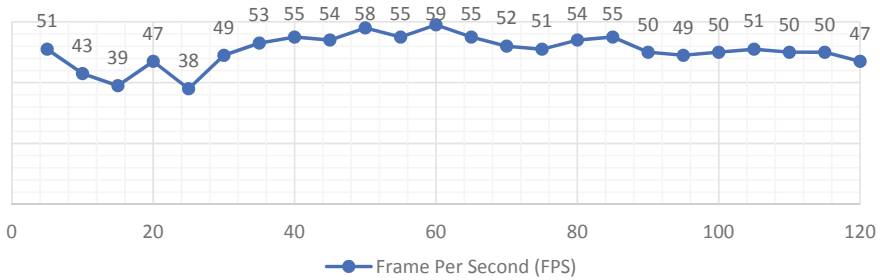


Fig. 12 FPS observation with time intervals

was monitored for 2 h since the total run time duration for the VR laboratory session is 2 h. The frame rate was measured at 5 min intervals.

FPS observation with time intervals is shown in Fig. 12.

3.3 Experiment and Result Evaluation

Research hypothesis was described using three components of HCI:

1. Participants (human)
2. Apparatus (computer)
3. Procedure (interaction)

To evaluate the proposed method's results, a mixed-method design (pre-assessment survey, post-assessment survey) was conducted. The survey questions were developed according to the current aspects of the ongoing situation. Pre-assessment survey was conducted to determine considering whether the students have previous experience with VR, access to their existing biology laboratory, various specimen related to syllabus, and impact on their learning efficacy during pandemic. During the pre-assessment survey, altogether 105 students from various high schools were recruited to test the proposed method. Out of 105 students, 71(67.6%) of them were male and 34(32.4%) of them were female students. When asked if they had any prior experience with VR, almost 75% of them responded negatively. Almost all of the recruits marked yes to if they had a biology laboratory in their institution. While 87% of the students pointed out they have microscopes in their laboratory, majority of the students said they do not have specimens like amoeba, animal cell, plant cell, fungal cell, bacteria cell, etc. According to the responses, 97.1% of the students did not have access to their biology laboratory during the lockdown period of the ongoing pandemic, which affected their efficacy of biology laboratory works. The feedback clearly portrays during such uncertain times, there's a need for an alternative learning solution in their laboratory (Fig. 14). The post assessment survey questionnaires were designed to take feedback from the students who participated

in the VR biology laboratory and consisted of five types of feedback for each question, viz. SA (strongly agree), A (agree), N (neutral), D (disagree), and SD (strongly disagree). Participants spent 45–80 min (mean 62.5) on virtual laboratory. According to the survey, the proposed model can hold the key to revolutionizing educational laboratory work and practical learning where expensive equipment is out of reach especially in third world countries. The main purpose of the proposed model was to replicate the existing biology laboratories the students have and also make sure they get to use the simulated VR experience in a very user-friendly manner. 64.8% of the students have vetted the proposed model, whereas about 10% of the population couldn't figure out where their opinion rests, so they remained neutral about the prospects and the remaining disagreed with the question. Similarly, 64 student have said that the proposed model makes learning better, whereas 15 of the students went against the statement. More than 65% of the students have “Strongly Agreed” and “Agreed” that the 3D human skeleton is well structured and the cell division simulation brings appeal in the learning process as well as brings forth newer depth in the learning process. According to the feedback provided, the VR laboratory experience is much more enjoyable than the orthodox laboratory work in the actual school laboratory. The feedback insists on creating similar laboratory simulation in other fields like chemistry and biology too in the future. A comparison study of the feedbacks reveals how the proposed method had a favorable influence on students in terms of its learning efficacy (Fig. 15). During the pre- and post-assessment survey, participants were particularly asked to rate their learning efficacy of biology practical laboratory work during ongoing pandemic lockdown through online video conferencing platforms vs the proposed VR method. The feedback consisted of five types of responses: very bad, bad, neutral, good, and very good. Their feedback illustrates they prefer our solution over other conventional learning methods. In the feedback, in response to their learning efficacy through online learning majority of the participants (62 out of 105, 59%) voted negatively while 14% stayed neutral and only 20% of the participants voted good. On the other hand, regarding the proposed VR learning method the participants (65 out of 105, 62%) took the opposite stand and voted positively, and only 20% of the students gave opposing feedback while only 4% remained neutral, which implies learning through learning efficacy before and after VR exposure is shown in Figs. 13 and 15.

VR has a significant effect on participants learning efficacy (Fig. 15).

Limitations:

This study also has few limitations, which may have an impact on the results and create a danger of bias in the findings. First, because of the lockdown, the sample size for this pre- and post-assessment research was modest, and a greater number of teachers will be assessed once the prototype is improved and ready for actual deployment. Next, in order to avoid overwhelming participants with evaluation instruments, the amount of engagement in VR was not examined, so we don't know whether there are any potential links between engagement and changes in acceptance of VR-based learning. However, assessing the pragmatic features of the VR experience is something we absolutely want to do. This entails doing a practical exam right away.

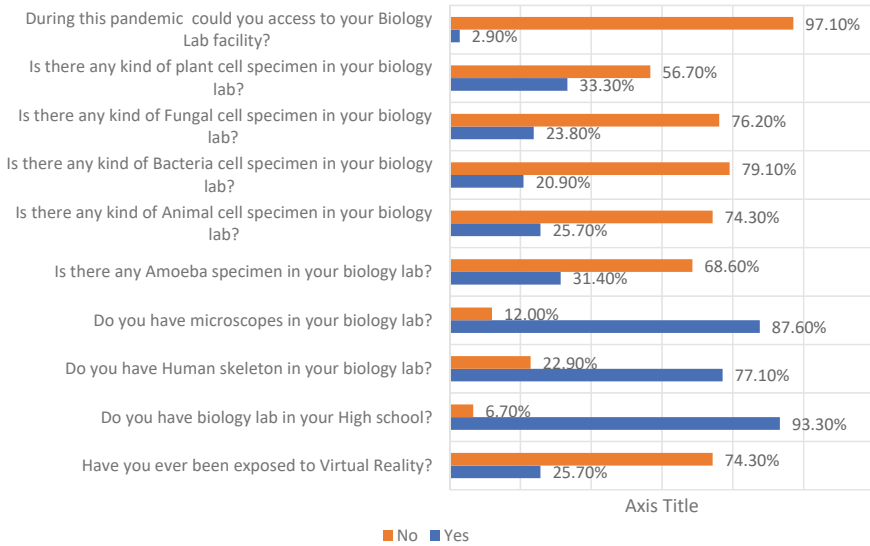


Fig. 13 Pre-assessment survey

Finally, there was no monitoring group and no follow-up investigation to evaluate if any changes in long-term learning efficacy is maintained. The next phase is to investigate these parameters to see if there are any changes in learning through VR.

4 Conclusion and Future Work

The immersive virtual reality simulated biology laboratory based on the real-life environment and how a traditional biology laboratory works to facilitate the day-to-day curriculum of biology practical can be identified as a success. The proposed model has been through the test phase, and the results obtained from the feedback points out the current version is actually making the students understand better their biology laboratory works and their experiments and interaction with the laboratory equipment brings forth a positive outcome in the learning phase. The main aim of this project was to provide an equal footing to the underprivileged students in comparison with their privileged peers who have access to state-of-the-art biology laboratories in their school. The next phase of the development for this immersive VR simulated biology laboratory would include addition of more experiments that are included in the biology syllabus to make it an all-inclusive solution. Also, a connected platform where the multiple users can interact within the virtual simulation can be created. This will allow for teachers to actually see how the students work in the biology laboratory and mentor and grade them. This will also open doors to peer learning and group project executions leading to innovation. In similar way, an immersive

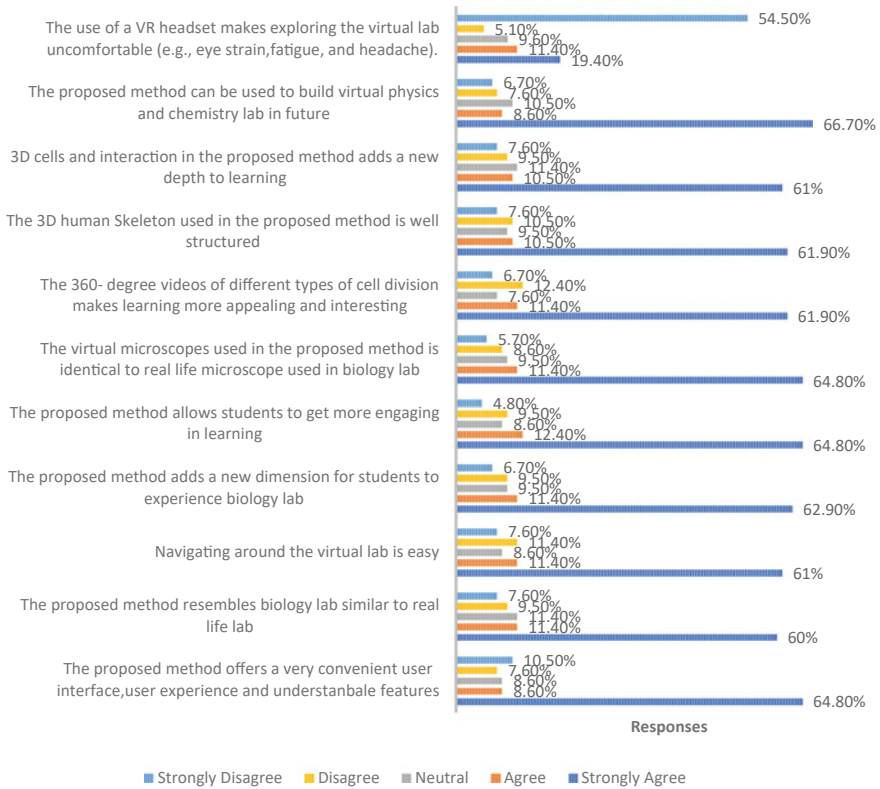


Fig. 14 Post-assessment survey

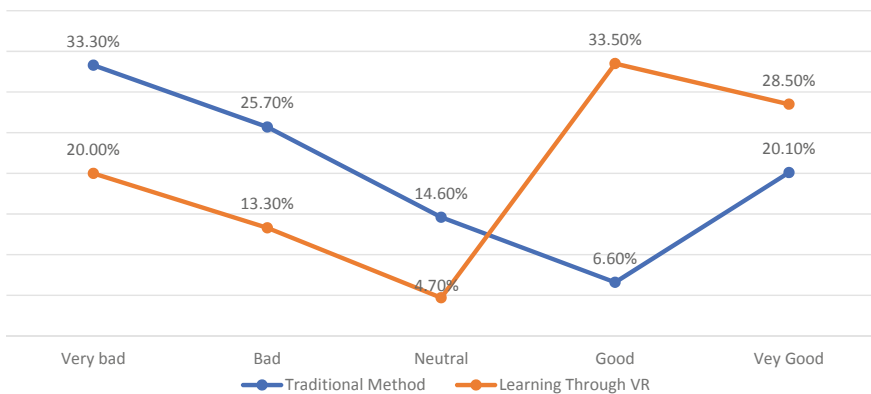


Fig. 15 Learning efficacy before and after VR exposure

virtual reality simulated physics and chemistry laboratory can be created apropos of the curriculum that is followed in Bangladesh.

Acknowledgements Declarations The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interest Statement On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Hamilton D, McKechnie J, Edgerton E, Wilson C (2021) Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *J Comput Educ* 8(1):1–32
2. Hodson D (1996) Laboratory work as scientific method: three decades of confusion and distortion. *J Curric Stud* 28(2):115–135
3. Kezar A, Eckel PD (2002) The effect of institutional culture on change strategies in higher education: universal principles or culturally responsive concepts? *J Higher Educ* 73(4):435–460
4. Widodo A, Maria RA, Fitriani A (2017) Constructivist learning environment during virtual and real laboratory activities. *Biosaintifika: J Biol Biol Educ* 9(1):11–8
5. Swan AE, O'Donnell AM (2009) The contribution of a virtual biology laboratory to college students' learning. *Innov Educ Teach Int* 46(4):405–419
6. Aripin I, Suryaningsih Y (2021) The development of prototype virtual laboratory through biology, technology, engineering, and mathematics (BTEM). *J Phys Conf Ser* 1806(1):012171 IOP Publishing.
7. Loftin RB, Engleberg M, Benedetti R (1993) Applying virtual reality in education: a prototypical virtual physics laboratory. In: Proceedings of 1993 IEEE research properties in virtual reality symposium. <https://doi.org/10.1109/vrais.1993.378261>
8. Dhaya R (2020) Improved image processing techniques for user immersion problem alleviation in virtual reality environments. *J Innov Image Process* 2(2):77–84
9. Kluge A (2014) Combining laboratory experiments with digital tools to do scientific inquiry. *Int J Sci Educ* 36(13):2157–2179
10. Špernjak A, Šorgo A (2018) Differences in acquired knowledge and attitudes achieved with traditional, computer-supported and virtual laboratory biology laboratory exercises. *J Biol Educ* 52(2):206–220
11. Cochrane T (2016) Mobile VR in education: From the fringe to the mainstream. *Int J Mob Blended Learn (IJMBL)* 8(4):44–60
12. Warburton S (2009) Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *Br J Edu Technol* 40(3):414–426
13. Setiawan A, Malik A, Suhandi A, Permanasari A (2018) Effect of higher order thinking laboratory on the improvement of critical and creative thinking skills. In: IOP conference series: materials science and engineering, vol 306, no 1, p 012008. IOP Publishing
14. Dziuban C, Graham C, Moskal P, Norberg A, Sicilia N (2018) Blended learning: the new normal and emerging technologies. *Int J Educ Technol Higher Educ* 15(1)
15. Moro C, McLean M (2017) Supporting students' transition to university and problem-based learning. *Medical Science Educator*. 27(2):353–361
16. Yang J, Chen C, Chang JM (2010) Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning. *Comput Educ* 55(3):1346–1356

17. Wu J, Guo R, Wang Z, Zeng R (2019) Integrating spherical video-based virtual reality into elementary school students' scientific inquiry instruction: effects on their problem-solving performance. *Interact Learn Environ* 29(3):496–509
18. Virvou M, Katsionis G (2008) On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Comput Educ* 50(1):154–178
19. Li J, Chen G, De Ridder H, Cesar P (2020) Designing a social VR clinic for medical consultations. In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems 2020*, pp 1–9
20. Basdogan C, Sedef M, Harders M, Wesarg S (2007) VR-based simulators for training in minimally invasive surgery. *IEEE Comput Graphics Appl* 27(2):54–66
21. Riener R, Harders M (2012) *Introduction to virtual reality in medicine*. Virtual reality in medicine. Springer, London, pp 1–12
22. Birt J, Stromberga Z, Cowling M, Moro C (2018) Mobile mixed reality for experiential learning and simulation in medical and health sciences education. *Information* 9(2):31
23. Sheikh A, Barry D, Gutierrez H, Cryan J, O'Keeffe G (2015) Cadaveric anatomy in the future of medical education: What is the surgeons view? *Anat Sci Educ* 9(2):203–208
24. Moro C, Štromberga Z, Raikos A, Stirling A (2017) The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat Sci Educ* 10(6):549–559
25. Ying L, Jiong Z, Wei S, Jingchun W, Xiaopeng G (2017) VREX: virtual reality education expansion could help to improve the class experience (VREX platform and community for VR based education). In: *2017 IEEE Frontiers Educ Conf (FIE)*
26. Seymour N (2007) VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World J Surg* 32(2):182–188
27. Maresky H, Oikonomou A, Ali I, Ditzkofsky N, Pakkal M, Ballyk B (2018) Virtual reality and cardiac anatomy: exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clin Anat* 32(2):238–243
28. Smys S, Raj JS, Krishna Raj N (2019) Virtual reality simulation as therapy for posttraumatic stress disorder (PTSD). *J Electron Inform* 01(01):24–34

Multi-label Classification of Cervix Types with Image Size Optimization for Cervical Cancer Prescreening by Deep Learning



Maryna Tomko, Mykhaylo Pavliuchenko, Ivan Pavliuchenko, Yuri Gordienko , and Sergii Stirenko 

Abstract Cervical cancer is one of the most pressing and embarrassing problems in the medical, psychological, and social aspects of women’s lives. The multi-class classification problem and multi-label classification problem with image size optimization (on the basis of the updated 6-class labeling) are considered on the basis of the dataset published during Intel and MobileODT Cervical Cancer Screening competition on Kaggle platform. In the multi-class classification problem, the better metrics were obtained after training the standard DNNs with the initial weights obtained after pretraining on ImageNet dataset with the conclusion as to possible useful usage of the smaller versions of DNNs (like MobileNetV2, NASNetMobile). In the multi-label classification problem with image size optimization, EfficientNetB0 model was used as an example for investigation of the effective method of metric improvement by image size optimization. It allowed us to improve the mean AUC values by 2.7–2.8% (in comparison with the standard image sizes of 224 * 224 pixels used for pretraining the standard models on ImageNet dataset) in the limits of the standard deviation 0.3–1.8% by the proper selection of the optimal input image size. In general, this approach based on averaging values of metrics obtained by various regimes of DNN training (with and without DA) and verification (validation and testing) for the pairs of smallest and biggest available image sizes with further extrapolation of metric tendencies could be the useful strategy for image size optimization in other use cases.

Keywords Multi-label classification · Neural networks · Deep learning · Data augmentation · Cervix uteri · Cervical cancer · Optimization

M. Tomko · I. Pavliuchenko · Y. Gordienko (✉) · S. Stirenko
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”,
Kyiv, Ukraine
e-mail: yuri.gordienko@gmail.com

M. Pavliuchenko
Zaporizhzhia State Medical University, Zaporizhzhia, Ukraine

1 Introduction

Cervical cancer is one of the most pressing and exciting problems in the medical, psychological, and social aspects of women's lives [23]. However, cervical cancer is very easy to prevent if detected in the precancerous stage. However, partly due to the lack of experience in this field, one of the biggest problems of cervical cancer detection and treatment programs is determining the appropriate treatment, which may vary depending on the physiological differences of patients. To choose the right treatment, you need to know the type of cervix. Different types of cervix have different transformation zones and are not always visible, so some patients need further testing and others do not.

Among all problems mentioned below, in this paper we are interested in determining the type of the cervix (its type by the transformation zone [5]) using deep learning approaches to help to identify possible negative variants of the disease and choose effective treatment for further complete overcoming of cancer in patients. Section 2 contains description of the state of the art, Sect. 3 describes dataset, models, experiments, and the whole workflow, Sect. 4 gives the results obtained during the experiments, Sect. 5 contains discussions of results and resumes them.

2 Background and Related Works

Cervical cancer is one of the most pressing and embarrassing problems in the medical, psychological, and social aspects of women's lives. According to World Health Organization (WHO) more than 600 thousand new cases (3.1% of all cancer cases) and more than 340 thousands deaths (3.4% of all cancer deaths) were recorded worldwide in 2020 [17]. This problem is especially relevant in Ukraine, where the cervical cancer mortality rate in 1997 was the highest one in Europe and was equal to 5.22 (per 100,000 women) [22]. Moreover, the stable, though not very large, increase in the number of cases was observed since 1997. Most cervical cancers begin in cells in the area of transformation. Basically, the assessment of the transformation zone and classification into appropriate transformation zones of types 1, 2, or 3 is usually performed after the application of 5% acetic acid. Determining the location of the junction between the multilayered squamous and columnar epithelium is a prerequisite for classifying the transformation zone of types 1, 2, or 3.

The boundary between the stratified squamous and columnar epithelium is transformation zone—an area of metaplastic immature epithelium, which is vulnerable to human papillomavirus. Pathological processes on the cervix usually begin in this area, and in this case the doctor observes the line of contact of the abnormal epithelium with the columnar. The joint can be “fully visible,” “partially visible,” or even “invisible.” The junction of multilayered squamous and columnar epithelium is considered fully visible when it can be traced 360° and partially visible when a sector or part of the junction is located in the cervical canal or the lesion overlaps the junction of

multilayered squamous epithelium with a columnar inner edge in the cervical canal. The junction of the stratified squamous epithelium is considered invisible when it is located in the cervical canal and cannot be traced [4, 13].

If a transformation zone of types 1 or 2 is identified, the boundary between the stratified squamous and columnar epithelium will be fully visualized. The transformation zone is classified as type 1 when it is completely located on the ectocervix (without the endocervical part). The transformation zone of types 2 and 3 always has an endocervical part, in the boundary between the squamous and columnar epithelium extends to the cervical canal. If at the endocervical location the border is completely visualized, it is called the zone of transformation of type 2. The “inner” boundary of the transformation zone of types 1 and 2 is “completely visible,” but the difference between types 1 and 2 of the transformation zone is important, especially when planning treatment. The inner boundary of the transformation zone is the newly formed junction of the multilayered squamous and columnar epithelium, and the outer boundary is the last open gland. The entire area of the transformation zone is particularly vulnerable to exogenous factors. As a rule, precancerous processes occur in the transformation zone, so its visualization is important in the study [4, 13].

Deep learning (DL) methods based on deep neural networks (DNNs) recently were applied to medical imaging as a whole [2, 6, 15, 20] and for cervical cancer prescreening in part [23]. Several recent reviews contain the thorough studies of such attempts for cervical cancer prescreening [1, 21, 23].

The following promising approaches should be emphasized. For example, the fully automated pipeline for cervix detection and cervical cancer classification was proposed recently [3]. It consists of two pretrained DL models for the automatic cervix detection and cervical tumor classification. The first DL model detects the cervix region achieving an accuracy of 0.68 in terms of intersection of union (IoU) measure. Self-extracted features are used by the second model to classify the cervix tumors. These features are learned using two lightweight models based on convolutional neural networks (CNN). The proposed DL classifier is characterized by an area under the curve (AUC) score of 0.82. It is very important that the pipeline accuracy, speed, and lightweight architecture make it very appropriate for mobile phone deployment for application under field conditions which is the very hot area of investigations now [18].

In another approach, the novel methodology was proposed where image enhancement on cervigrams was followed by segmenting regions of interest (RoIs) and then classifying the RoI to determine the appropriate treatment [9]. For the classification problem, the hierarchical convolutional mixture of experts (HCME) algorithm was proposed which is capable to avoid overfitting, taking into account that small datasets are typical and create problems in the field of biomedical computer vision. The proposed methodology has outperformed several existing methodologies on publicly available Intel and MobileODT Kaggle dataset with accuracy of 96.77% [12].

3 Methodology

3.1 Dataset

The main source of inspiration for this research was Intel and MobileODT Cervical Cancer Screening competition on Kaggle platform, from which the main set of data for further work was taken [12]. The aim of the competition, according to the description on the resource, was to develop algorithms for the correct classification of cervical types based on images of the cervix. All of these types of cervix are considered normal (not cancerous) in the dataset, but because transformation zones are not always visible, some patients require further testing and others do not. This decision is very important for the healthcare professional and critical for the patient. Identifying transformation zones is a challenge for healthcare professionals, so the algorithm solution will significantly improve the quality and effectiveness of cervical cancer screening for these patients. As it is already mentioned, in this work was used dataset provided by Intel and MobileODT Cervical Cancer Screening competition [12].

Multi-class Classification. The dataset consists of 8727 pictures of not cancerous cervix, which is splitted into 3 types (Fig. 1). As far as in this labeling scheme each image contained only single label then this case was considered as multi-class classification problem (see below).

Multi-label Classification. After a detailed review of data by medical experts, the dataset was labeled by additional and more detailed six classes (Normal, Minor, Major, Suspected, NotSpecific, and Misc, see some examples in Fig. 2) that are used in the current practice. As far as in this labeling scheme some labels can be prescribed to the same image then this case was considered as multi-label classification problem (see below).

Content and Structure of the Dataset. The detailed exploratory data analysis (EDA) was performed to understand the data representation by classes, weight, height, and

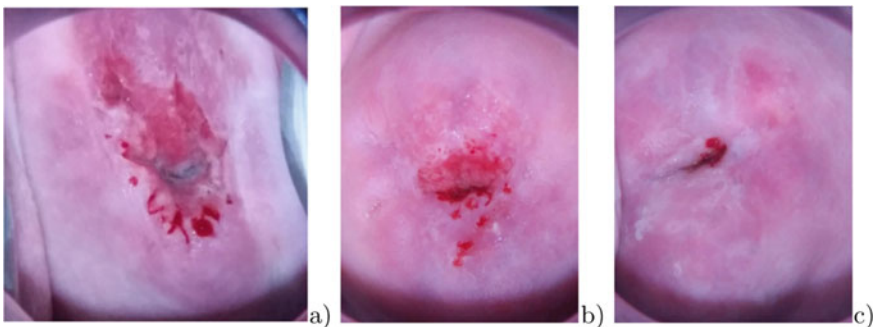


Fig. 1 Examples of objects for the multi-class (3 classes) classification problem from Intel and Mobile ODT cervical cancer screening competition [12]: **a** Type 1, **b** Type 2, **c** Type 3

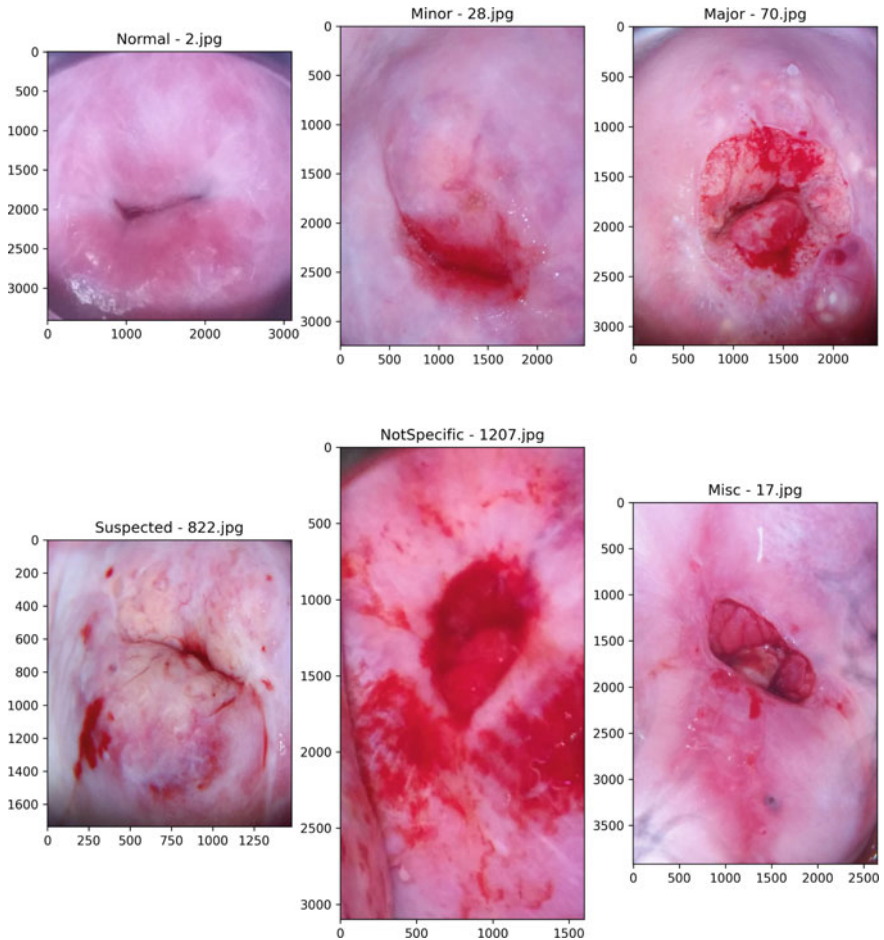


Fig. 2 Examples of objects for the multi-label (6 classes) classification problem with labels and filenames (above each image) and image sizes in pixels (shown on the axes)

area. Below distribution of images by classes in counts (Fig. 3a) and proportions (Fig. 3b) is shown, where the pronounced bias is evident. During training to avoid such unbalanced configuration, the weights were applied that are proportional to the actual representation of images by the classes. It should be noted that the same distribution of images by classes was preserved for train, validation, and test subsets of the dataset (Fig. 3b).

Additional very important aspect is related with the desire of practitioners to work with the selected RoIs of raw images, especially taking into account the high efficiency of some ROI segmenting approaches and following classification [9]. That is why some manual cropping was performed by medical experts during manual labeling by 6 classes according to their medical expertise. The final distribution of

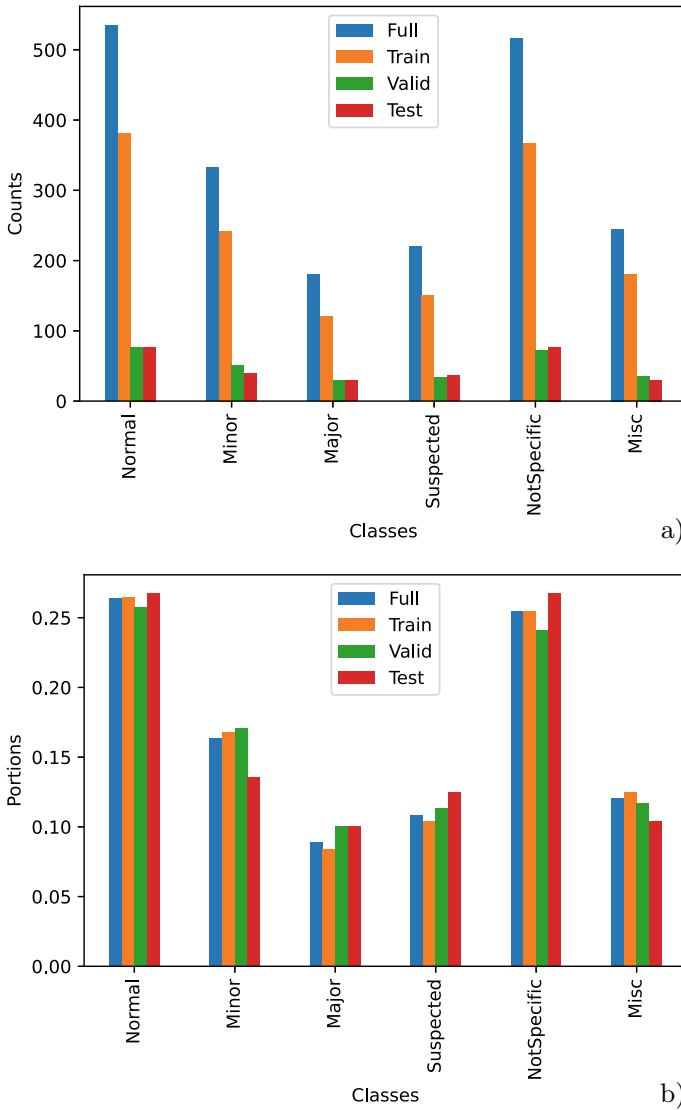


Fig. 3 Distribution of images by classes in: **a** counts, **b** proportions

the cropped images by sizes is shown in Fig. 4a by height and width, and Fig. 4b by area. Despite the current attempts to use the standard DNNs pretrained on ImageNet dataset for image with sizes of 224 * 224 (or 300 * 300, 331 * 331, etc.) pixels, the current dataset demonstrates the quite wider range of sizes from ~200 * 200 up to ~3000 * 4000 pixels with nearly uniform distribution (except for some narrow peaks at tails). That is why the necessity to investigate dependence of DNN performance for various input image size is of great interest.

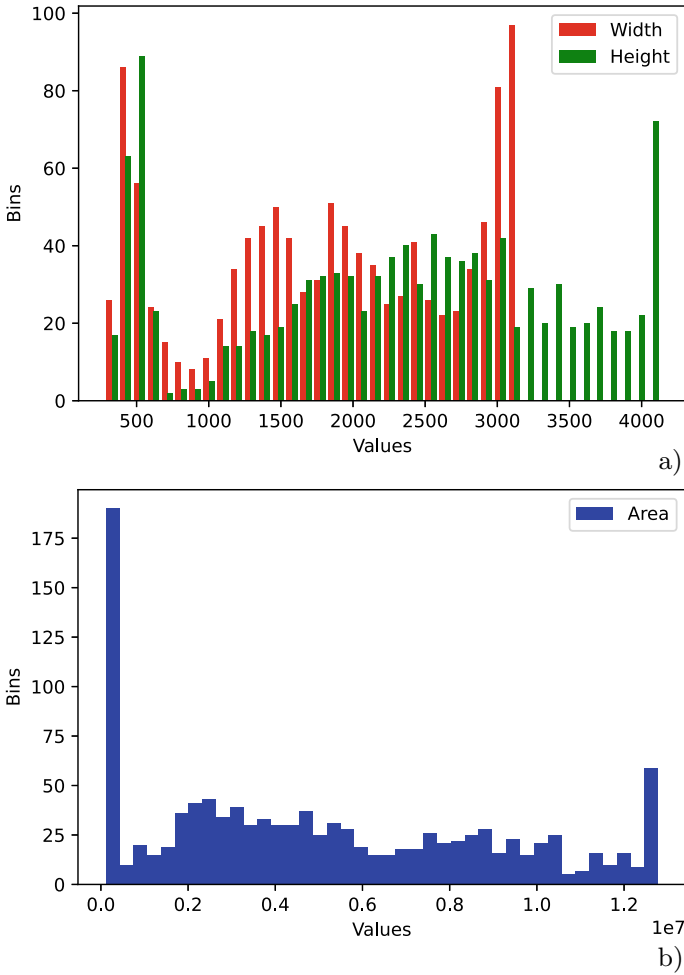


Fig. 4 Distribution of images by: **a** width and height, **b** area

Preprocessing and Data Augmentation. For multi-class classification (where each image was assigned to one and only one label), the images were reduced to 224 * 224 size and normalized to 0.0–1.0 range by dividing the value of each pixel by 255. Then several DNNs were trained with data augmentation (DA) using a data generator with the following parameters: resizing photos, rotating them, stretching them horizontally and vertically, enlarging and reducing the image size, and mirroring them horizontally.

For multi-label classification (where each image was assigned to a set of labels), the image sizes were changed to various different sizes (see “Image Size” column in Tables 2 and 3) to investigate dependence of DNN performance versus various input image sizes. To visualize similarity and dissimilarity of cervix images for

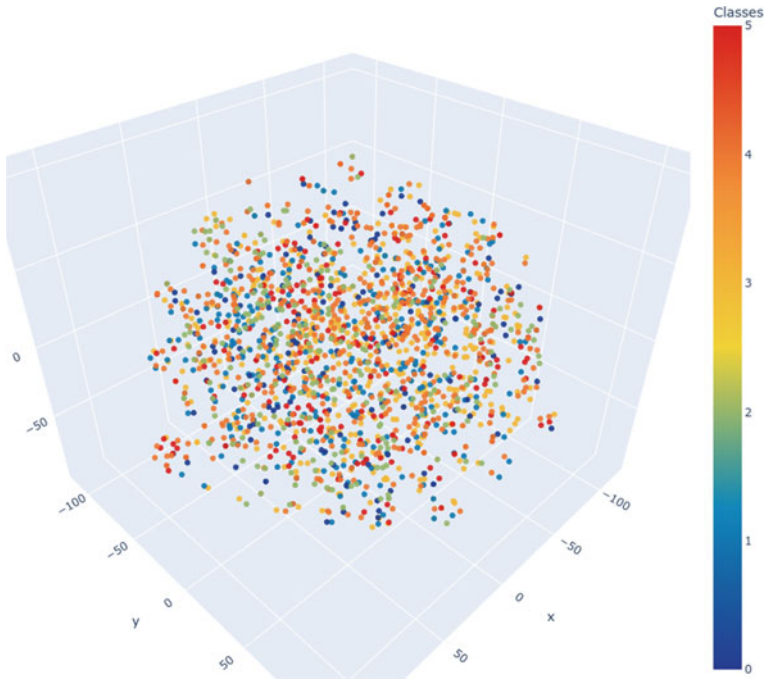


Fig. 5 Visualization of t-SNE model for a cervix image by a three-dimensional point, where similar cervix images are modeled by nearby points and dissimilar cervix images are mapped to distant points. The colors correspond to 6 classes of cervix types (see description in the text)

different classes, the well-known t-distributed stochastic neighbor embedding (t-SNE) technique was applied [16, 19]. It allowed us to embed high-dimensional cervix image data into a 3D space, which can then be visualized in a scatter plot (Fig. 5). It should be emphasized that actually, this t-SNE map of cervix images in 3D space allows us visually understand the very close similarity of objects from different classes that hardly can be distinguished by simple approaches.

3.2 Models and Workflow

Multi-class Classification Workflow. To solve the multi-class classification problem, the following popular standard DNN models were used in the work: InceptionV3, ResNet50, MobileNetV2, and NASNetMobile. They were selected with regard to their high performance on ImageNet dataset (InceptionV3 and ResNet50) and relatively high performance taking into account the lower size that can be important for Edge Computing devices (MobileNetV2 and NASNetMobile). They were used in two training modes: (i) initial random weights, (ii) initial weights pretrained

on ImageNet dataset. A pretrained network is usually trained on a huge set of data for large-scale image classification task, namely ImageNet here. Thus, the spatial hierarchy of features learned by DNNs can be very effective as common DNN models for various problems, even if they applied to completely different classes than those in the original problem for which the DNN model was trained. Because the selected pretrained DNN models by default accept $224 * 224$ input image size, so the input image size for all models was set to this size. Initially, all DNNs were studied in 20 epochs, because as it is turned out, that was enough for saturation of validation metric growth. The batch size for the initial runs was selected as 32. Accuracy, precision, AUC, and loss were chosen as the main metrics. The standard architecture of DNNs used was partially changed to accommodate the number of classes used; namely, the last classification layer was taken away, the dense layer with 512 nodes was added, the dropout layer was added dropout ratio as 0.3, the classification dense layer with 3 nodes was added, which determines the final image category. To adopt multi-label dataset (with 6 classes) to the original multi-class configuration (with 3 classes), cervix images were sorted into 3 appropriate folders (according to the availability of the corresponding label) with the following numbers of images: for the first type—1191 images, for the second—3567 images, for the third—1976 images. In total, 6734 training images in all these folders. Along with 512 images for testing (sorted by these types also), the total number of images was equal to 7246.

Multi-label Classification Workflow. To solve the multi-label classification problem and investigate dependence of DNN performance versus various input image sizes EfficientNetB0 DNN architecture was selected because it was one the smallest and powerful enough architecture that can be ported to devices with the limited computational resources [10, 11]. It was used in a single training mode with initial weights pretrained on ImageNet dataset (as a result of the previous multi-class classification problem, see details below), but in two regimes of DA: (i) no DA, (ii) DA with lossy transformations like vertical and horizontal flips, random crops (up to 10%), small Gaussian blur (with random sigma between 0 and 0.5), scaling down (80%) and up (120%), translations (in and out by 20%), rotations (between -25 and 25°), and shears (between -8 and 8°). DA procedures were necessary to increase the amount of relevant data in the dataset and the level of possible variability of the data, because the number of images in the dataset was not very high by the current standards. The initial learning rate (LR) was equal to 0.001 with the patience parameter used for reducing LR on plateau by factor of 10 every epoch to the minimum LR equal to 10^{-8} . The batch size was equal to 32 for the input square image sizes 256, 384, 448, 512, 8—for 528, 600, 640, 768, and 2—for 1024, 1536, 2048. The models were implemented by Tensorflow 2.0 Keras application programming interface and run on NVIDIA GPU accelerators A100.

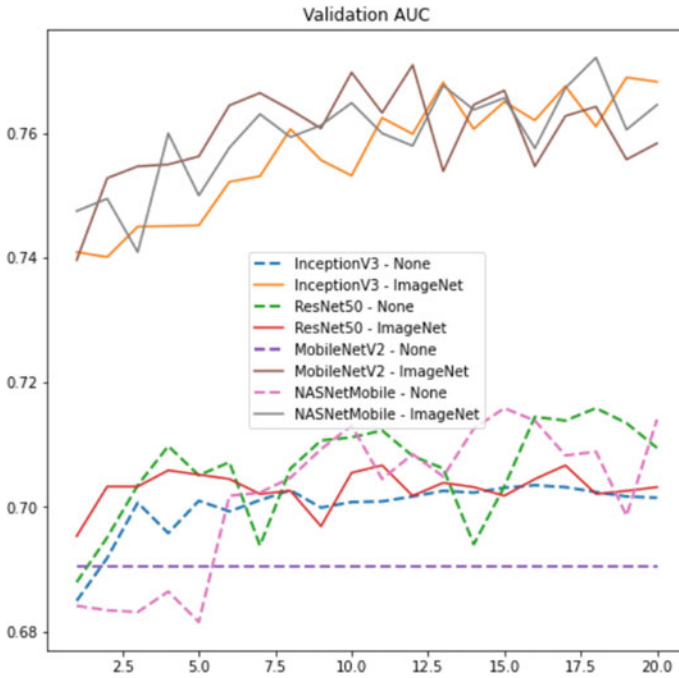


Fig. 6 Comparison of validation AUC evolution during training for various DNNs

4 Results

4.1 Multi-class Classification

The evident observation from the training results for various DNNs (Fig. 6 and Table 1), all DNNs work best after pretraining on ImageNet dataset, despite some recently reported results that the best DNNs pretrained on the general-purpose datasets (like ImageNet) not always perform well on the medical datasets. At this stage of the experiments, these conclusions were based on AUC values, because this metric takes into account several thresholds, and the other metrics have a fixed threshold 0.5. Therefore, all DNNs considered are better for the versions pretrained on ImageNet dataset (except for ResNet50, which gives the same low result for untrained and pretrained versions of DNNs and this will be investigated in the future research).

4.2 Multi-label Classification

As it was mentioned above, for multi-label classification the image sizes were changed to various different values (see “Image Size, L ” column in Tables 2 and

Table 1 Comparison of the best metric values for different models

Model	Weights	Accuracy	AUC	Precision	Loss
InceptionV3	None	0.548	0.702	0.800	0.999
InceptionV3	Imagenet	0.578	0.779	0.612	0.870
ResNet50	None	0.548	0.706	0.591	0.984
ResNet50	Imagenet	0.549	0.703	0.556	0.992
MobileNetV2	None	0.548	0.691	0.548	0.986
MobileNetV2	Imagenet	0.570	0.758	0.586	1.127
NASNetMobile	None	0.548	0.714	0.564	0.983
NASNetMobile	Imagenet	0.573	0.765	0.610	0.910

Table 2 AUC values for different image sizes

Image size, L	AUC_V^{DA}	AUC_t^{DA}	AUC_V^{noDA}	AUC_t^{noDA}	Mean, AUC	Stdev
128	0.747	0.700	0.747	0.752	0.736	0.025
224	0.768	0.750	0.748	0.769	0.759	0.011
256	0.772	0.762	0.781	0.781	0.774	0.009
384	0.785	0.787	0.791	0.785	0.787	0.003
448	0.794	0.756	0.794	0.781	0.781	0.018
512	0.796	0.781	0.803	0.763	0.786	0.018
600	0.788	0.719	0.808	0.763	0.769	0.038
640	0.754	0.753	0.815	0.787	0.777	0.030
768	0.779	0.755	0.790	0.782	0.776	0.015
1024	0.694	0.703	0.781	0.774	0.738	0.046
1536	0.711	0.699	0.693	0.691	0.698	0.009
2048	0.666	0.627	0.713	0.671	0.669	0.036

Table 3 OLS regression results

Data	Image size, L	A_s	A_t	A_{stdev}	B_s	B_t	B_{stdev}	R^2
Raw	$L < \overline{L(AUC_{max})}$	$-3.9e-03$		0.006	$-2.0e-04$		$3.9e-05$	0.929
Raw	$L \geq \overline{L(AUC_{max})}$		$-4.2e-03$	0.003		$7.5e-05$	$4.7e-06$	0.973
Smoothed	$L < \overline{L(AUC_{max})}$	$-3.8e-03$		<0.001	$-5.1e-05$		$1.3e-06$	0.997
Smoothed	$L \geq \overline{L(AUC_{max})}$		$9.9e-04$	0.001		$7.7e-05$	$8.9e-07$	0.999

3) to investigate dependence of DNN performance versus various input image sizes. The values of AUC metrics for various training, validating, testing regimes, and for different image sizes are given on Fig. 7, where AUC_v^{DA} —the best AUC validation value (subscript v) for the DNN trained with DA (superscript DA), AUC_t^{DA} —the best AUC testing value (subscript t) for the DNN trained with DA (superscript DA) that was the best after validation, AUC_v^{noDA} —the best AUC validation value (subscript v) for the DNN trained without DA (superscript noDA), AUC_t^{noDA} —the best AUC testing value (subscript t) for the DNN trained without DA (superscript noDA) that was the best after validation, \overline{AUC} —the mean AUC value averaged over the 4 previous AUC values, and “Stdev”—its standard deviation. In Fig. 7, the values of these AUC metrics are denoted by symbols with their maximal numerical values observed for some image size L , the mean AUC value \overline{AUC} is given (in the top legend) with the correspondent standard deviation observed for some image size L also. The red filling color in Fig. 7 means the range of standard deviations for each image size. It should be noted that the mean AUC value ($\overline{AUC} = 0.759 \pm 0.011$) for the standard input image size $224 * 224$ used in many works (and for multi-class classification here) can be improved by the proper selection of the other image size, for example, up to ($\overline{AUC} = 0.787 \pm 0.003$) for $384 * 384$ and ($\overline{AUC} = 0.786 \pm 0.018$) for $512 * 512$. Unfortunately, the bigger increase of the input image size due to the better camera resolution does not allow to improve these results and discussion of the possible reasons which are given below.

These values were also smoothed by the locally weighted polynomial regression method [7] and denoted by the lines in Fig. 7 for the better representation of the validation and testing results. Figure 7 contains additional notations for the following values: $AUC_{v,sm}^{DA}$, $AUC_{t,sm}^{DA}$, $AUC_{v,sm}^{noDA}$, $AUC_{t,sm}^{noDA}$ —are AUC values that are similar to the aforementioned ones, but these were obtained after smoothing (subscript sm).

For all regimes of DNN training (with and without DA) and verification (validation and testing), the following two tendencies can be observed: (i) the dependencies of AUC versus image size are not monotonous and have maximum values, (ii) the maximal AUC values are observed for nearly the same image size (in the limits of some error).

5 Discussion and Conclusions

5.1 Multi-class Classification

As a result of work on the multi-class classification problem, it was found that the best metric values were demonstrated by DNNs like InceptionV3, MobileNetV2, NASNetMobile that were trained on ImageNet dataset. MobileNetV2 and NASNet-Mobile are of great interest because they are small size, they are fast to train and test and provide the high values of AUC and accuracy metrics. This means that these net-

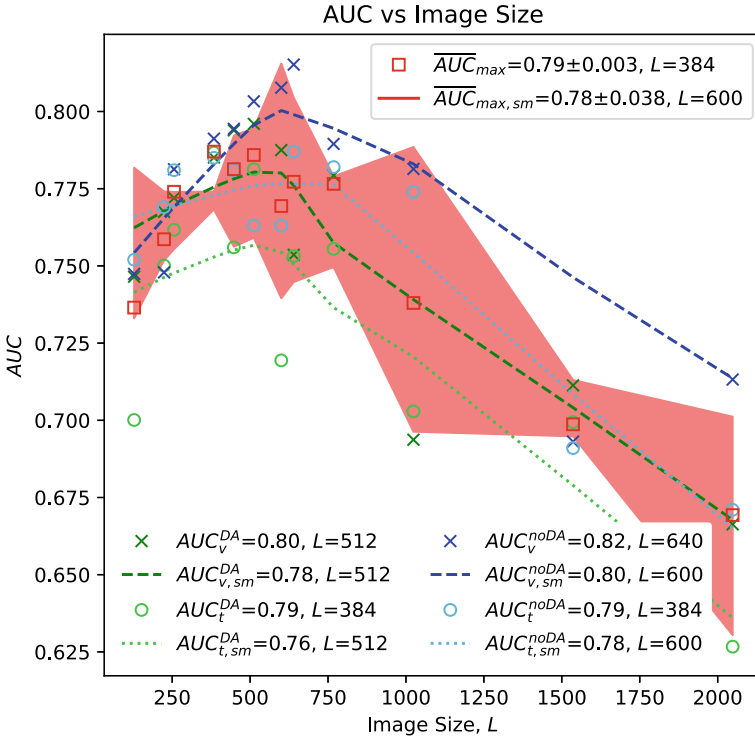


Fig. 7 Dependence of AUC versus image size (see the explanations of these notations in the text)

works are perfect for field conditions, including use cases with limited computational and memory resources.

5.2 Multi-label Classification

As to the multi-label classification problem, our research work was targeted on usage of one of the recent DNN like EfficientNetB0 with the small and powerful enough architecture that can be ported to devices with the limited computational resources. The additional motivation was to analyze the observed the non-monotonous dependence of AUC versus image size (Fig. 7).

Scaled Image Size Dependence. For this purpose, all of the dependences of AUC versus image size (Fig. 7) were rescaled with regard to maximum mean values of raw AUC (\overline{AUC}_{max}) and smoothed AUC ($\overline{AUC}_{max,sm}$). Finally, the scaled AUC (\overline{AUC}_{sc}) was a function of the scaled image sizes (L_{sc}):

$$\overline{AUC}_{sc} = f(L_{sc}), \tag{1}$$

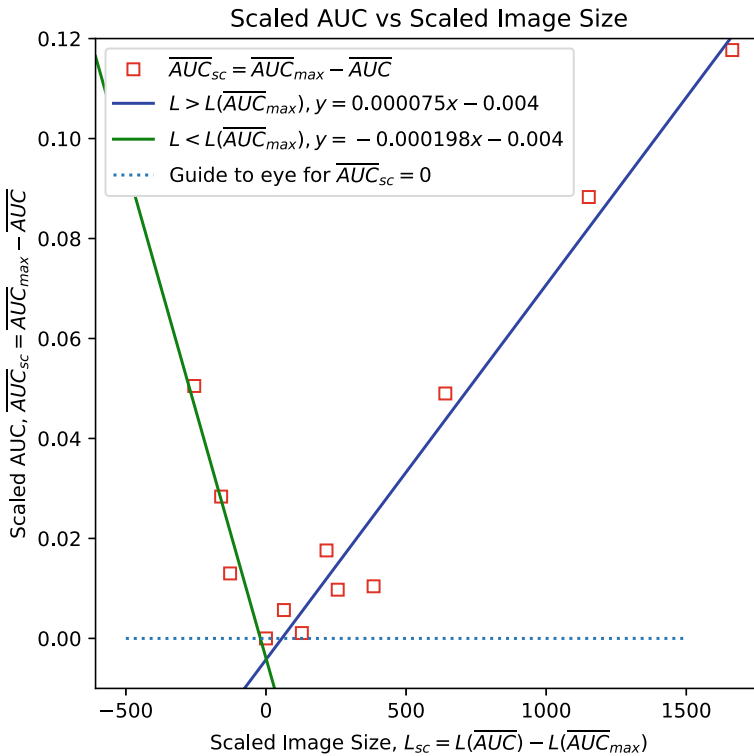


Fig. 8 Dependence of the scaled mean AUC values (\overline{AUC}_{sc}) versus the scaled image size values (L_{sc}) for raw original data. Equations in the legend correspond to Eq. 2

where the scaled mean AUC is $\overline{AUC}_{sc} = \overline{AUC}_{max} - \overline{AUC}$, and scaled image size $L_{sc} = L(\overline{AUC}) - L(\overline{AUC}_{max})$. The resulting dependencies for Eq. 1 are shown in Fig. 8 for raw original data and in Fig. 9 for smoothed data.

The resulting plots (for raw data in Fig. 8 and, especially, for smoothed data in Fig. 9) contain symbols that demonstrate strong visual linear dependence with decrease for negative values of L_{sc} and increase for positive values of L_{sc} . To check the linear dependence, the linear regression model was applied (see below).

Linear Regression for Rescaled Image Size Dependence. Table 3 contains results of ordinary least squares (OLS) method applied for estimating the unknown parameters (line intercept and slope) in a linear regression model, where A —is the fitting line intercept, A_{stddev} —is intercept standard deviation, B —is the fitting line slope, B_{stddev} —is the slope standard deviation, and R^2 — R -Squared or the coefficient of determination.

The data in Table 3 allow us to rewrite Eq. 1 in the following shape:

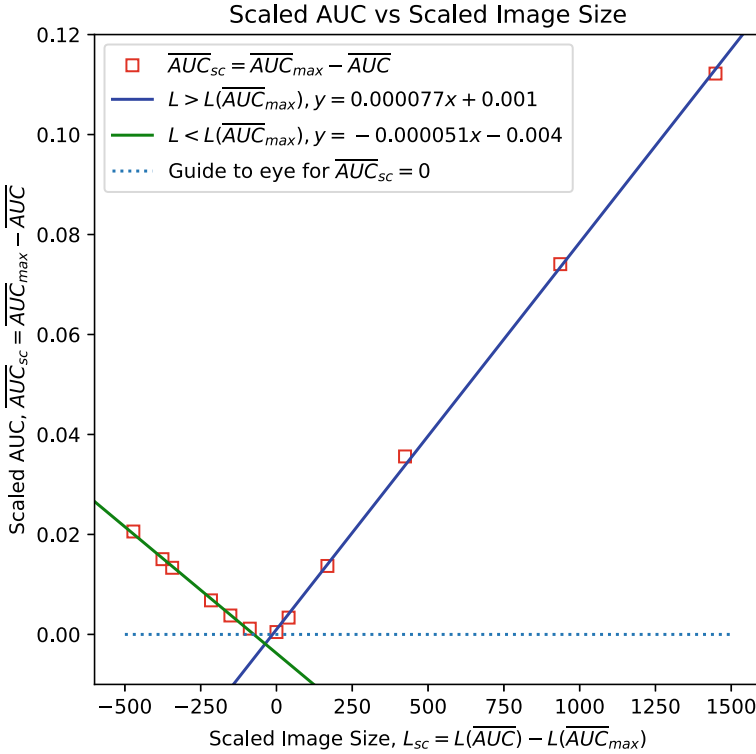


Fig. 9 Dependence of the scaled mean AUC values (\overline{AUC}_{sc}) versus the scaled image size values (L_{sc}) for smoothed data. Equations in the legend correspond to Eq. 2

$$\overline{AUC}_{sc} = \begin{cases} A_s + B_s * L_{sc} & \text{if } L < L(\overline{AUC}_{max}) \\ A_1 + B_1 * L_{sc} & \text{if } L \geq L(\overline{AUC}_{max}) \end{cases} \quad (2)$$

where $L_{sc} = L(\overline{AUC}) - L(\overline{AUC}_{max})$ and $\overline{AUC}_{sc} = \overline{AUC}_{max} - \overline{AUC}$.

It should be noted that both dependencies of the scaled mean AUC values (\overline{AUC}_{sc}) versus the scaled image size values (L_{sc}) for raw and smoothed data demonstrated the strong linear dependence for decrease and increase of the scaled AUC with the scaled image size. The scaled AUC is actually difference of AUC value from its maximal value, and the scaled image size is actually difference of L value from L value at the maximal AUC value. In the view of the much higher values of A_{stddev} in comparison with A_s and A_1 , they cannot be distinguished from 0 in the limits of standard deviations. Then it allows us to rewrite Eq. 2 in the following shape:

$$\overline{AUC} = \begin{cases} \overline{AUC}_{max} + B_s * (L - L(\overline{AUC}_{max})) & \text{if } L < L(\overline{AUC}_{max}) \\ \overline{AUC}_{max} + B_1 * (L - L(\overline{AUC}_{max})) & \text{if } L \geq L(\overline{AUC}_{max}) \end{cases} \quad (3)$$

The decrease of AUC for $L > L(\overline{\text{AUC}}_{\max})$ (Fig. 7) can be explained by the influence of the finer details inside images with the higher resolution that does not have useful information, but add some noise with the decrease of a signal-to-noise ratio (SNR): the finer details the lower SNR due to increase of noise. In reverse the decrease of AUC for $L < L(\overline{\text{AUC}}_{\max})$ can be explained by the roughening image with exclusion of some pertinent details and the subsequent decrease of SNR: the coarser details the lower SNR due to decrease of signal (the characteristic features).

Image Size Optimization. As it was mentioned above the mean AUC value ($\overline{\text{AUC}} = 0.759 \pm 0.011$) (see Table 2) for the standard input image size $224 * 224$ used in many works (and for multi-class classification here) can be improved by the proper selection of the other image size. Usually, the “brute force” search of the optimal image size takes the significant investment of resources to iterate over all possible image sizes in grid-like manner. But fortunately, at least in this use case the more effective image size optimization method can be used. Taking into account the results from Table 3 and Eq. 3, one can perform training DNNs in the aforementioned regimes (with and without DA): (i) for the two largest image sizes (e.g., for $1536 * 1536$ and $2048 * 2048$), calculate two $\overline{\text{AUC}}$ for these two L values, and determine B_l value; (ii) for the two smallest image sizes (e.g., for $128 * 128$ and $224 * 224$), calculate two $\overline{\text{AUC}}$ for these two L values, and determine B_s value. Then using Eq. 3, one can extrapolate these linear dependencies and find their intersection point, that actually will be located at $(L(\overline{\text{AUC}}_{\max}), \overline{\text{AUC}}_{\max})$ coordinates in Fig. 7. Despite the statistical reliability (Table 3), it should be noted that the level of generalization of the results obtained for other use cases is unknown and should be thoroughly investigated in the future research. In general, the proposed optimization method can be used along with other tuning methods used by us and scientific community to get the better DNNs with higher performance [8, 14].

Conclusions. On the basis of the dataset published during Intel and MobileODT Cervical Cancer Screening competition on Kaggle platform the two following problems were considered: (i) multi-class classification problem and (ii) multi-label classification problem with image size optimization (on the basis of the updated 6-class labeling). In the former case, despite the quite predictable outcome as to the better metrics results obtained after training the standard DNNs with the initial weights obtained after pretraining on ImageNet dataset, the conclusion could be made as to possible useful usage of the smaller versions of DNNs (like MobileNetV2, NASNet-Mobile). They are usually more efficient for the smaller datasets that are quite usual case for the specific medical datasets like in the cervix use case considered here. That is why for the latter case, namely, for multi-label classification problem with image size optimization, EfficientNetB0 model was used as an example for investigation of the effective method of metric improvement by image size optimization. Finally, it allowed us to improve the mean AUC values by 2.7–2.8% (in comparison with the standard image sizes of $224 * 224$ pixels used for pretraining the standard models on ImageNet dataset) in the limits of the standard deviation 0.3–1.8% by the proper selection of the optimal input image size. In general, this approach based on averaging values of metrics obtained by various regimes of DNN training (with and

without DA) and verification (validation and testing) for the pairs of smallest and biggest available image sizes with further extrapolation of metric tendencies could be useful strategy for image size optimization in other use cases. But, again, the level of generalization of the results obtained for other use cases is unknown and will be thoroughly investigated in our future research.

Acknowledgements Supported by “Knowledge At the Tip of Your fingers: Clinical Knowledge for Humanity” (KATY) project funded from the European Union’s Horizon 2020 research and innovation program under grant agreement No. 101017453.

References

1. Aina OE, Adeshina SA, Aibinu A (2019) Deep learning for image-based cervical cancer detection and diagnosis—a survey. In: 2019 15th international conference on electronics, computer and computation (ICECCO). IEEE, pp 1–7
2. Alienin O, Rokovyi O, Gordienko Y, Kochura Y, Taran V, Stirenko S (2022) Artificial intelligence platform for distant computer-aided detection (CADe) and computer-aided diagnosis (CADx) of human diseases. In: The international conference on artificial intelligence and logistics engineering. Springer, pp 91–100
3. Alyafeai Z, Ghouti L (2020) A fully-automated deep learning pipeline for cervical cancer classification. *Expert Syst Appl* 141:112951
4. Bedell SL, Goldstein LS, Goldstein AR, Goldstein AT (2020) Cervical cancer screening: past, present, and future. *Sex Med Rev* 8(1):28–37
5. Bornstein J, Bentley J, Bösze P, Girardi F, Haefner H, Menton M, Perrotta M, Prendiville W, Russell P, Sideri M et al (2012) 2011 colposcopic terminology of the international federation for cervical pathology and colposcopy. *Obstet Gynecol* 120(1):166–172
6. Chen YW, Jain LC (2020) Deep learning in healthcare. Springer
7. Cleveland WS, Grosse E, Shyu WM (2017) Local regression models. In: Statistical models in S. Routledge, pp 309–376
8. Gang P, Zeng W, Gordienko Y, Kochura Y, Alienin O, Rokovyi O, Stirenko S (2019) Effect of data augmentation and lung mask segmentation for automated chest radiograph interpretation of some lung diseases. In: International conference on neural information processing. Springer, pp 333–340
9. Gorantla R, Singh RK, Pandey R, Jain M (2019) Cervical cancer diagnosis using CervixNet—a deep learning approach. In: 2019 IEEE 19th international conference on bioinformatics and bioengineering (BIBE). IEEE, pp 397–404
10. Gordienko Y, Kochura Y, Taran V, Gordienko N, Rokovyi A, Alienin O, Stirenko S (2020) Scaling analysis of specialized tensor processing architectures for deep learning models. In: Deep learning: concepts and architectures. Springer, pp 65–99
11. Gordienko Y, Kochura Y, Taran V, Gordienko N, Rokovyi O, Alienin O, Stirenko S (2021) “Last mile” optimization of edge computing ecosystem with deep learning models and specialized tensor processing architectures. *Adv Comput* 122:303–341
12. Kaggle competition: Intel & MobileODT cervical cancer screening (2022) <https://www.kaggle.com/competitions/intel-mobileodt-cervical-cancer-screening>. Accessed 30 June 2022
13. Keppler D, Lin AW (2015) Cervical cancer: methods and protocols. Springer
14. Kochura Y, Stirenko S, Gordienko Y (2017) Comparative performance analysis of neural networks architectures on H2O platform for various activation functions. In: 2017 IEEE international young scientists forum on applied physics and engineering (YSF). IEEE, pp 70–73

15. Pandian AP (2019) Identification and classification of cancer cells using capsule network with pathological images. *J Artif Intell* 1(01):37–44
16. Schmidt P (2017) Cervix EDA and model selection. Notebook. <https://www.kaggle.com/philschmidt>. Accessed 30 June 2022
17. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 71(3):209–249
18. Taran V, Gordienko Y, Rokovyi O, Alienin O, Kochura Y, Stirenko S (2022) Edge intelligence for medical applications under field conditions. In: *The international conference on artificial intelligence and logistics engineering*. Springer, pp 71–80
19. Van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9(11)
20. Vijayakumar T (2019) Neural network analysis for tumor investigation and cancer prediction. *J Electron* 1(02):89–98
21. William W, Ware A, Basaza-Ejiri AH, Obungoloch J (2018) A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. *Comput Methods Programs Biomed* 164:15–22
22. Wojtyla C, Janik-Konieczny K, La Vecchia C (2020) Cervical cancer mortality in young adult European women. *Eur J Cancer* 126:56–64
23. Xue P, Wang J, Qin D, Yan H, Qu Y, Seery S, Jiang Y, Qiao Y (2022) Deep learning in image-based breast and cervical cancer detection: a systematic review and meta-analysis. *NPJ Digit Med* 5(1):1–15

Faecal Image-Based Chicken Disease Classification Using Deep Learning Techniques



S. Suthagar , G. Mageshkumar , M. Ayyadurai , C. Snegha, M. Sureka, and S. Velmurugan

Abstract Various diagnostic methods are used in chickens, including counting oocytes in the stool or intestinal tract, detecting the virus and using polymerase chain reaction (PCR) procedures, which require multiple diagnoses. The diseases are transmitted through contaminated feed and excrement from infected poultry. As a result of late diagnoses or a lack of credible specialists, many domesticated birds are lost by farmers. The most common ailments affecting chickens can be easily identified in the pictures of chicken droppings using artificial intelligence and machine learning methods based on computer screening and image analysis. In this paper, a model for early detection and classification of poultry diseases with high accuracy using wildlife database is proposed. The dataset contains 6812 images of four different classes such as healthy chicken, Coccidiosis, Salmonella and Newcastle images is proposed. A deep learning method based on convolutional neural networks (CNN) is used to predict whether chicken faecal image belongs to any of the four categories. The pre-trained DenseNet model, the Inception model and the MobileNet model were used to predict whether chicken faecal belonged to four categories with minimum loss. In comparison with the above-mentioned models, DenseNet method produced the best results with an accuracy of 97% which is recommended for poultry diagnostic application.

Keywords Chicken · Deep learning · DenseNet model · Faecal disease · Inception model · MobileNet model

S. Suthagar (✉) · G. Mageshkumar · C. Snegha · M. Sureka · S. Velmurugan
Kongu Engineering College, Perundurai, Erode 638060, India
e-mail: susjou@gmail.com

M. Ayyadurai
SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
S. Smys et al. (eds.), *Inventive Computation and Information Technologies*, Lecture Notes
in Networks and Systems 563, https://doi.org/10.1007/978-981-19-7402-1_64

903

1 Introduction

A broiler is a bird that has been bred and kept only for the purpose of producing meat. Most commercial broilers achieve slaughter weight between the ages of four and seven weeks; however, slower-growing types reach slaughter weight around 14 weeks. Typical broilers have white feathers and yellowish skin. Due to the growing demand for egg and meat products, the fowl sector is one of the country's best livestock industries [2]. Poultry waste is a mixture of chicken manure, feathers, bedding, spilled food, medicines and water produced in large quantities during the process. Due to its relatively high nutrient content, especially nitrogen, due to its high protein content and amino acids, waste can be useful in soil supplementation as organic fertilizer and infections with *Escherichia coli*, Salmonella, Paratyphoid, Fowl Cholera and *Riemerella anatipestifer*. The most frequent disease in chicken is Mycoplasma, which causes necrotizing enteritis. Newcastle, Coccidiosis and Salmonella illnesses have a significant impact on poultry productivity. The infections have resulted in significant losses and the expenditure of a substantial sum of money to cure affected chickens, resulting in a low yield. The chicken disease has terrible financial consequences on farmers, resulting in significant losses incurred, the failure to compete in export markets and the transmission of disease to humans, as shown earlier. Laboratory procedures, such as the polymerase chain reaction (PCR) process, are currently used to diagnose poultry diseases. The findings take 3–7 days to arrive, making them time-consuming and costly. The high death of the fowls is due to the treatment of illnesses after late identification. Therefore, the objective of this research is to create a complete research project model for rapid illness identification in poultry [15]. Vaccinations were the most popular countermeasure, albeit do not work for all diseases and farmers want solutions to problems in order to increase yield. Deep learning's recent application in disease identification emphasizes the importance of contributing to accurate diagnoses.

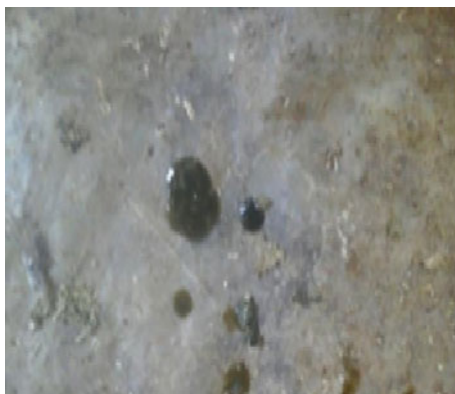
1.1 Coccidiosis

Coccidiosis is caused by Eimeria bacteria, which infect poultry intestines [11]. In chickens, there are seven different species of Eimeria. Coccidiosis is among the most prevalent reasons for the mortality of chickens, with *Eimeria tenella* being one of the most harmful parasites [4]. A common diagnostic technique for estimating the lesion score includes counting the percentage of embryos in the faeces. and evaluating the digestive tract. Bloody/brown diarrhoea is one of the clinical symptoms as shown in Fig. 1.

Fig. 1 Faecal image of chicken affected by Coccidiosis



Fig. 2 Faecal image of chicken affected by Newcastle



1.2 Newcastle

Newcastle disease is caused by the virulent Newcastle disease virus infecting domestic chickens and other bird species. It is a global epidemic that is often seen as a serious respiratory illness, but it may also appear as sadness, panic or diarrhoea [5]. The severity depends on the severity of the infection and the risk of the host. The incidence of this disease must be reported, and trade limits are set. One of the clinical indications of this condition is greenish watery diarrhoea as shown in Fig. 2.

1.3 Salmonella

Salmonella is a type of bacterial infection that causes illness in chickens and humans. *Salmonella pullorum* (SP) and *Salmonella gallinarum* (SG) viruses cause pullorum and poultry typhoid, respectively. *Pseudomonas enteritidis* and *Salmonella typhi* have been connected to human illnesses. ST strains are spread through a variety of poultry feed and poultry products. To detect and identify distinct Salmonella strains, the PCR technique is utilized. White diarrhoea is one of the clinical signs as shown in Fig. 3.

Fig. 3 Faecal image of chicken affected by Salmonella



2 Literature Review

Bao et al. [1] developed an artificial intelligence-based detection method for dead and unwell chickens on big farms. The use of artificial intelligence to detect sensors is proposed. The ZigBee network is used by the recovery term to gather data from the foot ring sensors. A machine learning system may be used to assess whether the chicken is dead or alive. Actual results show that the program is 91.6% accurate in caring for dead and sick chicks when compared to human performance, the cost of operating the program for four years can be cut in half. Zhuang and Zhang [18] proposed a supervised learning method that was used to detect unwell livestock and poultry. This study proposes a model framework for identifying sick broilers in a herd using digital imaging and in-depth study. For this, a chicken database (B1) was created with 14,728 images, while the PASCAL VOC 2012 dataset with 17,125 images was used as training background images. IFSSD can detect broilers and determine their health status at the same time. The model has a median accuracy (mAP) of 95.7% where the intersection-over-union (IoU) is more than 0.5 and 48.1% if the IoU is greater than 0.9. Overall, the data imply that unwell broilers in the herd can be spotted automatically, which has the potential to improve herd control. Zhuang et al. [17] have designed a model to identify extremely sick livestock and poultry, an advance warning algorithm that is being developed. Chicken photographs are captured, and two different types of classification methods are used to isolate the livestock and poultry from the environment so that the bone and shape information of the broiler chicken may be extracted. The test results mean that the algorithms created in this study can effectively discriminate against chickens in the background, retrieve poultry information and measure the health status of poultry accurately and quickly using SVM. Algorithms for digital image processing and machine learning are tested on deer health status and are optimized for accuracy, stability and normal performance, as well as the ability to deliver early warning signals.

Chen and Jiang [3] developed a Chicken Litters Microbiological Safety or Chicken Litters Microbiological Biofertilizers based on chicken litter are returned into the soil on a regular basis to help agricultural land enhance its structure and fertility. Hot processing of poultry manure or poultry manure before applying soil is an effective way

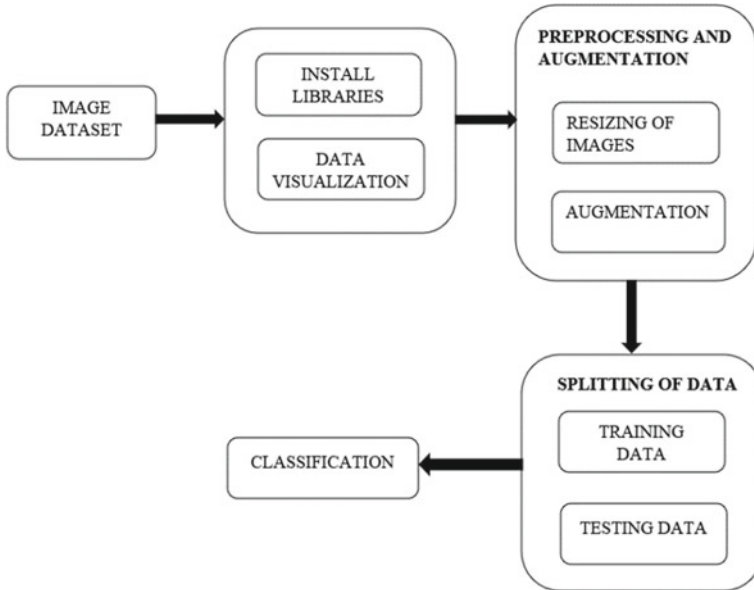


Fig. 4 Block diagram of the proposed system

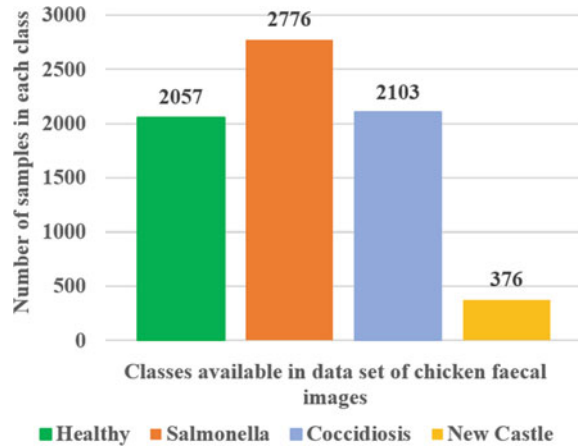
to disinfect. During building or composting, however, some persons may acquire accustomed to a hostile environment and develop heat resistance with opposite protection during the most recent heat treatment.

According to the literature survey conducted, it is found that the accuracy of the existing methods in identifying dead and unwell hens. They lack in ability to classify the diseases that are affecting the hens. Most of the techniques are computationally heavy and takes more time to produce result.

3 Methodology

A solution based on supervised learning is included in the suggested technique (CNN) to predict whether chicken poop is healthy, Coccidiosis, Salmonella or images of a Newcastle. Create a CNN from the ground up, then train it to create a model for detecting chicken sickness. Second, pretraining the learned models will increase the usage of deep learning algorithms. The sick chicken’s clinical indications could be intestinal or respiratory [14]. The colour of the droppings is affected by the presence of sickness in the chicken, and hence, this research focuses on the digestive symptoms of the disease. To detect illnesses in faeces photographs, DenseNet, MobileNet and Inception designs were used. The proposed method’s schematic representation is shown in Fig. 4.

Fig. 5 Number of samples taken in each category of dataset containing four classes of chicken faecal images



3.1 Data Collection

Faecal images were obtained from poultry farms and inoculation sites. There were 2057 healthy sample images, 2103 Coccidiosis images, 376 Newcastle images and 2776 Salmonella images as shown in Fig. 5. For the model development, a total of 6812 images were collected. The model using the transfer learning approach, the dataset was deemed appropriate. In addition, a variety of methods have been used to increase the size of the database.

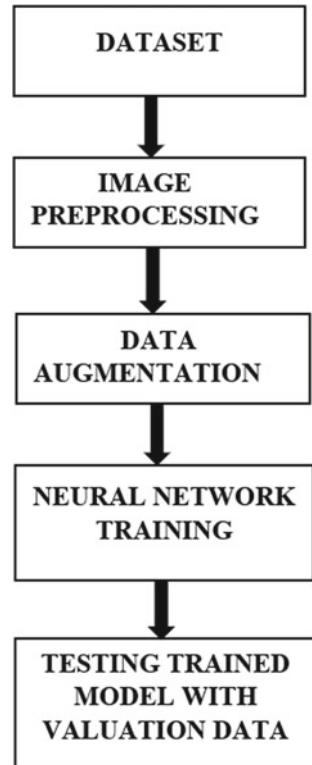
3.2 Research Pipeline

The pipeline is the framework that depicts the research strategy. To begin, gather data (faecal photographs) and create a dataset. After that, the dataset is pre-processed by being labelled, scaled or supplemented [7]. Data that have been pre-processed are saved in the cloud. Google Drive was used to save the information. Before the models were trained, the database was separated into train and test sets. Data are generated on neural networks that have already been trained in this project. The flowchart of the proposed method is shown in Fig. 6.

3.3 Dataset

At all stages of object recognition research, from the training phase to the evaluation of the performance of awareness algorithms, appropriate datasets are required. Poultry faecal pictures make up the annotated collection of poultry disease diagnostics

Fig. 6 Process flow of the proposed system



for small- and medium-scale poultry farms. The healthy class of normal faeces was collected from chicken farms, as was the cocci class of Coccidiosis sickness. After one week, the hens were infected with Salmonella and faecal photographs were collected from the ill chicks for the salmo class. The hens were also given a Newcastle disease vaccine, and faeces for the NCD class were obtained three days later. Some of the images used in the collection were retrieved from the Internet and found by searching for them by name in a variety of languages. The annotated dataset of poultry disease diagnostics for small- to medium-scale poultry farmers consists of poultry faecal images. The photographs of chicken faeces were obtained between September 2020 and February 2021 in Tanzania’s Arusha and Kilimanjaro areas, using the Open Data Kit (ODK) programme for mobile phones, which is available on the Zenodo Website. The ‘healthy’ class of normal faeces was collected from chicken farms, as was the ‘cocci’ class of Coccidiosis sickness. After one week, the hens were given a Salmonella infection and faecal photographs were collected from the ill birds for the ‘Salmo’ class. The hens were also given a Newcastle disease vaccine, and faecal pictures for the ‘NCD’ group were collected in three days time [8].

3.4 Image Preprocessing

Images acquired online came in a variety of formats, resolutions and quality levels [16]. The final images intended for use as a deep neural network separator database were pre-processed like removal of noise in the images using various filters like mean, median, anisotropic diffusion method can be done to achieve consistency and improve features [9]. In addition, image processing in the past involved capturing images in order to highlight the region of interest.

3.5 Image Augmentation

Image augmentation can be used to enhance the training dataset automatically by generating modified versions of images on the database. The primary goal is the supplement is to enhance the size of the dataset and add a slight distortion to the images, which helps to reduce overcrowding during the training phase. Image data augmentation is a technique for increasing the amount of training data by creating customized versions of images in the database by performing some operations such as crop, rotation, shift, increase brightness, zoom-in and zoom-out [10]. In-depth network neural network models trained in additional data are competent, and extension processes can provide image diversification that helps appropriate models make new images more numerous. Augmentation is done with a strong training system to provide accurate images based on a faecal image database. There is no augmentation in the verification and testing of sets.

3.6 Neural Network Training

The goal of network training is to teach the neural network the properties that distinguish one class from the others [13]. As a result of the increased use of augmented pictures, the network odds of learning the required features have increased. In four classes, Healthy, Salmonella, Newcastle and Coccidiosis and three baseline models, MobileNet, DenseNet and Inception, were trained. These models are chosen for their special features. MobileNet combines the use of pointwise and depthwise convolution to reduce computation. Inception model has less computation cost. DenseNet provides better accuracy when the loss is due to vanishing gradient.

3.7 Transfer Learning

CNN is frequently used in huge datasets with far more than almost one million pics, and it performs better when the layers are dense and interconnected. In the health

sector; however, it is challenging to find a significant collection of images that have been carefully tagged. As a result, CNN-based methods for medical problems like illness categorization are immediately limited to a relative number of datasets which can easily lead to problems with overcrowding of deep networks, and thus, the results obtained are not solid and scientific [12]. Transferred learning is a form of in-depth learning in which a mono model is used to train a production version. Substantial can not only prevent congestion problems when the data are tiny, but it can also cut trained time can be reduced.

3.8 Inception V3 Architecture

Inception V3 is a 48-layer deep convolutional neural network and uploads a pre-trained network version to over 5000 images from the ImageNet website. The input size of the image is $224 * 224 * 3$ used for this design. Globalized global integration will be used in the final convolutional layer output, and with that, the output of the Inception model will be a 2D tensor. The exit layer of the starting v3 model is flat; the size of the compound is reduced; a compact layer with the function of the rest of the compound is added to the Inception V3 model; and the convolution layer is expanded. The number of neurons is 1024. During the training phase, the dropout layer changes the input values to 0 and the frequency ranges to 0 randomly for every step, preventing congestion. 2,102,276 trainable parameters and 21,802,784 untrained parameters in the pre-trained V3 model.

3.9 DenseNet-121 Architecture

DenseNet works with the idea that convolutional networks can be more in-depth, more accurate and more efficient to train if they have short connections between near and input layers and those near output. DenseNet-121 was chosen because, while having 121 layers, it has fewer parameters than other models and is well-suited to dealing with the vanish gradient problem. Faecal images are 32×32 and 3 channels (RGB). Image feeding the upgrade to 224×224 is done in a double rendering mode. Class labels are converted to one hot code coding. In the DenseNet-121 architecture, the last fully connected layer was completed, and in its place, creating a dense layer of 1024 nodes, then a dense layer of node 128, and finally, a dense layer with softmax output on the output. The last layer used ReLU activation. Collection routine and 40% density of dense layers were also used. The training was done in a group size of 32 and 10 epochs. There are only 148,228 training parameters in the dense layer parameters. There are 7,037,504 untrained parameters. These value parameters will be taken directly from the pre-trained DenseNet model.

3.10 MobileNet Architecture

The MobileNet model is downloaded using the MobileNet function in the TensorFlow Keras applications module. After loading the original model, the summary is summarized. The images in the feecal images are 32×32 and 3 channels (RGB). Feature photography upgrades to 224×224 are done in duplicate translation. Class labels are converted to one hot code coding. MobileNet model layers used for 1000 classes have now been removed and need to add new layers that can work with 4 classes in the database. The modified model contains a variation of the basic model, modified by adding new layers at the top to make it fit the dataset. The basic variable model at the top of the list is followed by at least one FC layer with the same number of neurons as the number of classes in the database. There are two layers which are a central integration layer and a fully integrated layer of 30 neurons using a compact class. Only 17,467,780 parameters are trained in FC layer parameters. Combining layers do not have training parameters. There are 3,228,864 untrained parameters. These price parameters will be taken directly from the pre-trained MobileNet.

4 Results and Discussion

4.1 Training and Testing the Model

Three baseline models, MobileNet, DenseNet and Inception, were trained in four classes: Healthy, Salmonella, Newcastle and Coccidiosis. The healthy images are 2057, the Coccidiosis images are 2103, Salmonella images are 2276 and Newcastle images are 376 as shown in Table 1. The dataset was split while training the models, with 70% used for training and 30% used for validation.

Different hyperparameters such as learning rate, epochs, batch size and optimizer were used to train the baseline models as shown in Table 2.

- **Learning Rate:** When the model weights are changed, the parameter determines how much the model should change in response to the estimated mistake.
- **Epoch:** The number of iterations a model undergoes when training a batch of data.
- **Batch Size:** In a single round, the total sample size is sent via the network.

Table 1 Dataset split for training and validation of the proposed model

Class	Number of images	Training samples	Validation samples
Healthy	2057	1439	618
Salmonella	2276	1593	683
Coccidiosis	2103	1472	631
Newcastle	376	263	113

Table 2 Hyperparameter used to train MobileNet, DenseNet and Inception models

Parameter	Value
Learning rate	0.001
Epochs	10
Batch size	132
Optimizer	Adam

- **Optimizer:** An extended class that contains additional information for training a certain model. SGD, Adam and RMSprop are some examples of regularly used optimizers.

4.2 Classification Results and Analysis

To enhance the size of the dataset, the photographs with several strategies when training the model are enriched. A large number of datasets have two possible advantages: first, it prevents over-fitting, and second, it allows the model to learn from previously unknown datasets. Picture flipping, cropping and padding, as well as adding random image saturation, are among the augmentation techniques employed in our model.

4.3 Evaluation of Models

The performance of the models was tested. In testing the models, precision metrics were used. From the first to the 10th period, the accuracy of the training, the loss of training, the accuracy of the verification and the loss of validation are all measured during training and validation as shown in Table 3.

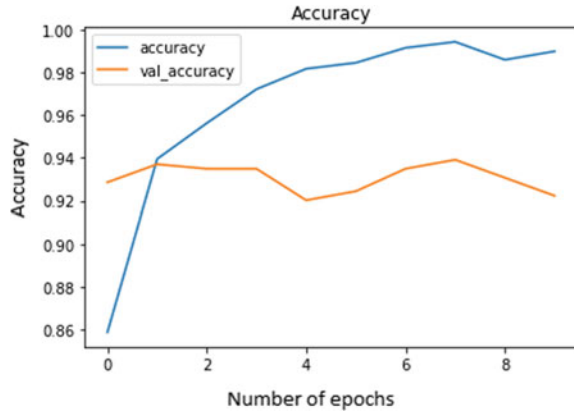
- **Training Accuracy:** When the model is applied to the training data, it achieves accuracy.
- **Training Loss:** This is the error on the data training set.
- **Validation Accuracy:** This is the precision with which the model performance is assessed.
- **Validation Loss:** After passing the validation set of data through the trained network, this is the error.

After pretraining the three models such as MobileNet, DenseNet and Inception, DenseNet-121 has the higher training and validation accuracy; it gives the best performance among those three models, so it is considered the best model for the classification of chicken diseases.

Table 3 Performance comparison of MobileNet, DenseNet and Inception models for the classification of chicken faecal images

Models	Training accuracy	Training loss	Validation accuracy	Validation loss
Inception V3	0.978	0.07	0.92	0.18
MobileNet	0.993	0.009	0.94	0.62
DenseNet 121	0.997	0.002	0.97	0.30

Fig. 7 Accuracy of training and validation of DenseNet model



The validation and training and accuracy are increased from the initial point which is moving in the range of 90–100. Losses can be seen as the distance between true class values and model-predicted values. The greater the loss, the greater the errors that you have made in the data. Hence, the training loss is decreasing enormously and reaches saturation and the validation loss increases at first, and after a few epochs, it remains as constant as shown in Figs. 7 and 8. The accuracy of the classifier model is calculated using the formula given in Eq. 1

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{1}$$

4.4 Confusion Matrix

The confusion matrix is an N matrix multiplication matrix used to check the functioning of a split model, where N is the number of class labels. The matrix compares the actual numerical value to the predictions of the classification model. This provides us with a complete picture of how well the classification algorithm operates and what

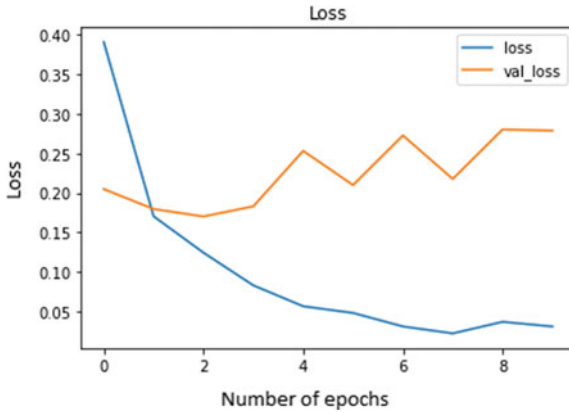


Fig. 8 Loss of training and validation of DenseNet model

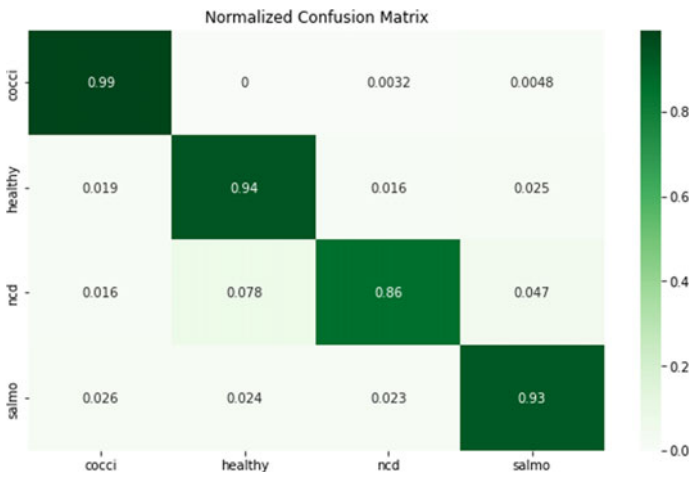


Fig. 9 Confusion matrix of DenseNet model

sorts of issues it makes. 99% of coccidiosis images, 94% of healthy images, 86% images of Newcastle disease (NCD) images and 93% of Salmonella images were correctly classified by the model as shown in Fig. 9.

4.5 Classification Report

In deep learning, the classification report is an indicator of performance testing [6]. Used to illustrate the correctness, retention, *F1* score and support of your trained classification model. It gives us a clearer understanding of our pre-trained model's overall performance as shown in Table 4.

Table 4 Classification report of DenseNet model

	Precision	Recall	F1 score	Support
Cocci	0.98	0.98	0.98	649
Healthy	0.92	0.95	0.94	601
NCD	0.88	0.62	0.73	103
Salmo	0.93	0.94	0.94	691
Macro average	0.93	0.87	0.89	2044
Weighted average	0.94	0.94	0.94	2044
Accuracy	–	–	0.97	2044

5 Conclusion

In this research, we describe an innovative chicken sickness approach to detection based on a pre-trained network and learning techniques. The use of machine technology and reliable data is two key aspects that can improve field officers’ and poultry growers’ efficiency in the rapid diagnosis of chicken illnesses. A modern-day demand is the development of image processing technology that can assist farmers. These approaches are beneficial in reducing losses and increasing output, and it is obvious that infections may be recognized early on before causing chicken fatalities. For the past several decades, machine vision research has been attempting to close the gap by developing automated systems that can interpret pictures and make decisions using computers. To create the CNN model specifically, which learns the hidden pattern among the many faecal photographs in our dataset. Coccidiosis, health, Newcastle and Salmonella are the four categories predicted by the supervised learning method. According to the findings, the proposed model is accurate up to 97% of the time. Our the suggested method outperforms the Inception, DenseNet and MobileNet designs by a wide margin. The results of the experiments show that our technology is effective in detecting illnesses in chickens and may be used for reliable diagnostics.

References

1. Bao Y, Lu H, Zhao Q, Yang Z, Xu W (2021) Detection system of dead and sick chickens in large scale farms based on artificial intelligence. *Math Biosci Eng* 18(5):6117–6135
2. Bhatnagar N, Ryan D, Murphy R, Enright AM (2020) Trace element supplementation and enzyme addition to enhance biogas production by anaerobic digestion of chicken litter. *Energies* 13(13):3477
3. Chen Z, Jiang X (2014) Microbiological safety of chicken litter or chicken litter-based organic fertilizers: a review. *Agriculture* 4(1):1–29
4. Fatoba AJ, Adeleke MA (2018) Diagnosis and control of chicken coccidiosis: a recent update. *J Parasit Dis* 42(4):483–493
5. Gohm DS, Thür B, Hofmann M (2000) Detection of Newcastle disease virus in organs and faeces of experimentally infected chickens using RT-PCR. *Avian Pathol* 29(2):143–152

6. Jacob IJ, Darney PE (2021) Design of deep learning algorithm for IoT application by image based recognition. *J ISMAC* 3(03):276–290
7. Kyakuwaire M, Olupot G, Amoding A, Nkedi-Kizza P, Ateanyi Basamba T (2019) How safe is chicken litter for land application as an organic fertilizer? A review. *Int J Environ Res Public Health* 16(19):3521
8. Machuve D, Nwankwo E, Mduma N, Mbelwa H, Maguo E, Munisi C (2021) Machine learning dataset for poultry diseases diagnostics. <https://doi.org/10.5281/zenodo.4628934>
9. Mageshkumar G, Suthagar S, Tamilselvan K (2018) Performance comparison of adaptive filters for speckle noise reduction in SAR images. In: 2018 international conference on intelligent computing and communication for smart world (I2C2SW). IEEE, pp 195–197
10. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6(1):1–48
11. Stephens C, Hampson D (1999) Prevalence and disease association of intestinal spirochaetes in chickens in eastern Australia. *Avian Pathol* 28(5):447–454
12. Suthagar S, Snegha C, Sureka M, Velmurugan S (2022) Analysis of breast cancer classification using various algorithms. In: 2022 6th international conference on computing methodologies and communication (ICCMC). IEEE, pp 1286–1291
13. Suthagar S, Tamilselvan K, Balakumar P, Rajalakshmi B, Roshini C (2020) Translation of sign language for deaf and dumb people. *Int J Recent Technol Eng* 8(5):4369–4372
14. Suthagar S, Tamilselvan K, Priyadharshini M, Nihila B (2021) Determination of apple, lemon, and banana ripening stages using electronic nose and image processing. In: *Innovations in cyber physical systems*. Springer, pp 755–769
15. Weldekidan H, Strezov V, Li R, Kan T, Town G, Kumar R, He J, Flamant G (2020) Distribution of solar pyrolysis products and product gas composition produced from agricultural residues and animal wastes at different operating parameters. *Renew Energy* 151:1102–1109
16. Zhang Z, Han Y (2020) Detection of ovarian tumors in obstetric ultrasound imaging using logistic regression classifier with an advanced machine learning approach. *IEEE Access* 8:44999–45008
17. Zhuang X, Bi M, Guo J, Wu S, Zhang T (2018) Development of an early warning algorithm to detect sick broilers. *Comput Electron Agric* 144:102–113
18. Zhuang X, Zhang T (2019) Detection of sick broilers by digital image processing and deep learning. *Biosyst Eng* 179:106–116

Fake News Detection Using Supervised Machine Learning Classification Algorithms



Akanksha Singh and Sanjay Patidar

Abstract The online world has become an essential part of everybody's life in today's society. Almost everyone is using social media to gain information about all around the world. The news or information shared on the Internet could be beneficial as well as harmful. On one hand, it is the most inexpensive, easy, and convenient way of getting information in no time; on the other hand, it also prevails fake news. False news/information has a tremendous impact on our social lives, in fact, in all fields, particularly politics and education, and organization. The propagation of false information has the potential to create significant social and emotional harm, as well as have potentially dangerous consequences. Spreading incorrect data via online media to stand out enough to be noticed or monetary and political increase is common on social media these days. The focus of the research is to develop a detection that incorporates multiple machine learning classification methods in order to present an analytical method for detecting false news. To identify fake news, we used five machine learning classification approaches. The five supervised ML classifiers are logistic regression, decision tree, Naïve Bayes, Random Forest, and support vector machine. We have calculated the accuracy of different classifiers and gave a comparative analysis of accuracy along with all performance measures. The model's output has a 99.59% accuracy when employing feature extraction approaches like term frequency and inverted document frequency, and a decision tree as a classifier.

Keywords Fake news · Logistic regression · Decision tree · Naïve Bayes · Random forest · Support vector machine

1 Introduction

For a long time, social media has taken over a meaningful place in people's life. Fake news primarily prevails via social media and articles available online. Fake news indulges politics, democracy, education as well as finance and business at risk.

A. Singh · S. Patidar (✉)

Department of Software Engineering, Delhi Technological University, New Delhi, India
e-mail: sanjaypatidar@gmail.com

Even while false news is not a new issue, people these days place a larger focus on social media, which leads to the acceptance of deceitful remarks and the subsequent propagation of the same wrong information. It's getting harder to tell the difference between accurate and misleading news these days, which leads to confusion and complications. Manually recognizing fake news is tough; it is only achievable when the individual identifying the news has extensive expertise in the subject. Fake news can destroy someone's career and if it is political and harm the nation and citizens of that country as well as it can also affect businesses, products, and reputations. It is now easier to manufacture and circulate fake news because of the recent advances in computer science, but it is much more difficult to determine if the information is accurate or not.

As a result, we carried out and compared five various methods in this study. Here, we are using supervised ML classification techniques, i.e., logistic regression, decision tree, Naïve Bayes, Random Forest, and support vector machine to determine if the news being transmitted is authentic or not. We have used a dataset available on the Kaggle website which contains two datasets for there are both authentic and fraudulent news items.

The paper structure could be defined here given there are Sect. 5 in this paper. In Sect. 1, we have provided the formal introduction of the research being carried out and its motive. In Sect. 2, we have discussed the relevant research that has been done in this field and the algorithm that we used in this project. Then in Sect. 3, we've gone through the methodology, which includes information on the flowchart, dataset, and machine learning techniques that we employed in this research. In Sect. 4, we have discussed the implementation and the results obtained in this study. After that in Sect. 5, we have concluded the study so far and the results with accuracy. Also, we proposed the future work for the study.

2 Literature Survey

The primary goal of this study is to discover the best effective classification system for detecting and quantifying false news. To find out, we looked at several classification techniques and used them in our model. We have applied five classification techniques here. Further, we are providing a brief review of the papers we have studied.

In 2017, Granik and Mesyura presented a methodology for detecting bogus news utilizing Naive Bayes on news posts on Facebook and got a 74% accuracy rate. And concluded that AI techniques could be used successfully to handle these kinds of problems [1].

In 2017, Ahmed, Traore, and Saad have suggested a methodology for identifying false news, researchers have developed an AI model using n-gram analysis. The classifier used here was the support vector machine, which has a precision of 92% [2].

In 2017, Campan et al in their study proposed a model how fake news spread on social media and how the Internet affects the diffusion of false information in creating and spreading. They also discussed the solutions to reduce the dissemination of false information and provided the future research aspects in this area [3].

In 2017, Perez-Rosas et al, Klienberg, and colleagues suggested a model that automatically detects fake news for online resources. They developed a computer algorithm and tools to detect bogus news. They work with two different datasets. The first came via the Internet, and the second resulted from a mix of human data collecting and Internet assistance [4].

In 2018, Aphiwongsophon and Chongstitvatana have a study on using Naïve Bayes, SVM, and neural networks to detect the fake news and calculated the performance measures they have found that Naïve Bayes has 96.08% and neural network and SVM 99.90% accuracy. Through this experiment, they found out that neural networks and support vector machines are having significant accuracy and high confidence [5].

In 2018, Gahirwal and colleagues suggested a support vector machine news detection based model for false or real news that has an accuracy of 87%. She had recognized comedy, negative words, ridiculousness, syntax, and punctuation using five predictive features. Its goal was to ensure that the substance of a news piece was accurate [6].

In 2019, Ozbay and Alatas have used AI techniques for detecting fake news. In the first phase, they preprocessed the dataset to transform unstructured data into structured data, and then they used text mining to construct about twenty-three supervised AI algorithms. They applied these algorithms to about three real-world datasets and found the accuracy and performance measures accordingly. The best average value they got was by using a decision tree, ZeroR, CVPS, and WIHW algorithms [7].

In 2019, Agarwal et al and colleagues taken a dataset namely the Liar dataset, and given a comprehensive study of various approaches. They offer a stacking model in this research that fine-tunes the informational knowledge obtained from user input at each level before attempting to predict something [8].

In 2019, Riece et al are working on looking for a range of elements in news articles, postings, and stories that might assist to identify false news with increased precision. He demonstrated the significance of these new qualities in evaluating bogus news. Discrimination, integrity, involvement, domain location, and temporal patterns are some of these characteristics. They used 2282 BuzzFeed items in their analysis (news articles). Using KNN, Naive Bayes, Random. Forest, XGBoost, support vector machine (SVM), and they analyzed and described the strengths and limits of this technology and discovered that XGBoost performed better when compared with other with an accuracy of 0.86 [9].

In 2020, Zhang and Ghorbani's study elaborates that false information has been a serious concern for the industry as well as academia as it is widely utilized to confuse and persuade online users with skewed facts. Furthermore, the Internet generates and disseminates a vast volume of fantastic and incorrect information. It has emerged as a potential danger to social networking groups and has had a major adverse influence on online activities such as online commerce and networking sites [10].

In 2020, Shaikh and Patil have used three AI techniques for giving the detection model for the detection of false news. Three algorithms that they used were SVM, Naïve Bayes, and passive aggressive classifier, respectively, with SVM giving the highest accuracy of 95.05% [11].

In 2020, Smitha and Bharath have illustrated the model and different methodologies to identify and quantify fake news with the help of ML techniques and NLP techniques. Seven different classification algorithms are proposed here, and accuracy, F1score, recall, and precision are compared [12].

In 2020, Kesarwani et al demonstrated a basic strategy for detecting false news on social media using a K -nearest neighbor classifier, which obtained an accuracy of roughly 79% when evaluated against a sample of Facebook news articles [13].

In 2021, Khanam et al and colleagues carried out the research by reviewing it in two phases: firstly, they used multiple supervised learning algorithms to define the essential principles and criteria of false news found in web-based media. They proposed using scikit-learn library for processing text data. They performed techniques for feature selection to select the best fit [14].

In 2021, Nagaraja and colleagues showed in their study that false information mostly circulates through social media and is propagated further without investigating the true data. They applied various NLP techniques and two ML algorithms, i.e., Naïve Bayes and SVM which gives 63% and 75% accuracy, respectively [15].

A significant amount of prior and ongoing studies is based on fake news detection. The misleading information has always been a serious concern worldwide due to its bad influence on social, religious, educational and civilization, and many more fields. We studied research papers in order to carry out this study and filtered out the papers for extensive literature review and summarized them. The proposed model in this study has been drawn from various research being done, and the machine learning algorithms that have been applied in research papers to propose a better way for detection of fake news. We have taken the supervised algorithms which performed best in various study given in [2, 5, 6, 7, 9, 11 and 15] and implemented those algorithms to find out the results.

3 Methodology

Various stages involved in this experiment are given in the following flowchart Fig. 1.

3.1 Dataset

We obtained the dataset for this study from the Kaggle website [16] which contains two files. Out of these two, one contains real news articles and another one contains fake news articles. Real articles are around 21,417 and fake articles are around 23,481

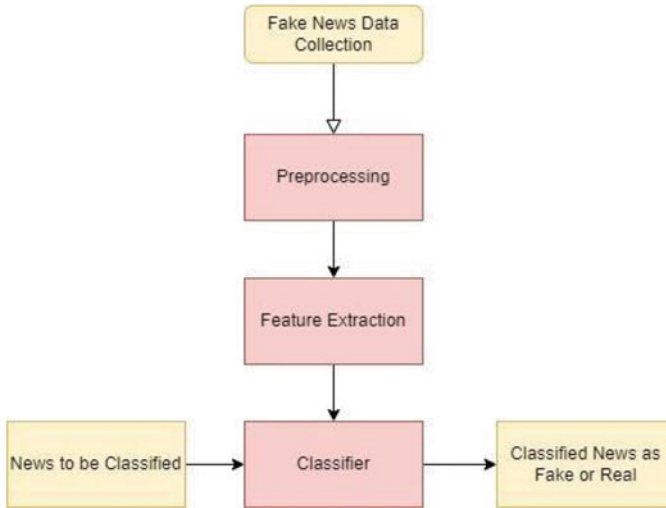


Fig. 1 Flowchart for fake news detection

	title	text	subject	target
0	While You Were Distracted, Republicans Gutted...	On Monday, Missouri Gov. Eric Greitens signed ...	News	fake
1	WATCH WHAT HAPPENS WHEN CHRISTIAN MAN ASKS 13 ...	I was wondering if I could get a cake that sa...	politics	fake
2	LISTEN: Dispatch Audio Reveals Cop Who Killed...	As much as Jeronimo Yanez, the police officer ...	News	fake
3	South Africa's Zuma asks court to set aside re...	JOHANNESBURG (Reuters) - South African Preside...	worldnews	true
4	Bombs kill at least 17 people in Somali capita...	MOGADISHU (Reuters) - Two car bombs killed at ...	worldnews	true

Fig. 2 Head of the data

with a total of 23,481. To further proceed with both the datasets, we had combined the dataset that contains the combination of both fake as well as real. Figure 2 shows the head of our data.

The dataset we are using is having five features, i.e., title, text, subject, date, and target. So, we dropped the unnecessary columns, i.e., date and title as we will be working with only text here.

The data we have used here is available in text as we know that text data requires preprocessing in order to be changed over into a suitable structure for information display. There are several methods for transforming text data, including natural language text processing approaches, which we employed. After removing the date and title columns that were no longer needed for modeling and converting the text to lowercase. We have performed stop word removal because there are many words in a text that occur very frequently in a document and have not much information such as ‘a,’ ‘is,’ ‘the,’ and ‘am.’ To improve the accuracy of analysis, these words are generally ignored using natural language toolkit (NLTK) library for stop word removal. After that, here are many words in a text that occur very frequently in a document

and have no much information such as ‘a,’ ‘is,’ ‘the,’ and ‘am.’ To improve the accuracy of our analysis, these words are ignored using natural language toolkit (NLTK) library for stop word removal. After that, punctuation removal was performed as punctuation like commas and full stops don’t add much importance to text so they should be filtered out. We split the data into train and test before feeding it into the machine learning model. We have separated the 30% data into test set and 70% into train set. A subset of the dataset used to train the model has already revealed the outcome. The detection model is tested on a subset of the dataset, and the test set is utilized to forecast the outcome. For feature extraction, we have concentrated on two distinctive choice techniques: term frequency and term frequency inverted document frequency. This retrieval methodology considers the frequency of a phrase as well as the inverse document frequency.

A term’s frequency in a text may be determined using term frequency. The term n in the formula represents the number of times the phrase appears in each document or text. As a result, each term has a TF value.

3.2 Machine Learning Algorithms

In this paper, we are presenting five different supervised machine learning classification algorithms. The following is a quick rundown of all the algorithms used:

3.2.1 Logistic Regression

It’s a tool for categorizing binary data. For binary classification, usually, linear regression is used to create the best fit line. When two classes can be separated linearly, logistic regression is used. It is within the supervised machine learning algorithm category. It’s a machine learning-based categorization problem-solving approach. In logistic regression, a type of predictive analysis, the probability assumptions are applied. To complete a binary classification job, a linear equation is used as input, and the logistic function and log odds are used in the logistic regression model. It employs a more complicated function when compared to linear regression.

$$y = d_0 + d_1 * x \tag{1}$$

$$Q = 1/(1 + e^{-y}) \tag{2}$$

$$\ln(Q/(1 - Q)) = d_0 + d_1 * x \tag{3}$$

where in Eq. (3), d_0 is slope, d_1 is intercept, and x are a data point. Equation (2) is a sigmoid function where Q has been used to eliminate the outlier’s effect.

3.2.2 Decision Tree

It's an ML supervised classification algorithm which means we have to clarify what the information is and what the relating yield is in the preparation information. It is a tree-like construction where the information is consistently parted by a specific boundary. The elements of a dataset are addressed by the inner nodes and branches that address the decision rules and each leaf addresses the ultimate results or choices.

3.2.3 Naïve Bayes

The Naive Bayes approach, a supervised machine learning methodology based on the well-known Bayes theorem, is used to tackle classification issues. It is most commonly used for text classification with a big training dataset. One of the most simple and effective classification methods is the Naive Bayes classifier. It allows for the rapid building of machine learning models as well as effective training and testing to make speedy predictions. It's a probabilistic classifier, which implies the algorithm's whole basis is built on probabilities that have been computed, and it predicts based on an item's likelihood.

Naïve Bayes Equation:

$$P(R|S) = P(S|G) * P(R)/P(S)$$

where $P(R|S)$ is the posterior probability. $P(S|R)$ is the likelihood. $P(R)$ is the class prior probability. $P(S)$ is the predictor of prior probability.

3.2.4 Random Forest

It is a supervised ML technique. It is basically established on the outfit learning techniques where different classifiers are united to deal with an issue and to chip away at the display of the presentation of the model. Random Forest is a classifier that calculates the dataset's predicted precision by averaging the results of many decision trees applied to different subsets of the dataset.

3.2.5 Support Vector Machine

This approach aims to find a hyperplane (where N is the number of characteristics) that clearly arranges the principal elements in an N -dimensional space. There is an assortment of hyperplanes from which to isolate the two kinds of informative items. Our point is to track down the plane with the biggest edge or distance between relevant items from the two classes. Boosting the edge distance gives some support, making it more straightforward to arrange ensuing information points.

3.3 *Evaluating Measures*

Evaluation metrics are frequently used to assess categorization performance. As a result, performance measurements are the most common. So, we have used different metrics to evaluate our classifiers given as follows.

3.3.1 Accuracy

It gives the comparison of actual and predicted labels, i.e., it measures how often a classifier predicts accurately. It can be formulated as

$$\text{Accuracy} = \text{Correct prediction} / \text{Total data points}$$

3.3.2 Precision

Precision is a proportion to tell how exact a classifier is performing. Precision P can be formulated as the ratio of total true positives to total predicted instances

$$\text{Precision}(P) = \text{TP} / (\text{TP} + \text{FP})$$

3.3.3 Recall

It tells what percentage of positive instances were successfully identified and its formula is

$$\text{Recall}(R) = \text{TP} / (\text{TP} + \text{FN})$$

3.3.4 F-Measure

It is represented as a harmonic mean of precision and recall and can be formulated as

$$F - \text{Measure}(F) = 2P \cdot R / (P + R)$$

where

TP (True Positive) belongs to a class it belongs actually,

FP (False Positive) belongs to a class it doesn't belong actually, FN (False Negative) doesn't belong to a class it actually should belong, and TN (True Negative) doesn't belong to a class it actually doesn't belong.

4 Implementation and Results

In Figs. 3, 4, 5, 6 and 7, we are giving the various confusion matrices that we obtained after applying five supervised machine learning algorithms. As discussed in methodology, each confusion matrix contains the four values TP (True Positive), FP (False Positive), TN (True Negative), and FN (False Negative). In our experiment, TP represents that the news which was actually fake is also predicted as fake. FP represents that the news which was actually real is predicted as fake. FN represents that the news which was actually fake but is predicted as real. TN represents that the news which was actually real is also predicted as real. The real and fake true label and predicted label are shown in confusion matrix. We have implemented five supervised machine learning algorithms here, i.e., logistic regression (LR), decision tree (DT), Naive Bayes (NB), Random Forest (RF), and support vector machine (SVM).

Figures 8 and 9 represent the word cloud obtained from both real and fake news sets, respectively, to represent the significant textual data points/words in the dataset.

Table 1 indicates the accuracy % and performance metric of all five classifiers, i.e., precision, recall, and F1-score.

As described in the methodology, we have implemented all the five ML classification algorithms and calculated the accuracy and performance measures. From

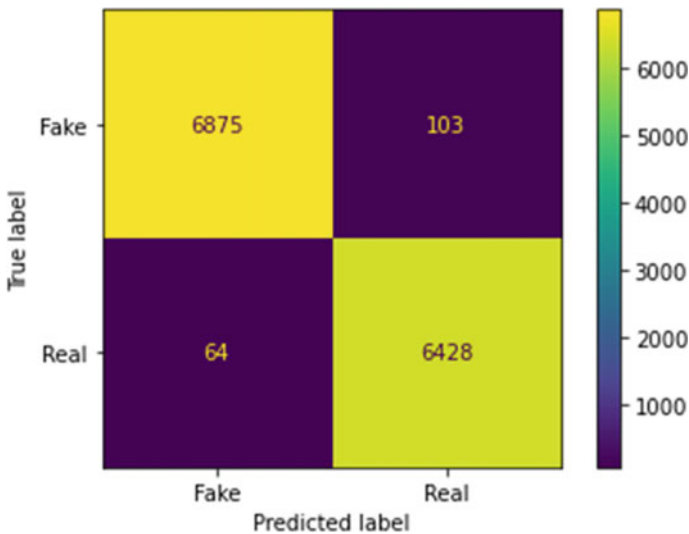


Fig. 3 Confusion matrix using LR

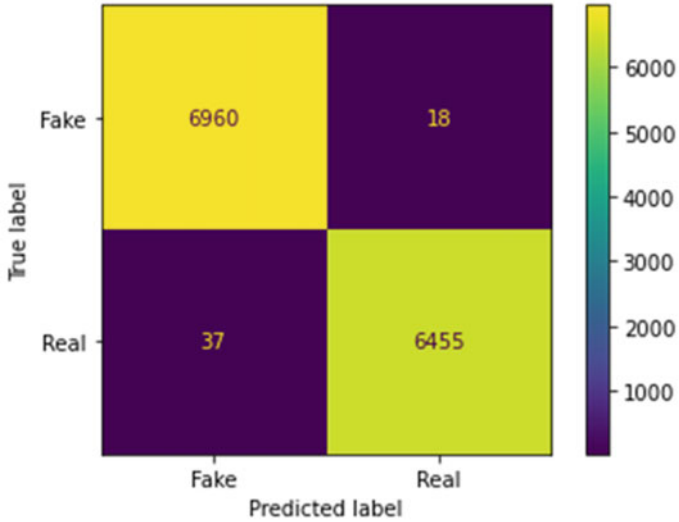


Fig. 4 Confusion matrix using DT

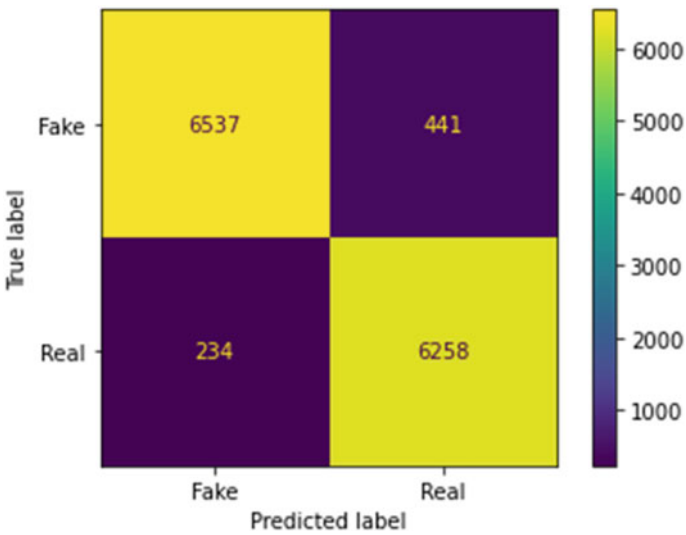


Fig. 5 Confusion matrix using NB

Table 1, we can see that decision tree outperforms here which has the accuracy 99.59% and all performance metrics (precision, recall, and F1-score) as 1.00 perform the best. After decision tree, SVM performs with a very negligible difference in accuracy when compared. Naïve Bayes has the lowest accuracy, i.e., 94.99% and all performance metrics, i.e., (recall, F1-score, and precision) as 0.95.

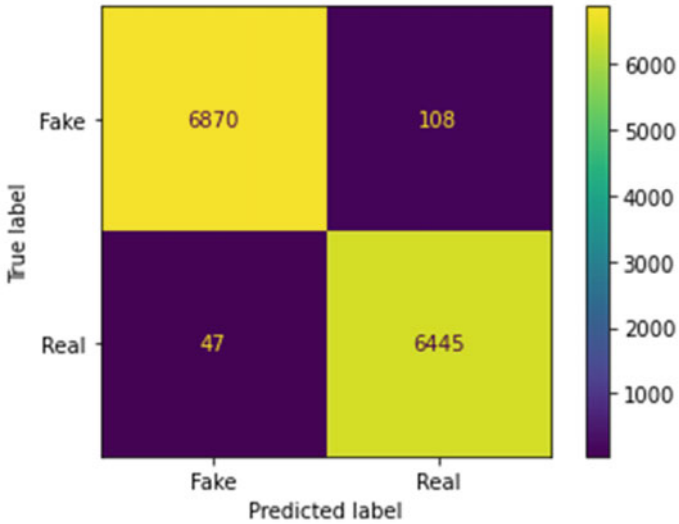


Fig. 6 Confusion matrix using RF

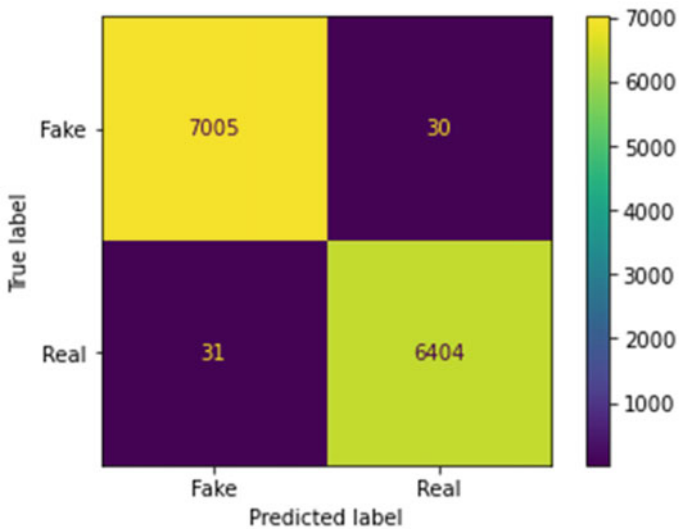


Fig. 7 Confusion matrix using support vector machine

Figure 10 graph shows the comparison of all the models used in this experiment which shows that decision tree has the highest accuracy of 99.59% and the Naïve Bayes classifier has the lowest accuracy of 94.99%.

In our experiment, decision tree is performing best. As the data is categorical that is the news is either fake or real, so in such cases this algorithm performs

Table 1 Accuracy and performance measures

ML classifier	Accuracy (%)	Precision	Recall	F1 Score
Logistic regression	98.76	0.99	0.99	0.99
Decision tree	99.59	1.00	1.00	1.00
Random forest	98.85	0.99	0.99	0.99
Naïve Bayes	94.99	0.95	0.95	0.95
Support vector machine	99.58	1.00	1.00	1.00

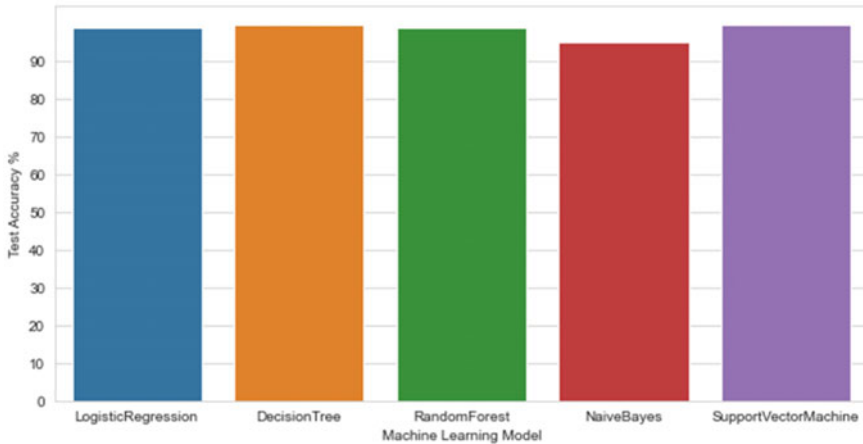


Fig. 10 Plot for comparison of all the classifiers used

better as compared to other supervised ML algorithm. When compared, NB is a generative model while DT is discriminative model. When compared, SVM solves nonlinear issues using the kernels method, whereas decision trees handle the problem by deriving hyper-rectangles in input space.

Hence in our study, decision tree is performing best with an accuracy of 99.58%.

5 Conclusion and Future Scope

Fake news has a huge influence on our social life, as well as in other domains, such as politics and education. Fake news may create significant social and societal harm, as well as have potentially dangerous consequences. It is becoming more difficult for the citizens/consumers to obtain the information that is precise and error free and reliable because of increasing the dimensions of social media. It's critical to discover such false information early on in order to avoid the global harm it can do. As a result, in this paper, we designed a methodology for detecting false news that combines NLP

techniques with supervised learning classification algorithm. In this work, we have presented a machine learning approach using various machine learning classifiers to detect fake news. After comparing the performance of each model, the conclusion can be drawn that decision tree outperforms the other algorithms being used, i.e., with the accuracy of 99.59%, and secondly, SVM performs well with the accuracy of 99.68%. This approach would be helpful to identify fake news effectively and with higher accuracy in future.

Future work could include comparing multiple deep learning approaches and new ensemble learning methods to the classification techniques used in this study and determining the best strategy for detecting fake news. Also, we may integrate a larger dataset from different sources like various URLs and news publication sites as it would be having bigger journalese and could be used for obtaining better results in a generalized manner.

References

1. Granik M, Mesyura V (2017) Fake news detection using naive Bayes classifier. In: 2017 IEEE 1st Ukraine conference on electrical and computer engineering UKRCON, pp 900–903. <https://doi.org/10.1109/UKRCON.2017.8100379>
2. Ahmed H, Traore I, Saad S (2017) Detection of online fake news using N-Gram analysis and machine learning techniques. In: Traore I, Woungang I, Awad A (eds) Intelligent, secure, and dependable systems in distributed and cloud environments, pp 127–138. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-69155-8_9
3. Campan A, Cuzzocrea A, Truta TM (2017) Fighting fake news spread in online social networks: actual trends and future research directions. Presented at the December 1. <https://doi.org/10.1109/BigData.2017.8258484>
4. Pérez-Rosas V, Kleinberg B, Lefevre A, Mihalcea R (2018) Automatic detection of fake news. COLING 2018 - 27th Int. Conf. Comput. Linguist. Proc. 3391–3401 (2018). <https://doi.org/10.48550/arXiv.1708.07104>
5. Aphiwongsophon S, Chongstitvatana P (2018) Detecting fake news with machine learning method. In: 2018 15th International conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON), pp 528–531. IEEE, Chiang Rai, Thailand. <https://doi.org/10.1109/ECTICon.2018.8620051>.
6. Gahirwal M (2008) Fake News Detection 3
7. Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. Phys A Stat Mech Appl 540:123174. <https://doi.org/10.1016/j.physa.2019.123174>
8. Agarwal V, Sultana HP, Malhotra S, Sarkar A (2019) Analysis of classifiers for fake news detection. Proc Comput Sci 165:377–383. <https://doi.org/10.1016/j.procs.2020.01.035>
9. Reis JCS, Correia A, Murai F, Veloso A, Benevenuto F (2019) Supervised learning for fake news detection. IEEE Intell Syst 34:76–81. <https://doi.org/10.1109/MIS.2019.2899143>
10. Zhang X, Ghorbani AA (2020) An overview of online fake news: characterization, detection, and discussion. Inf Process Manage 57:102025. <https://doi.org/10.1016/j.ipm.2019.03.004>
11. Shaikh J, Patil R (2020) Fake news detection using machine learning. In: 2020 IEEE international symposium on sustainable energy, signal processing and cyber security (IS-SSC), pp 1–5. IEEE, Gunupur Odisha, India (2020). <https://doi.org/10.1109/ISSC50941.2020.9358890>

12. Smitha N, Bharath R (2020) Performance comparison of machine learning classifiers for fake news detection. In: 2020 Second international conference on inventive research in computing applications (ICIRCA), pp 696–700. IEEE, Coimbatore, India (2020). <https://doi.org/10.1109/ICIRCA48905.2020.9183072>
13. Kesarwani A, Chauhan SS, Nair AR (2020) Fake news detection on social media using K-nearest neighbor classifier. In: 2020 International conference on advances in computing and communication engineering (ICACCE), pp 1–4. IEEE, Las Vegas, NV, USA. <https://doi.org/10.1109/ICACCE49060.2020.9154997>
14. Khanam Z, Alwasel BN, Sirafi H, Rashid M (2021) Fake news detection using machine learning approaches. In: IOP conference series: materials science and engineering, vol 1099, p 012040. <https://doi.org/10.1088/1757-899X/1099/1/012040>
15. Nagaraja A, KN S, Sinha A, Rajendra Kumar JV, Nayak P (2021) Fake news detection using machine learning methods. In: International conference on data science, E-learning and information systems 2021, pp 185–192. ACM, Ma'an Jordan. <https://doi.org/10.1145/3460620.3460753>
16. <https://www.kaggle.com/clmentbisailon/fake-and-real-news-dataset>

The Effect of E-Commerce on Increasing Consumption Behavior During a Pandemic



Ford Lumban Gaol, Sherina Gloria Tomponu, Jessica Adine Larissa, Arista Wijaksana, Ferdianto Prajnowira, and Tokuro Matsuo

Abstract The rapid development of technology and ease of information today greatly affects all aspects of life. During the COVID-19 pandemic, technology played an important role in maintaining community activities continued smoothly despite being hampered by several health policies during the COVID-19 pandemic. Many phenomena develop with new technologies in several fields, one in the economic field is e-commerce. E-commerce is a site that makes it easier for buyers to get the goods desired by the community, as well as the various interesting promos that are the attractiveness of the community to constantly shop in e-commerce so that the community applies consumptive behavior. So that the problem that will be discussed in this study is the development of e-commerce technology that affects people's behavior in shopping which tends to be consumptive and hedonic. The purpose of this study is to determine the impact or effect of the use of e-commerce on consumptive behavior during the pandemic period and provide an overview of the potential regarding e-commerce in the future. This research method is quantitative. The data collection technique used for this study is to use a questionnaire. Thus, this research is expected to minimize the negative impacts that can be caused by the increasing consumptive behavior of society during the pandemic and provide wise e-commerce use. The

F. L. Gaol (✉)

Binus Graduate Program—Doctor of Computer Science, Jakarta, Indonesia

e-mail: fgaol@binus.edu

S. G. Tomponu · J. A. Larissa · A. Wijaksana · F. Prajnowira

School of Information System, Bina Nusantara University, Alam Sutera, Indonesia

e-mail: sherina.tomponu@binus.ac.id

J. A. Larissa

e-mail: jessica.larissa@binus.ac.id

A. Wijaksana

e-mail: arista.wijaksana@binus.ac.id

F. Prajnowira

e-mail: ferdianto.prajnowira@binus.ac.id

T. Matsuo

Advanced Institute of Industrial Technology, Tokyo, Japan

e-mail: matsuo@aait.ac.jp

results of this study can be seen from calculations using a formula that respondents tend to be more extravagant during the pandemic and for future potential, respondents will certainly enjoy the convenience provided by e-commerce.

Keywords E-commerce · COVID-19 pandemic · Consumptive behavior · Increasing · Impact

1 Introduction

E-commerce is a marketplace that can bring together the seller and buyers online without meeting directly. Now, many producers or resellers have promoted their products through this marketplace. The number of sellers in e-commerce is certainly a factor. This happens because of the many consumers who now tend to choose to shop at e-commerce. Consumers can easily find items they want and make transactions in this e-commerce. E-commerce provides a variety of interesting promos from the seller [1]. That is certainly able to increase customer loyalty to continue shopping on the site.

During the pandemic COVID-19 condition in 2021, the use of e-commerce continues to increase in Indonesia. This is due to the increase of new users who constantly use e-commerce as a spending tool. Indonesia is a country that has the most Internet users who uses e-commerce. The Indonesian state noted that as many as 88.1% of Internet users in Indonesia use e-commerce to buy certain items in recent months [2]. Thus, the existence of e-commerce causes a different lifestyle in society. The lifestyle applied by the community tends to be consumptive and hedonic [3–5].

Consumptive behavior is an action as a consumer in getting, using, and making decisions in choosing something that has not become a necessity and not a major virtue, just because you want to follow fashion, try new products, and gain social recognition with the dominance of emotional factors [6–9].

According to Suprana [10], a hedonic lifestyle is a lifestyle that directs its activities to find life pleasure, such as spending more time outside the home, more playing, happy with the hustle of the city, often buying expensive items just to fulfill the fun, and always want to be the center of attention [11].

The following are problem formulation that must be solved in this research:

1. How does people's consumptive behavior affect the economy during the pandemic?
2. Why has e-commerce become a major factor in the increase of consumptive behavior?
3. What are the effects of the increasing consumptive behavior during the pandemic?

There are several reasons for e-commerce to become a community driving factor to be a more consumptive and hedonic lifestyle, as follows

1.1 Short Term Perspective and Awareness

In the buying process, buyers tend not to consider the long term and immediately transact to get what they want. Shoppers love promotions and discounts on products they may not need. This makes buyers chase discounts provided by e-commerce even though the products offered are not an important need. On the other hand, public awareness and knowledge of consumptive behavior are still minimal and there is a need for education about the importance of long-term needs such as starting to invest.

1.2 Technology Adaptation

Technology is a means of solving the basic problems of every human civilization. Without the use of technology, this will cause many problems that cannot be solved properly and perfectly [12]. The current generation quickly adapts to technology, such as gadgets, apps, and more. Technology that continues to develop continuously will facilitate the quality and all aspects of life, where on the one hand, technological adaptation has a negative impact, one of which is in the field of spending which creates a great opportunity for people to become more consumptive. The adaptability of this technology makes it easier for buyers to conduct e-commerce transactions. In addition, purchase products at retail outlets and offline stores with the e-wallet application. The number of promotions offered by e-wallets also encourages many buying and selling transactions.

1.3 Consumption Satisfaction

In today's modern era, today's society is becoming increasingly consumptive, this is as expressed by Stearns [13] in *Consumerism in World History: the global Transformation of Desire*. We can see the change in society which is increasingly shifting toward a consumptive society. However, consumption is no longer just an activity to fulfill basic and functional needs as a human being.

Consumer satisfaction in conducting online shop transactions is indicated to have an effect on consumer trust which will ultimately affect consumer attitudes in making repeat purchases [11]. Someone who buys goods based on desire without prioritizing the usefulness and benefits of an item will only make that person a consumptive [14–17]. Consumptive behavior is a person's tendency to behave excessively in buying something irrationally and prioritize wants over needs. If consumptive behavior continues to occur, it will result in financial conditions becoming uncontrolled, besides that it will lead to wasteful actions and result in the accumulation of goods due to excessive or continuous purchases [11].

2 Research Technique

The research technique that we use is a questionnaire using Google Form. According to Sugiyono [18], a questionnaire is a data collection technique by which researchers provide a list of questions or written statements to be answered by respondents [19]. In this study, we distributed the questionnaires via online or on the Internet. The steps in this questionnaire technique are that first, we will look for random responders like people around our campus, home, or neighborhood who are willing to answer the questionnaire given. We also included a step-by-step guide to the community on how to answer the questionnaire when we are distributing the questionnaire.

Before distributing the questionnaires, we have determined the research model and research hypotheses that are useful as a reference in making questionnaire questions. This questionnaire survey aims to determine the impact or influence of the use of e-commerce on people's consumptive behavior during the pandemic. In this survey, it is also known that all respondents use e-commerce as a shopping medium, so the data generated is quite valid.

Table 1 is a list of questions as the research questions contained in the Google Form questionnaire:

3 Research Method

This research was conducted to see the effect of E-commerce on increasing people's consumptive behavior during the pandemic. This research uses a quantitative research method to collect information by comparing various sources of data that we got from the questionnaire.

According to Sunyoto [20], quantitative research is a number that is certain so that it can be assembled, easier to read and easier for researchers to make an understanding [21]. Then, according to Sugiyono [18], quantitative data is a research method based on positivistic (concrete data), research data in the form of numbers that will be measured using statistics as a calculation test tool, related to the problem being studied to produce a conclusion [22, 23].

Researchers used a Likert scale of the variables given to each answer to each question contained in the questionnaire. The scale technique used in this study is itemized rating scales using a Likert scale (1–4) is shown in Table 2.

Analysis of the sample data in this study was carried out using the Slovin formula model [24] to calculate the minimum sample size if the behavior of a population is not known with certainty. This formula was first introduced by Slovin in 1960.

$$n = \frac{n}{1 + Ne^2} \quad (3.1)$$

Description:

Table 1 List of questionnaire questions

No.	Statement	Suitability level			
		SD	D	A	SA
1	I immediately buy things that I find interesting				
2	If there is a buyer who offers me an item, then I will immediately buy the item				
3	I immediately buy the item I like without seeing the quality of the item				
4	I bought the product out of necessity				
5	Buying goods because of the promo				
6	I only buy expensive things				
7	I have to prepare a special budget for online shop shopping				
8	The product I bought because of the desire factor				
9	I consider carefully when I want to buy an item				
10	I bought an item just because it was trending				
11	I buy things just to get social recognition from other people				
12	I bought things unplanned				
13	I buy things because I am influenced by opinions or promotions from other people				
14	In my opinion, expensive goods affect the quality				

Table 2 Scale description

Code	Description
1	Strongly disagree
2	Disagree
3	Agree
4	Strongly agree

n Sample

N Population

e 95% precision value or sig. = 0.05

Analysis of data based on respondents' answers in the questionnaire with a Likert scale of 1–4 as the basic formula for calculating the assessment of respondents' interpretations in the form of percentages.

$$\text{Formula} = T \times Pn \tag{3.2}$$

Table 3 Percentage category

Scale	Percentage limit interval	Rating category
1	0–25	Strongly disagree
2	26–50	Disagree
3	51–75	Agree
4	76–100	Strongly agree

Description:

T Total number of respondents who chose

Pn Choice of Likert score numbers

Calculation score interpretation.

In order to get the interpretation results, it must first be known the highest score (*X*) and the lowest score (*Y*) for the assessment item with the following formula:

Y = Likert’s highest score *x* number of respondents

X = The lowest score Likert *x* number of respondents

The following is the interval formula (range of distance) and percent interpretation in order to find out the assessment using the method of finding the percent score interval (*I*).

$$I = \frac{100}{\text{Totalscore (Likert)}} \tag{3.3}$$

The percentage category rating is shown in Table 3.

The following is the final completion formula to get the percentage results from the questionnaires that have been distributed.

$$\text{Final Completion} = \frac{\text{Total Score}}{Y \times 100} \tag{3.4}$$

4 Research Result

The following is the research sample data with calculations using the Slovin formula model [24]

$$n = \frac{n}{1 + Ne^2}$$

$$n = \frac{40}{1 + 40(0.05)^2}$$

$$n = 36.36$$

So, based on the results of calculations using the Slovin formula model [24], the minimum number of samples is 36 respondents. The figures related to the results are shown in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15.

After distributing questionnaire, there were 40 respondents who filled out the questionnaires that had been distributed. Of the 40 respondents, 56.8% of respondents were aged 20–26 years, 35.1% of respondents were aged 1–19 years, and 8.1 respondents aged 27–39 years.

Statement 1: I immediately buy things that I find interesting

From the 40 respondents, there were one respondent who strongly disagree, three respondents who disagree, 17 respondents who agreed, and 19 respondents who strongly agreed.

Statement 2: If there is a buyer who offers me an item, then I will immediately buy the item.

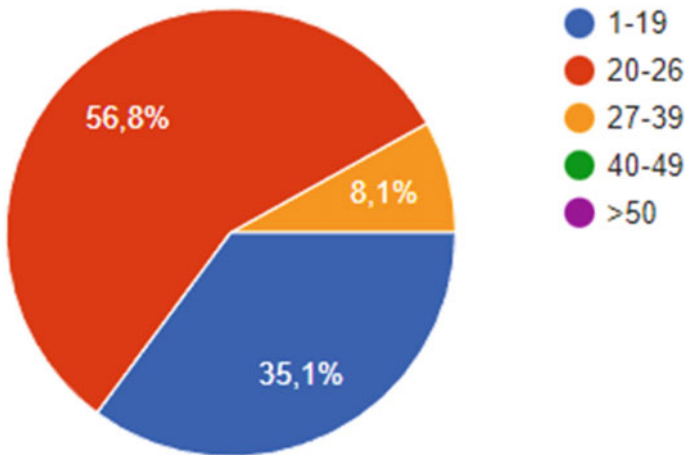


Fig. 1 Questionnaire result chart. Blue: age 1–19, red: age 20–26, yellow: age 27–39, green: age 40–49, purple: age > 50

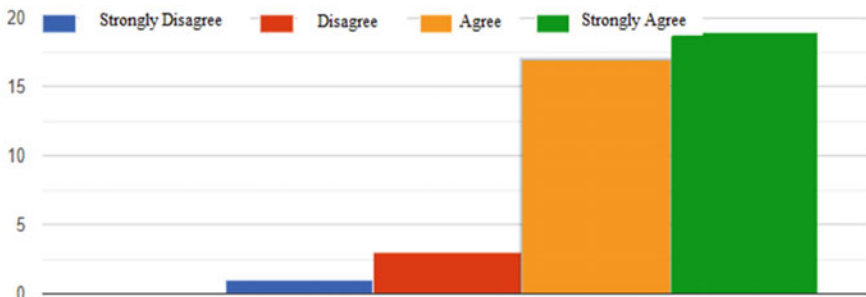


Fig. 2 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

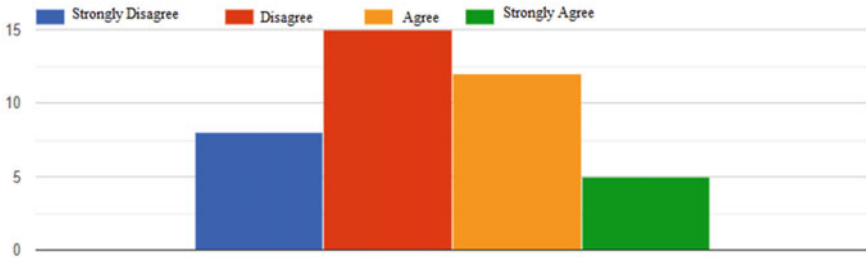


Fig. 3 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

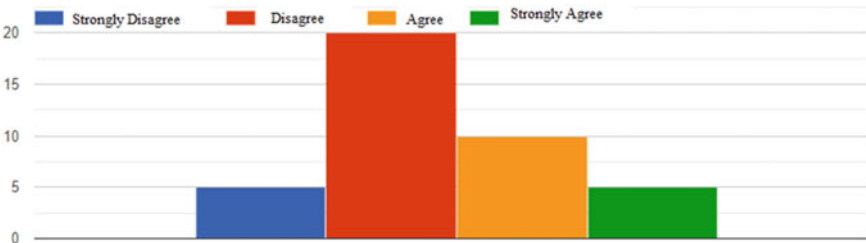


Fig. 4 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

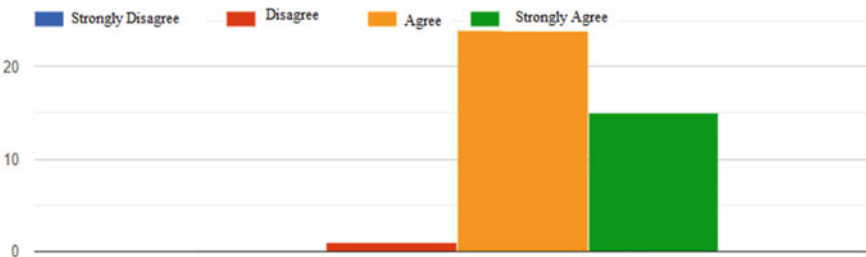


Fig. 5 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

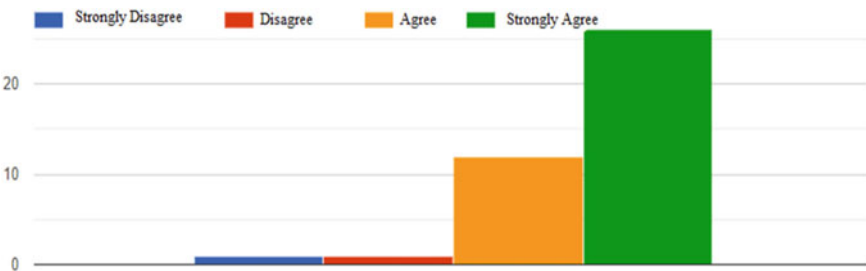


Fig. 6 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

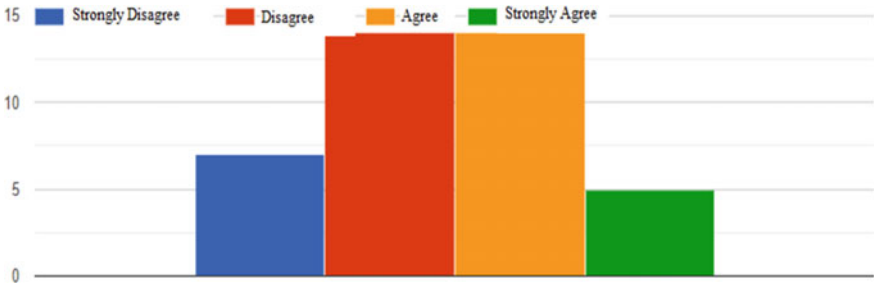


Fig. 7 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

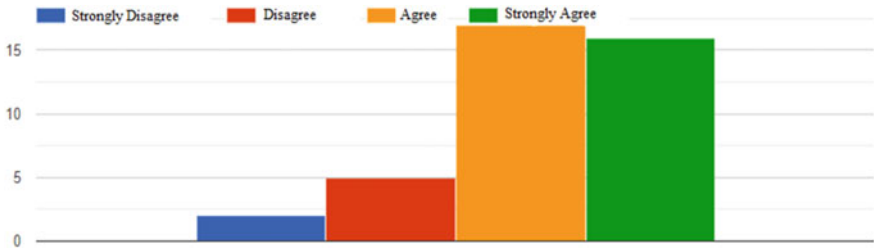


Fig. 8 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

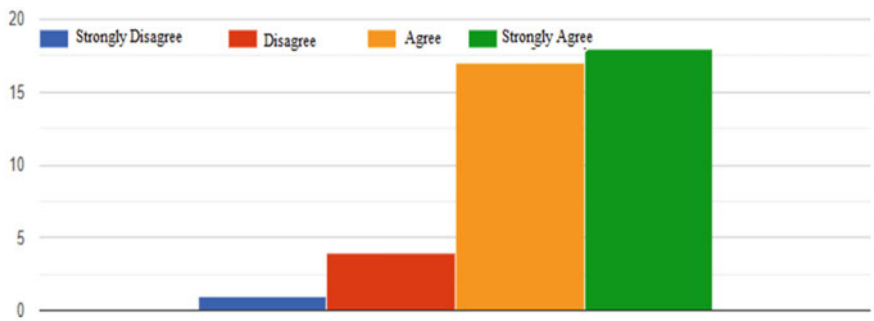


Fig. 9 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

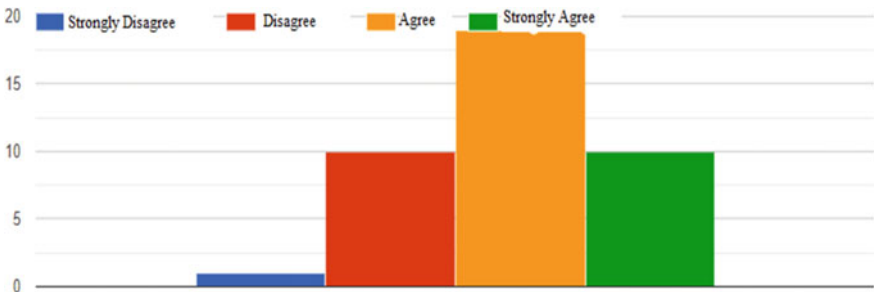


Fig. 10 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

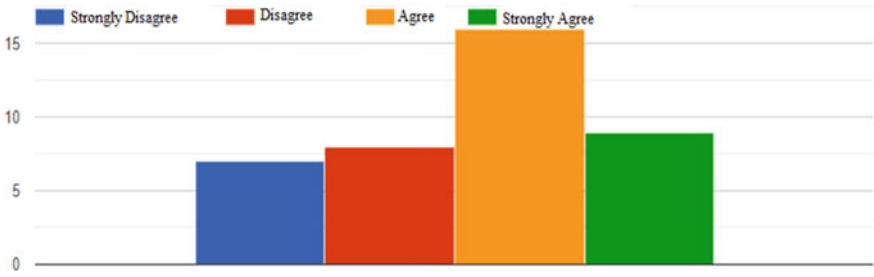


Fig. 11 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

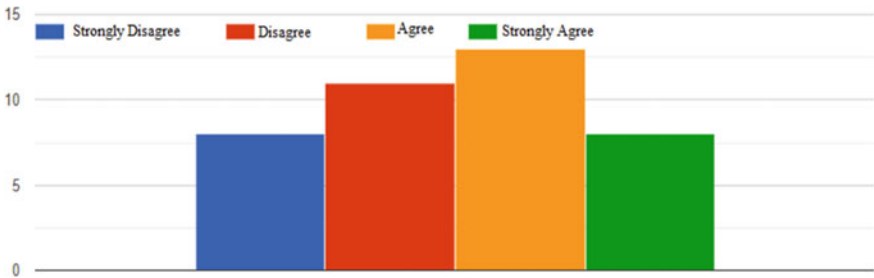


Fig. 12 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

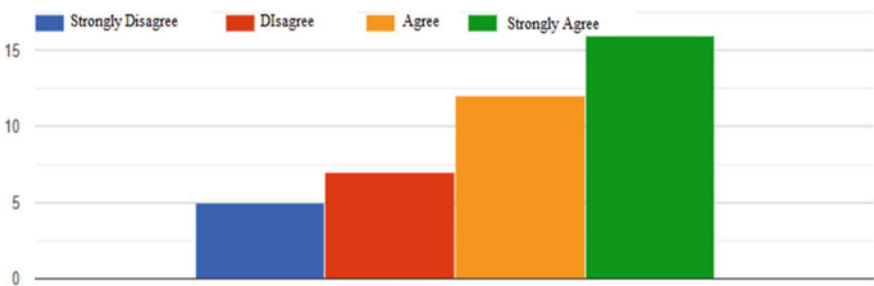


Fig. 13 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

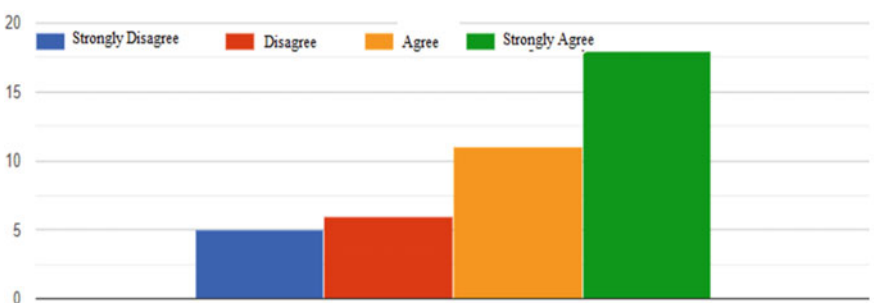


Fig. 14 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

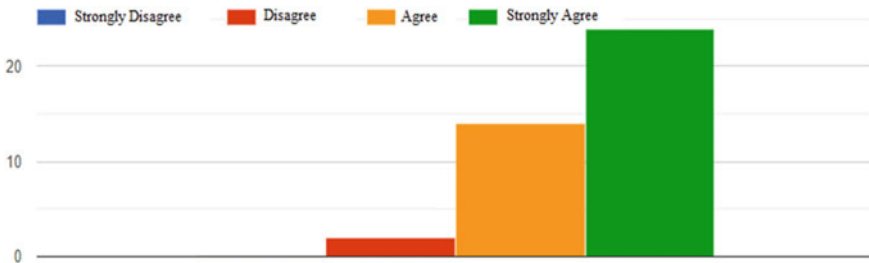


Fig. 15 Questionnaire result chart. X-axis: Likert scale, Y-axis: total respondent

From the 40 respondents, there were eight respondents who strongly disagree, 15 respondents who disagree, 12 respondents who agreed, and five respondents who strongly agreed.

Statement 3: I immediately buy the item I like without seeing the quality of the item.

From the 40 respondents, there were five respondents who strongly disagree, 20 respondents who disagree, ten respondents who agreed, and five respondents who strongly agreed.

Statement 4: I bought the product out of necessity.

From the 40 respondents, there were 0 respondent who strongly disagree, one respondent who disagree, 24 respondents who agreed, and 15 respondents who strongly agreed.

Statement 5: Buying goods because of the promo.

From the 40 respondents, there were one respondent who strongly disagree, one respondent who disagree, 12 respondents who agreed, and 26 respondents who strongly agreed.

Statement 6: I only buy expensive things.

From the 40 respondents, there were seven respondents who strongly disagree, 14 respondents who disagree, 14 respondents who agreed, and five respondents who strongly agreed.

Statement 7: I have to prepare a special budget for online shop shopping.

From the 40 respondents, there were two respondents who strongly disagree, five respondents who disagree, 17 respondents who agreed, and 16 respondents who strongly agreed.

Statement 8: The product I bought because of the desire factor.

From the 40 respondents, there were one respondent who strongly disagree, four respondents who disagree, 17 respondents who agreed, and 18 respondents who strongly agreed.

Statement 9: I consider carefully when I want to buy an item.

From the 40 respondents, there were one respondent who strongly disagree, ten respondents who disagree, 19 respondents who agreed, and ten respondents who strongly agreed.

Statement 10: I bought an item just because it was trending.

From the 40 respondents, there were seven respondents who strongly disagree, eight respondents who disagree, 16 respondents who agreed, and nine respondents who strongly agreed.

Statement 11: I buy things just to get social recognition from other people.

From the 40 respondents, there were eight respondents who strongly disagree, 11 respondents who disagree, 13 respondents who agreed, and eight respondents who strongly agreed.

Statement 12: I bought things unplanned.

From the 40 respondents, there were five respondents who strongly disagree, seven respondents who disagree, 12 respondents who agreed, and 16 respondents who strongly agreed.

Statement 13: I buy things because I am influenced by opinions or promotions from other people.

From the 40 respondents, there were five respondents who strongly disagree, six respondents who disagree, 11 respondents who agreed, and 18 respondents who strongly agreed.

Statement 14: In my opinion, expensive goods affect the quality.

From the 40 respondents, there were 0 respondent who strongly disagree, two respondents who disagree, 14 respondents who agreed, and 24 respondents who strongly agreed.

Questionnaire result list is shown in Table 4.

5 Discussion

The research reported here described the people's increasing consumptive behavior caused by e-commerce during the pandemic. The results we obtained through questionnaire and document research are in accordance with what has become the problem foundation. Most of our respondents show that they buy things that they find interesting and prepare a special budget for online shopping. Our research opens a new avenue of study regarding people's behavior during the pandemic, which in our case, it studies the consumptive behavior of others. If this topic is more analyzed, it can have further research on the state of the economy of the buyers and sellers during the pandemic. The response for every questionnaire answered is used to prove the truth of our research. It is also important to see the point of view from other people regarding our research topic or problems.

6 Conclusion

The e-commerce platform has contributed to the increase in people's consumptive behavior during the pandemic. With the pandemic situation going around forcing people to stay at their home, e-commerce becomes the solution of getting every

Table 4 Questionnaire result list

No.	Statement	Percentage	Category
1	I immediately buy things that I find interesting	83.75	Strongly agree
2	If there is a buyer who offers me an item, then I will immediately buy the item	58.75	Agree
3	I immediately buy the item I like without seeing the quality of the item	59.375	Agree
4	I bought the product out of necessity	83.75	Strongly agree
5	Buying goods because of the promo	89.375	Strongly agree
6	I only buy expensive things	60.625	Agree
7	I have to prepare a special budget for online shop shopping	79.375	Strongly agree
8	The product I bought because of the desire factor	82.5	Strongly agree
9	I consider carefully when I want to buy an item	73.75	Agree
10	I bought an item just because it was trending	66.875	Agree
11	I buy things just to get social recognition from other people	63.125	Agree
12	I bought things unplanned	74.375	Agree
13	I buy things because I am influenced by opinions or promotions from other people	76.25	Strongly agree
14	In my opinion, expensive goods affect the quality	88.75	Strongly agree

necessity. This situation has contributed to the increase of people's consumptive behavior. More and more people started shifting to online shopping. This is causing people to be more attracted to buy more stuff than they need to. Either it is caused by promos, discounts, or even boredom, more and more people started purchasing through e-commerce.

Consumptive behavior is one of the most detrimental things for most people, and for that consumptive nature has been studied a lot recently and it turns out that there are several ways to reduce consumptive nature. One way to reduce consumption is to use the needs priority method. This method is to compile a list of priority needs by making a priority list and then prioritizing what is needed more than just pleasure. So, with this method, if there is a desire to buy something other than a need, you can think longer, whether the item is on the priority list of needs or just a desire. However, this method has not been effective in reducing consumptive consumption, because it is very difficult to instill the thought that it is necessary to prioritize only the goods that are needed.

If we keep purchasing stuff online, we might come up with an addiction. It will ruin the buyer's economy, while giving sellers more income. More research is needed to come up with a solution to reduce the people's consumptive behavior to a normal or minimum level while still maintaining the economy for both buyer and seller.

Acknowledgements The research is funded by the Research and Technology Transfer Office of Bina Nusantara University. The success of this research paper is not possible without the help of various individuals. The researchers wish to give gratitude to all these people who was involved.

References

1. Ainy ZN (2020) Pengaruh E-Commerce Terhadap Perilaku Konsumtif Masyarakat Di Kelurahan Karang Panjang Kota Ambon 4:2–5
2. Mading R (2021) Ecommerce, Solusi Mempertahankan Ekonomi di Tengah Pandemi? <https://mading.id/perspektif/ecommerce-solusi-mempertahankan-ekonomi-di-tengah-pandemi/#:~:text=Terdapat%2088%2C1%20persen%20pengguna%20internet%20berusia%2016-64%20tahun,2020%2C%20nilai%20transaksi%20di%20ecommerce%20mencapai%20Rp266%2C3%20triliun>
3. Rachmawati AL (2019) Analisis Pengaruh E-commerce Terhadap Perilaku Konsumtif Mahasiswa (Studi Kasus Pada Mahasiswa Di Prodi Manajemen Universitas Tidar) 1:1–4
4. Mufarida HA (tahun) E-commerce consumer behavior among adolescents urban (Studies consumptive lifestyle and culture among youth perspective Surabaya city of cultural studies), pp 1–10
5. Hermawan J, Wafa MA, Shadiqin S (2020) Dampak Media Online Shopee Terhadap Perilaku Konsumtif Karyawan PT.Bank Pembangunan Daerah Jawa Barat dan Banten TBK Cabang Banjarmasin, pp 1–7
6. Dinisari MC (2020) E-commerce Dorong Perekonomian Indonesia, selama Pandemi Covid-19. <https://ekonomi.bisnis.com/read/20200417/12/1228750/e-commerce-dorong-perekonomian-indonesia-selama-pandemi-covid-19>, 24 Januari 2022
7. Authors Guide (2019) Apa yang Menyebabkan Perilaku Konsumtif di Indonesia? <https://www.sirclo.com/blog/apa-yang-menyebabkan-perilaku-konsumtif-di-indonesia/>. 3 Februari 2022
8. Dina Anavia P (2021) Pengaruh Intellectual Capital dan Islamic Corporate Governance Terhadap Kinerja Keuangan Bank Umum Syariah di Indonesia Tahun 2017–2019. Skripsi Tesis. Sekolah Tinggi Ilmu Ekonomi Indonesia
9. Rasyid NM, Aziz SA, Tengah RY (2019) Development of international motif inventory motives. *Mov Health Exerc J* 8(2):57–66
10. Suprana J (2003) Naskah-Naskah Kompas Jaya Suprana. Elex Media Komputindo, Jakarta
11. Hafsyah AH (2020) Pengaruh Kepuasan Konsumen, Perilaku Konsumtif, dan Gaya Hidup Hedonis Terhadap Transaksi Online (E-commerce)
12. Authors Guide (2017) Butuh Teknologi Pintar, Perlu Cerdas Berteknologi. <http://ditjenppi.menlhk.go.id/kcpi/index.php/inovasi/408-butuh-teknologi-pintar,-perlu-cerdas-berteknologi>. 3 Februari 2022
13. Stearns PN (2003) *Anxious parents a history of modern childrearing in America*. NYU Press
14. Kuncoro AA (2021) Pengertian Teknologi Menurut Ahli. <http://teknik-informatika-s1.stekom.ac.id/informasi/baca/Pengertian-Teknologi-Menurut-Para-Ahli/a11e499ed0f91399988fc7b98c460cdb2769d0bb>, 3 Februari 2022
15. Hidayat A (2017) Cara Hitung Rumus Slovin Besar Sampel. <https://www.statistikian.com/2017/12/hitung-rumus-slovin-sampel.html>. 3 Februari 2022
16. Khaira AA (2021) Pengertian Skala Likert dan Contoh Cara Hitung Kuesionernya. <https://www.diedit.com/skala-likert/>. 3 Februari 2022
17. Tridestia S (2020) Gambaran Variasi Menu dan Tingkat Kepuasan Konsumen di Katering Taman Sari Sakato Kota Pekanbaru. Diploma Thesis. Poltekkes Kemenkes Riau, Riau
18. Sugiyono (2018) *Metode Penelitian Kombinasi (Mixed Methods)*. CV Alfabeta, Bandung
19. Mingseli (2021) 9 Pengertian Kuesioner Menurut Para Ahli. <https://www.mingseli.id/2021/01/pengertian-kuesioner-menurut-para-ahli.html>

20. Sunyoto D (2016) Metodologi Penelitian Akuntansi. PT Refika Aditama Anggota Ikapi, Bandung
21. Thabrani G (2021) Metode Penelitian Kuantitatif: Pengertian, Karakteristik & Jenis. <https://serupa.id/metode-penelitian-kuantitatif-pengertian-karakteristik-jenis/>. 3 Februari 2022
22. Salma (2021) Penelitian Deskriptif: Pengertian, Kriteria, Metode, dan Contoh. <https://penerbitdeepublish.com/penelitian-deskriptif/>. 23 Januari 2022
23. Kurniawati M (2020) Ancaman Perilaku Konsumtif di Tengah Pandemi Corona. <https://money.kompas.com/read/2020/04/02/191400326/ancaman-perilaku-konsumtif-di-tengah-pandemi-corona?page=all>. 24 Januari 2022
24. Slovin MJ (1960) Sampling. Simon and Schuster Inc., New York

Network Device Identity Management Using Cryptography



Rohith Kapudasu and Kurunandan Jain

Abstract With the increase in network devices daily, the need for identity management arises. This can be achieved through cryptography. The proposed framework uses cryptographically generated dynamic keys and functions to secure network devices. A unique framework with an algorithm has been proposed that can mitigate the majority of current identity-based threats. This framework has been compared to the previous schemes, and security, and performance analysis has been calculated.

Keyword Dynamic key · Light weight authentication · Cryptography · Network device

1 Introduction

Network device identification is important to confirm whether the one trying to access the device's resources or connecting to it is genuine or a fraud. As there is an increase in the number of network devices, it needs a lot of security in order to maintain them. This project aims to prevent identity-based attacks on network devices by implementing a better algorithm to identify a user trying to access a network device. Cryptography can be used in order to secure such network devices from fraud. When using authentication based on cryptography, an attacker listening to the network should not gain any information that would enable them to falsely claim the user's identity. A network device in the network must be resilient against any identity-based attacks. This project proposes a methodology of lightweight password-less authentication which uses dynamic key generation and is resilient against multiple identity-based attacks. Performance and security analysis that is done using simulation suggests that this proposed scheme is resilient to most attacks present.

R. Kapudasu (✉) · K. Jain
Amrita Vishwa Vidyapeetam, Kollam, Kerala 690525, India
e-mail: rohith.kapudasu@gmail.com

K. Jain
e-mail: kurunandanj@am.amrita.edu

The contributions of this project are

- Network device registration and authentication
- Prediction and mitigation of identity-based attacks using a randomized dynamic key and secret generation.

2 Related Work

Adeel et al. [1] proposed a multi-attack resilient, lightweight IoT authentication scheme that was resilient against multiple attacks such as MITM attacks, replay attacks, session key disclosure, forgery attack, and impersonation attack. They provided mutual authentication between the client and the server, assuming two devices are trying to identify each other as valid. This scheme is said to be able to work in a real-world scenario where security is required at the most such as in healthcare sectors, and they are further working on industrial IoT test cases. Faria and Cheriton [2] proposed detecting identity-based attacks in wireless networks using signal prints. They used signal prints, i.e., signal strength characterization of packets being transmitted, and by matching them, they were able to know that attacker was using random MAC addresses or forging them. The authors of this paper showed that even if the attackers are moving by forging MAC addresses, they can be identified by the signal prints, which are strongly correlated to the physical locations they are in. Even though signal prints are strong, they can be defeated when an attacker uses multiple synchronized direction antennas. Trounge et al. [3] proposed a scheme for improvement of the more efficient and secure ID-based remote mutual authentication with a key agreement scheme for mobile devices on ECC, which overcame the weakness of the previous scheme as it had a flaw of storing random values in the database and was prone to leak. But this scheme is prone to known session-specific attacks as the session key or ephemeral state information might get leaked, and also this scheme is not user-friendly as a user needs to remember their identity ID and the authentication key generated by the server.

Wang et al. [4] proposed a better scheme to Trounge et al. scheme and pointed out the various attacks their scheme is vulnerable to. They were able to provide better efficiency compared to the previous schemes and were able to close some security loopholes. Neuman and Theodore's [5] paper on the Kerberos authentication scheme for identity management tells us how mutual authentication takes place between the server and the user. Kerberos is a distributed authentication service that allows a process on behalf of a client to verify its identity to the server without sending data across the network, which might be eavesdropped on by the attacker trying to impersonate the user. Kerberos is most useful in networks that are in doubt of being monitored by the attacker, although it has its limitations [6] as it is weak against password guessing and replay attacks. We can improve this by using a challenge-response protocol. Raghu et al. proposed a lightweight authentication scheme for 5G mobile communications [5] in which they have implemented a dynamic adaptive key management system for validating nodes, users, devices, and servers. This

provides an increased level of security in the authentication. This random key generation is used for key randomization purposes in our project. Jiang et al. [19] designed a mutual authentication mechanism for wireless sensor networks using the asymmetric Rabin cryptosystem (WSN). Li et al. [7] designed a lightweight mutual authentication scheme for smart city applications using public-key cryptography. Because the majority of IoT devices are wireless, the aforementioned techniques are not practicable for many IoT networks. Limited in terms of processing power, memory, and storage, researchers are also using a different approach when it comes to creating designs. Lightweight parameters, such as X-OR and hash functions, are used in the authentication protocols [8–14].

3 Methodology

The proposed framework works by using a dynamic key generation model as Fig. 1 [6] in which multiple random keys are produced. These dynamic keys are used in order to increase randomness and to add more security. This is done while the authentication server wants to send secret seed pairs. The two functions, *A* and *B*, will be performing logical and bitwise operations that will generate unique random secrets, which are passed as input to the hashing algorithm. The SHA256 hashing algorithm generates a hash that is 256 bits long. This forms the dynamic key from the secret and salt provided as input.

This methodology shows the implementation of a new identity management algorithm which uses dynamic key generation algorithm, and the protocol was checked using AVIPSA tool.

Below are the symbols and their details:

- X* Network device
- AS Authentication server
- SN Secret number
- IDX Identity of device *X*
- AIDX Alias ID of *X*
- MK Master key

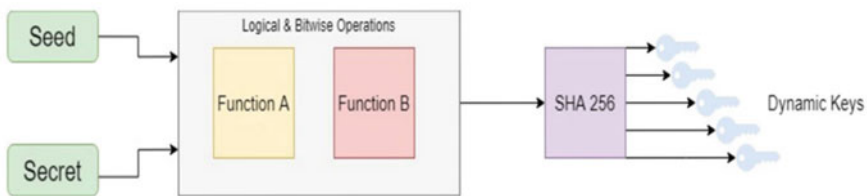


Fig. 1 Dynamic key generation model [6]

MIDX Master identity of X
 SKAS_X Secret key between AS and X
 || Concatenation

In the proposed framework, it has two phases which include

- Registration phase protocol
- Authentication phase protocol.

Dynamic keys are usually the most crucial cryptography keys [6] that form a sequence and are helpful in data encryption. This scheme provides keys in a sequential order originating from the initial parameters.

3.1 Registration Protocol

The sequential order of dynamic keys is mathematical and is presented as

$$n \in N, n > 1, (DK_i) = (DK_1, DK_2 \dots DK_n).$$

Each section of the cryptographic key encrypts only one message.

The generation of this involves:

- The authentication server does generation of distribution key, and it uses the exchange protocol to send it to the receiver
- The receiver together with the server is set to calculate the key of preference.

Device registers with the authentication server using its unique ID. Device ID should be unique. Device X sends its identity (IDX) to the AS.

The AS creates the alias identity (AIDX) for device X . The AS creates X 's master identity (MIDX) using the MK1. AS generates the secret key (SKAS_X) for X using the MIXD and the MK2.

$$AIDX = H(IDX \oplus SN) \tag{1}$$

$$MIDX = AIDX \oplus MK1 \tag{2}$$

$$SKAS_X = MIXD \oplus MK2 \tag{3}$$

Registration phase is shown in Fig. 2.

The AS computes the message digest $H(MIDX)$ in the end.

Authentication server sends AIDA, SKAS_X, $H(MIDX)$ to the device X .

$$\text{Hash1} = H(MIDX) \tag{4}$$

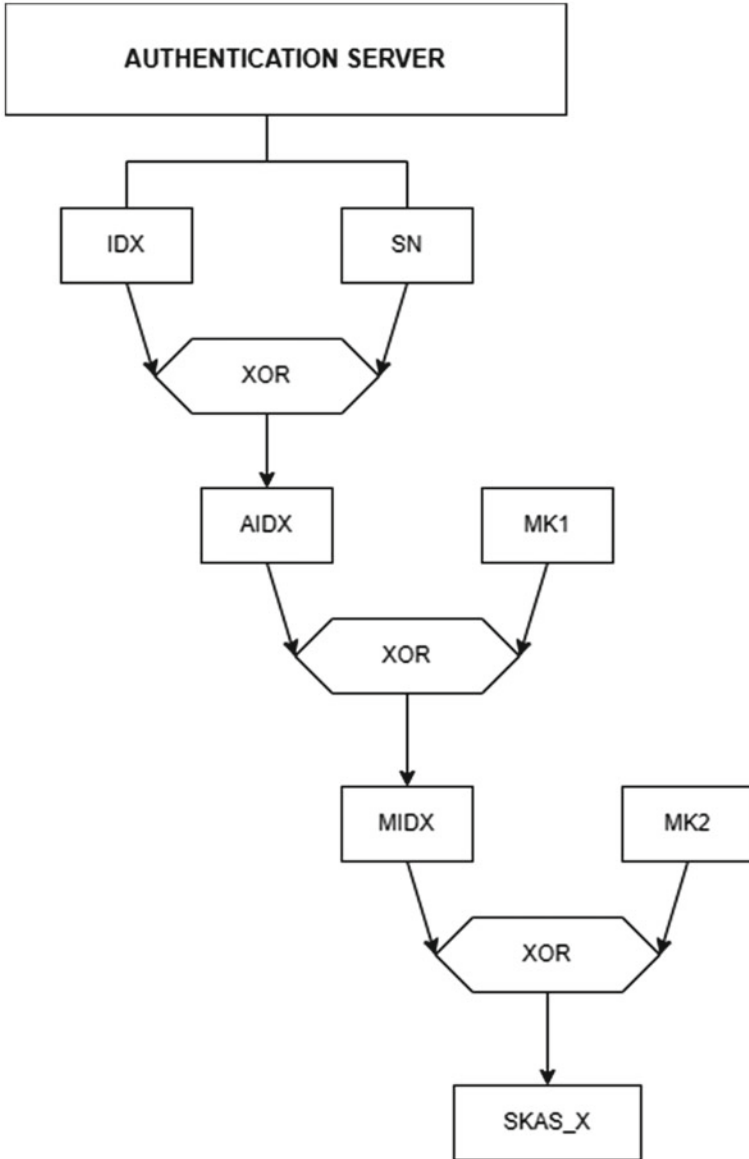


Fig. 2 Registration phase

3.2 Authentication Phase

The device calculates 2 secrets $S1, S2$ accordingly

$$S1 = SKAS_X \oplus R1 \tag{5}$$

$$S2 = H(MIDX) \oplus IDX \tag{6}$$

And finally device calculates its hash

$$\text{Hash2} = H(\text{AIDX} \parallel R1 \parallel S1 \parallel \text{IDX}) \tag{7}$$

and sends secrets $S1, S2$ along with Hash2 to the authentication server AS

Authentication phase is shown in Fig. 3.

Once the AS receives the secrets and message digest, it computes the digest Hash2' and compares it with Hash2 received from X. If both digests turn out to be equal, then the AS considers the secrets ($S1, S2$) authentic. If not, the authentication process is terminated when AS discards the ciphers.

$$\text{MIDX} = \text{AIDX} \oplus \text{MK1} \tag{8}$$

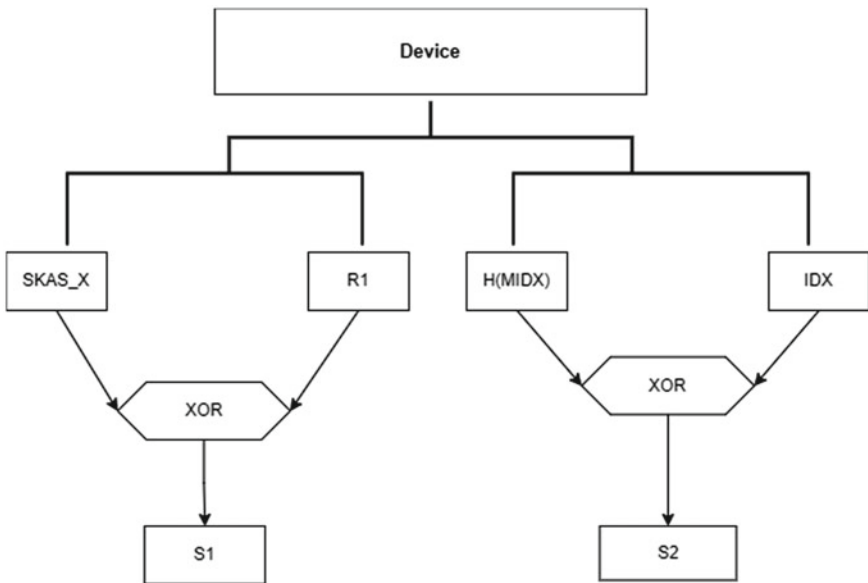


Fig. 3 Authentication phase

$$SKAS_x = MIDX \oplus MK2 \quad (9)$$

$$R1 = S1 \oplus SKAS_x \quad (10)$$

$$IDX = H(MIDX) \oplus S2 \quad (11)$$

$$\text{Hash2}' = H(\text{AIDX} || R1 || S1 || \text{IDX}) \quad (12)$$

Figure 4 shows the verification phase.

4 Security Analysis

Various attacks are deliberately malicious, and some vendors were performed mainly on dynamic key discussion stages, their aim being to explain the basis of operation of the mechanism in operation of the dynamic key, which includes an exhibition of security resilience against some attacks to disclose those keys in the mechanism. The attacks include:

- **Choosing of specific plaintext attack mechanism**

This is an attack which is in the form of crypto-analysis, which introduces the choice of a certain plain text to the encryption and decryption oracle field for the purpose of obtaining a certain cipher text that corresponds to it. This has a big consequence as it can lead to unallowed and unauthorized access of data stored and sensitive as there is open disclosure of symmetric keys.

- **Bruce forcing**

This attack is an exhaust key in methodology. Here, the adversary of the cipher is basically determining the proposed key by decryption of the cipher text by using the available space of the key.

The attacker using the brutal force is aware of the ciphertext and algorithm hence trying all the $2N$ keys, each one after the other, for decryption. The adversary of the key also utilizes a sample of known and establishable keys, and this helps him/her to increase finding the valid key through guessing. This, in return, is the cause and determinant of how strong is the cryptographic key delivered by the function of the algorithm. The whole role of this attack basically lies in deciphering the text.

- **Attack on preimage**

This is totally different from the previous two ones as it is variant and focuses on testing the hash operation algorithm against the indent preferred preimage resistance. In this case, there are two variant attacks inclusive of the first preimage and the second one of that kind.

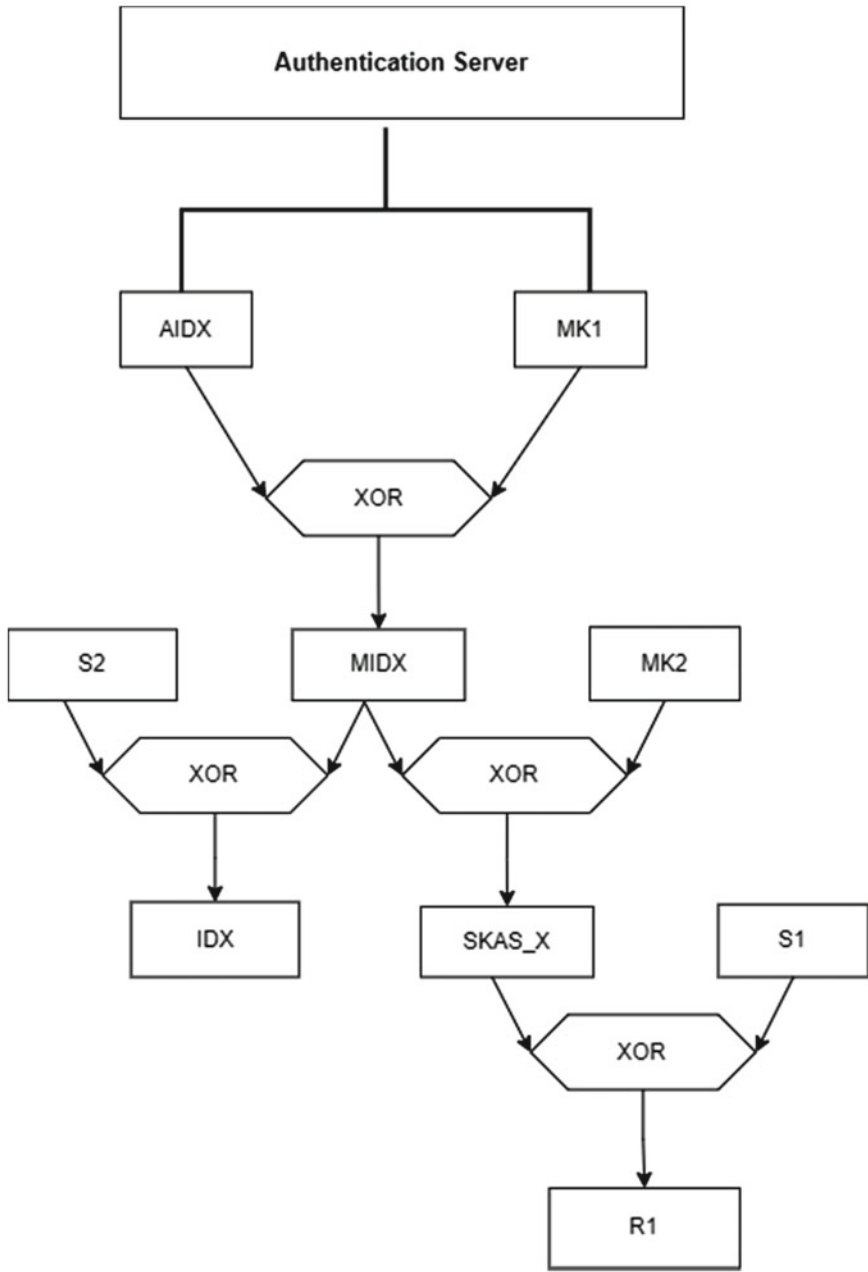


Fig. 4 Verification

Hash values help the adversary to pursue the text with less than $2N$, N standing for eligibility of the hash operation function. Its analysis resists, and essence, therefore, implies that for outputs that are not specified, it is extremely hard to check out any entry that hashes to the same display.

After the previous, there comes another attack of second preimage. The first message, $M1$, is known to the adversary, and this helps him/her to easily look for $M2$, which is obviously this second message keeping in mind that the first message is known. This, in turn, results in exactly computing hash figures for display while findings as $M1$ in less than $2N$ numbers as well.

- **Replay attack**

This attack is in the variance of man in the middle attack, i.e., MITM, where the adversary tends to interfere with the package on the communication channel and retransform a modified package on the wide area to yield the authentication mode display in nodes. It mainly constituents in areas of protocol authentication and reusability of authentication keys.

As a result, this attack, after being used by the adversaries, can lead to unauthorized access to network resources. It critically helps the adversary in the determination and synchronization of a certain scheme under investigation to know whether it can be utilized for the purposes of authentication in the topology of network design.

4.1 Security Analysis on AVISPA Tool

The suggested scheme's security is validated using the AVISPA tool. It is a tool used to know the correctness and security of any protocol. It assesses the security of protocols on the Internet and applications automatically. AVISPA specifies security protocols and their attributes using a high-level protocol specification language (HLPSL). The HLPSL is a rule-based language that defines the protocol's entities' roles. The environment part describes the intruder's understanding of the plan. The environment component helps analyze the level of the protocols' security. In the intruder knowledge section, multiple parameters can be defined to determine the point at which the protocol's security is breached. The purpose section outlines the protocol's security objectives. The on-the-fly-model checker (OFMC) may examine protocols for any known cyber-attack. An adversary is a component of the network in the Dolev-Yao(dy) [13] model. The attacker has the ability to listen in on all messages shared by legitimate network entities. Similarly, the attacker can alter communications and deliver them to genuine network entities. The suggested strategy is tested at the back ends of the AVISPA tool using the OFMC and the constraint logic-based attack searcher (CL-AtSe) and it shows to be safe.

4.2 Security of the Proposed Model

- **Resistance against impersonation attack**

Here, the network device is being registered before any transmission occurs between them with the device unique ID. So, this prevents impersonation attack, and the device unique ID cannot be spoofed as the attacker does not have the secret keys the device can use to authenticate itself to the server.

- **Replay attack resistance**

Replay attack cannot be possible because of the dynamic keys and dynamic secret randomization is done and only the random secret values can be known to the network device and the authenticating server. As only the real device knows the message digest hash the attacker will not be able to imitate as the network device. Attacker will not be able to replay.

- **Resistance against man in the middle attack**

Such attack is not possible because the secrets and keys are dynamically randomized, and only the parties communicating will know them. Even if the attacker gains access to the messages, he has no idea of how to decrypt them because the messages are not sent as plain text. The attacker will not be able to gain access because he does not know the random secret value.

- **Mutual authentication**

Here, mutual authentication is made sure because the network device needs to register itself, and only after sharing the required parameters does the server authenticates it. Similarly, the authentication server sends its parameters, and hence, it will be verified too.

- **Resistance against denial of service**

Assume a malicious device in the network tries to commence the authentication procedure with another device directly. The request is then canceled by the other valid device, which operates on the premise of communication over a verified channel.

5 Performance Analysis

In this section, we evaluate the proposed approach and compare it to other studies. The test configuration is illustrated in Fig. 5, where the code is written in Python 3.6. We used a Raspberry Pi 3 since it is often used in IoT and has all supported libraries loaded.

One Raspberry Pi board acts as the network device and another as the authentication server. We can see them being authenticated in Fig. 6.

Version	Tyoe
Instruction set	ARMvVv8-A (64/32-bit)
SoC	Broadcom BCM2837B0
FPU	VFPv4 + NEON
CPU	4x Cortex-A53 1.4 GHz
GPU	Broadcom VideoCore IV @ 250 MHz
Memory (SDRAM)	1GB

Fig. 5 Raspberry Pi configuration

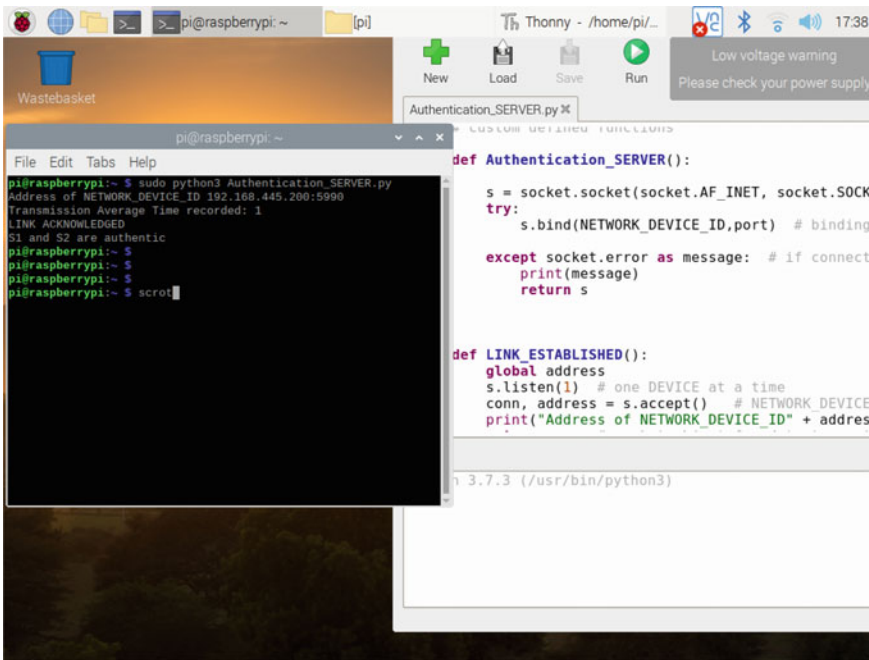


Fig. 6 Network devices being authenticated

5.1 Turnaround Time Comparison with Previous Schemes

The overall time required by the entities in protocol to complete the overall authentication process is referred to as turnaround time as shown in Fig. 7.

Scheme	Coding	Machine	OS	Time(ms)
Esfahani et al[8]	Java	Core i3	Windows 7	1441
Khemissa and Tandjaoui[9]	Java	Core i3	Windows 7	3855
Proposed scheme	Python	Core i3	Windows 7	1470

Fig. 7 Turnaround time comparison

Computation path	Messages	Computation overhead
A → AS	S ₁ , S ₂ , Hash ₁ , Hash ₂	4 * 128 = 512
AS → A	Verifier, Hash ₃	2 * 128 = 256
Total		768 bits

Fig. 8 Computation overhead

5.2 Computation Overhead

The computation overhead is the amount of time it takes entities to compute all of the messages sent throughout the authentication process as shown in Fig. 8.

5.3 Computation Time

Computation time has been compared with previous schemes as shown in Fig. 9.

Scheme	Coding	Machine	OS	Time(ms)
Esfahani et al[8]	Java	Core i3	Windows 7	1170
Khemissa and Tandjaoui[9]	Java	Core i3	Windows 7	2380
Proposed scheme	Python	Core i3	Windows 7	970

Fig. 9 Computation time comparison

6 Conclusion

Identity management is considered the main line of defense to mitigate any attacks. Therefore, this project aims to propose a new identity algorithm using dynamic keys, which has shown to mitigate many of the present identity-based attacks such as replay attack, forgery attack, session key disclosure, MITM attack, and impersonation attack. The framework's security analysis done using the AVISPA tool tells us that this framework can be deployed in a real-time environment and can further be studied.

References

1. Adeel A et al (2022) A multi-attack resilient lightweight IoT authentication scheme. *Trans Emerg Telecommun Technol* 33(3):e3676
2. Truong T-T, Tran M-T, Duong A-D (2012) Improvement of the more efficient and secure ID-based remote mutual authentication with key agreement scheme for mobile devices on ECC. In: 2012 26th international conference on advanced information networking and applications workshops. IEEE
3. Wang D et al (2016) On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices. *IEEE Syst J* 12(1):916–925
4. Neuman BC, Ts'o T (1994) Kerberos: an authentication service for computer networks. *IEEE Commun Mag* 32(9):33–38
5. Pothumarti R, Jain K, Krishnan P (2021) A lightweight authentication scheme for 5G mobile communications: a dynamic key approach. *J Ambient Intell Human Comput* 1–19
6. Faria DB, Cheriton DR (2006) Detecting identity-based attacks in wireless networks using signalprints. In: Proceedings of the 5th ACM workshop on wireless security
7. Shuai M et al (2020) A secure authentication scheme with forward secrecy for industrial internet of things using Rabin cryptosystem. *Comput Commun* 160:215–227
8. Esfahani A et al (2017) A lightweight authentication mechanism for M2M communications in industrial IoT environment. *IEEE Internet Things J* 6(1):288–296
9. Khemissa H, Tandjaoui D (2016) A novel lightweight authentication scheme for heterogeneous wireless sensor networks in the context of Internet of Things. In: 2016 wireless telecommunications symposium (WTS). IEEE
10. Hwang I, Young-Gab K (2017) Analysis of security standardization for the internet of things. In: 2017 international conference on platform technology and service (PlatCon). IEEE
11. Al-Fuqaha A et al (2015) Internet of things: a survey on enabling technologies, protocols, and applications. *IEEE Commun Surv Tutor* 17(4):2347–2376
12. Liu B et al (2017) Blockchain based data integrity service framework for IoT data. In: 2017 IEEE international conference on web services (ICWS). IEEE
13. Dolev D, Yao A (1983) On the security of public key protocols. *IEEE Trans Inf Theory* 29(2):198–208
14. Jiang Q, Zeadally S, Ma J, He D (2017) Lightweight three-factor authentication and key agreement protocol for internet-integrated wireless sensor networks. *IEEE Access* 5:3376–3392

Change in Outlook of Indian Industrial OEMs Towards IIoT Adoption During COVID-19



Prasanna V. Rao Deshpande and S. Kumar Chandar

Abstract Industrial Internet of Things (IIoT) is witnessing a steady increase in adoption by infrastructure and process industries. Industrial equipment manufacturers are one of the key stakeholders in this digitalization journey. The adoption of IIoT by the equipment manufacturers has been slower due to various valid reasons. The present pandemic COVID-19 created disruption in the factory operations in many parts of the world. This consequence has been hard on the manufacturing industry including the equipment manufacturers, and many of their strategic projects are slowing down or derailed. In India, a strict lockdown of three weeks which was later extended for another seven weeks was by far the longest lockdown effecting the industry and the equipment manufacturers. This study probes the impact of COVID-19 on the mindset of original equipment manufacturers (OEMs) towards adoption of IIoT.

Keywords IIoT · OEM · COVID-19 · Adoption

1 Introduction

Industry 4.0, the ongoing revolution on digitalization in manufacturing industry is impacting the product's design, production and after sales service. Many global infrastructure and process industries have started digital transformation in last few years. In India too, the pace of adoption has increased with large industrial end users, though the adoption by local Indian original equipment manufacturer (OEM) companies is yet to begin seriously. It is the IT services companies, many IT startups and traditional large automation companies who are entering this area. It is important that the Indian OEMs adopt digitalization considering the value derived from them and the progress being made by their counterparts—international OEMs. Three broad areas where OEMs can implement the digitalization include the business process, in

P. V. Rao Deshpande · S. Kumar Chandar (✉)
School of Business and Management, CHRIST (Deemed to Be University), Bangalore, India
e-mail: kumar.chandar@christuniversity.in

P. V. Rao Deshpande
e-mail: prasannadeshpande@res.christuniversity.in

the manufacturing and in the equipment which they manufacture. Digitalization in business process and manufacturing will yield productivity and quality improvements and the effective use of digitalization in the equipment can generate new business opportunity and create superior customer experience.

There are various technologies which are driving the digital transformation in the industries. These include big data and data analytics, Industrial Internet of Things (IIoT), cloud computing, cyber security, augmented reality, virtual reality, autonomous robots, additive manufacturing which includes 3D and 4D printing, simulation including digital twin, artificial intelligence, machine learning, system integration—horizontal and vertical. While most of these technologies can be used in the manufacturing units of OEMs, few of these could be predominantly used on the product they manufacture. IIoT is one of the technologies impacting.

There is a rapid growth in use of Internet of Things (IoT) in consumer products owing to availability of sensors at competitive prices, increased Internet usage and inclination of millennials and generation z for digital convenience. However, in industry, adoption of IIoT is yet to catch such momentum. Whereas there seems to be increase in pull from the end customers and as mentioned earlier the IT services companies are benefiting, the push from the OEMs is limited. OEMs have big stake to loose in the customer spend pie, if they do not bring additional meaningful services using IIoT.

2 Background of the Study

2.1 Equipment Ready for IIoT

The foundation of Industry 4.0 is connected device. OEMs play a vital role in ensuring that their equipment is connect ready. OEMs also have specific know-how of their equipment to determine performance and reliability related parameters. As the economy moves towards outcome based, OEMs profit stream will shift to more value-added services enabled through digitalization.

As a part of larger ongoing research, a preliminary study was done in 2019 regarding the barriers and enablers of Industrial Internet of Things (IIoT) by Indian OEMs in the equipment they produce. The OEMs in the study are industrial equipment manufacturers like pumps, fans, compressors, material handling equipment, and process equipment.

2.2 March 2020 COVID-19 Pandemic and Lockdown

On 24 March 2020 Government of India imposed a nationwide lockdown taking cue from many other major countries who were reeling under the rapid increase in cases

of COVID-19 pandemic. There was a strict lockdown for 3 weeks during which only very essential and critical services were permitted to operate. Considering the prevailing situation, painfully the lockdown was further extended twice with limited relaxation till first week of June 2020. During this complete period, movement of people was highly restricted. After 75 days of lockdown as the country opened in phased manner, there was still many checks and restrictions in people movement. This lockdown and post lockdown period effected the industries at large which included OEMs and their customers and partners.

The OEMs interviewed in 2019 were revisited post lockdown in India to understand their change in mindset. This paper covers the elements which have changed in OEM's perception about adoption of IIoT in the mid of this COVID-19 pandemic situation.

3 Literature Review

Within Industry 4.0, among various technology drivers of digitalization, physical components and biological digitalization would be fundamental and predominant driver [1]. In the last few years, IIoT has developed rapidly in many industrial fields and there has been good literature around IIoT but these have been mainly around technology or end user related [2]. Hardly 6% of the literature on IIoT is related to business management of which major part is end user oriented.

Awareness and attitude of industry players are topic of study for many researchers [3]. The awareness on some digital technologies by academia and industry are also studied [4–9].

4 Methodology

A questionnaire was designed to assess the awareness of various technologies in digitalization and Industry 4.0, know-how of specific elements in IIoT domain, customer journey, awareness of IIoT adoption by present competitors as well as global manufacturers of similar equipment.

Another set of questionnaires included benefits, barriers and enablers to the use of IIoT in their equipment, as well as their intent to adopt IIoT.

The questionnaire during 2019 which covered over 100 companies received complete responses from 52 OEMs, five consultants and 20 end users. The questions were administered to the senior and middle management of the firms through mainly face to face interactions and with few via telephonic interview. It was mentioned clearly that the response sought is of the firm and not individual level. The Likert scale '1–5' was used to rank their responses. The format also included possibility to include explanation for their responses. The respondents were encouraged to check with their relevant colleagues while responding to the questions. This was possible

due to acquittance of the interviewer with the respondents. Similar questionnaires were posed again to these 52 OEMs during July–August 2020 of which 33 OEMs responded to the requirement.

For analysis only these 33 OEMs who participated in both 2019 and 2020 have been considered.

Based on factor analysis of this study, following attributes have been compared using of the paired *t*-test.

- Awareness
- Enthusiasm to adopt
- End user pull

5 Definitions of the Variables

A. *Awareness*

Awareness in the study is the level of knowledge a firm possesses about digital technologies and specifically about IIoT technology. Awareness is measured in the study as it is the stimulant for further actions or inaction by the firm. This was measured by capturing inputs under following subheads. Various digital technologies can be applied to the equipment being sold and to service the product.

- IIoT technologies—sensors, protocols available, computing and analytical possibilities.
- Industry players in the technology.
- Adoption rate of IIoT in the relevant industry segment.
- Direct competitors and their moves on IIoT and digitalization.
- Best in class global equipment manufacturers and their advancement in using IIoT in the relevant applications.

B. *Enthusiasm to adopt*

Enthusiasm to adopt the IIoT by the firm is measured based on their perception of the following elements

- Impact on business performance which includes revenue and profit gains.
- Perceived risks and gains like warranty costs, product modification, integrating into existing system, hosting data securely and data analytics.
- Organizational capabilities including need for upskilling, integration of equipment data into process data—solutions including beyond equipment, handling different business models for services.

C. *End user pull*

End user pull is a market force which initiates adoption and tends to accelerate based on the pull. In this study, end user pull is measured by capturing OEM's perception of end user need, pains and gains through following elements.

- Benefits of IIoT ready equipment.
- Willingness to pay.
- Impact on equipment availability and OEE.
- Total cost of ownership (TCO).
- Data security and sensitivity.

6 Results and Discussions

A two-sample *t*-test was run using excel for each of the three variables assuming equal variances and comparing the data from 2019 and of 2020. 33 OEMs out of 52 OEMs who participated in 2019 (pre-COVID-19) responded in 2020 (during COVID-19), which was 63%. The response rate was good considering that key people interviewed in OEM organization are busy presently focusing heavily on the business and operations after completion of this lockdown period.

For all the three variables, the *p*-value is less than 0.01. This clearly signifies that there exists a difference in perception of all these attributes from pre-COVID-19 period to present COVID-19 period. The details of each of these variables are tabulated and discussed below.

A. *Awareness*

The mean awareness pre-COVID-19 was 2.5, and the awareness has increased substantially to a mean of 3.24. This is close to 30% increase as in Table 1. The main areas where the gain in awareness was in understanding technology applicable to the equipment, knowing about use cases and progress of global best in class companied in similar equipment range. The companies had utilized the lockdown period to understand on these aspects. Many of them had attended online seminars and trainings during this time.

Coincidentally the topic technology trends selected by OEMs during COVID-19 included cloud computing, on premise computing, virtual reality and augmented reality, IIoT and data analytics. Some OEMs have commented the importance of on premise considering the data security apprehension of many customers in allowing data to be sent to cloud. Technology trends identified by OEMs based on relevance is shown in Fig. 1.

B. *Enthusiasm to adopt*

The mean enthusiasm to adopt during pre-COVID-19 was 2.04, and the awareness has increased substantially to a mean of 2.67. This is close to 30% increase.

Table 1 Awareness two-sample-*t*-test pre-COVID-19—during COVID-19

	Awareness pre-COVID-19	Awareness during COVID-19
Mean	2.50	3.24
Variance	0.89	0.16
Observations	33	33
Pooled variance	0.125	–
Df	64	–
$P(T \ll t)$ two-tail	0.000	–
<i>t</i> critical two-tail	1.998	–

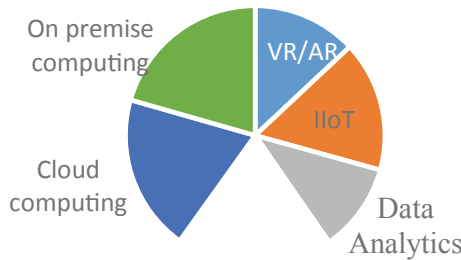


Fig. 1 Technology trends identified by OEMs based on relevance

The main areas where the gain was the expectation of increased revenue from existing and new markets, possibility to introduce new and more variant products, minimizing warranty costs and looking at Annual maintenance contracts (AMC). Other gain area was the possibility to improve the product performance as in Table 2.

Many companies scored lower on organizational capabilities during the COVID-19 interview and specifically on investment and resource allocation. This can be attributed to the current financial stress most of the companies are undergoing. However, interestingly eight of the companies who had not invested

Table 2 Enthusiasm 2-sample-*t*-test pre-COVID-19—during COVID-19

	Awareness pre-COVID-19	Awareness during COVID-19
Mean	2.04	2.67
Variance	0.498	0.358
Observations	33	33
Pooled variance	0.428	–
Df	64	–
$P(T \ll t)$ two-tail	0.000	–
<i>t</i> critical two-tail	1.998	–

Table 3 End user-pull 2-sample-*t*-test pre-COVID—during COVID-19

	Awareness pre-COVID-19	Awareness during COVID-19
Mean	2.93	3.2
Variance	0.161	0.191
Observations	33	33
Pooled variance	0.176	–
Df	64	–
$P(T \ll t)$ two-tail	0.009	–
<i>t</i> critical two-tail	1.998	–

or were low on investment pre-COVID-19 have scored high on investment. The feedback from them was that they wished to pilot study on priority to support their customers especially overseas clients.

C. *End user pull*

The mean of the perception of end user pull strangely has increased marginally to 3.2 from 2.93 in pre-COVID-19. The elements of remote maintenance, improved product availability, data driven decision and lower TCO were perceived to be stronger during COVID-19 interviews, however the perception around data security and willingness of end users to pay for the services was lower than the pre-COVID period as in Table 3.

7 Conclusion

The study gave a good insight of the change in mindset of OEMs to adoption of IIoT in their equipment by understanding their awareness, enthusiasm and capturing the perception of OEMs about their end user markets. Despite the numerous opportunities IIoT provides OEMs to serve their customers well, increase revenues and profitability, the adoption of IIoT by Indian OEMs is still low. There is, however, increased interest and change in mindset of the OEMs. This change in mindset along with improved ecosystem would accelerate the digital transformation. Change in mean value of variables studied is shown in Table 4.

Table 4 Change in mean values of variables studied

	Awareness pre-COVID-19	Awareness during COVID-19
Awareness	2.50	3.24
Enthusiasm to adopt	2.04	2.67
End user pull (OEM perception)	2.93	3.2

An interesting element from four OEMs to adopt IIoT was employee wellbeing. This surely reflects the increasing safety mindset of the OEM which was evident in the discussions.

8 Further Work

There is a need to further study the barriers and enablers for the OEMs to adopt IIoT. The aspect digital transformation is a big change for most of the companies and understanding the pains and gains of IIoT to OEMs is important. As the end user segment would itself have different pace of digital transformation, there is scope to study various categories of OEMs and catering to different industry segments. Lastly, this study can be repeated after the COVID-19 impact on the industry is normalized, to understand any further change in mindset.

Acknowledgements I wish to express my appreciation to the respondents who took their time to answer the interview questions, for patiently clarified the aspects and for insightful discussions. I am overwhelmed by the empathy shown by OEMs towards their employees especially those working in the field during these uncertain times as well as the enthusiasm they showed towards adopting the new technologies to give their customers better experience. I wish to also thank my research supervisor Dr S. Kumar Chandar who has always been a guide and support in spite of his busy schedules.

References

1. Guoping L, Yun H, Aizhi W (2017) Fourth industrial revolution: technological drivers, impacts and coping methods. *Sci Press* 626–637
2. Liao Y, Loures EFR, Deschamps F (2018) Industrial internet of things: a systematic literature review and insights. *IEEE Internet Things J* 4515–4525
3. Nwagwu W, Okuneye M (2016) Awareness and attitudes of small scale information technology business operators in Lagos, Nigeria toward e-waste hazards. *J Glob Inf Technol Manage* 267–282
4. Farayibi P, Abioye TE (2017) A study on the awareness level of additive manufacturing technology in south-western Nigeria. *Afr J Sci Technol Innov Dev* 157–162
5. Chen S, Xu H, Liu D, Hu B, Wang H (2014) A vision of IoT: applications, challenges, and opportunities with China perspective. *IEEE Internet Things J* 1(4):349–358
6. Lin J, Yu W, Zhang N, Yang X, Zhang H, Zhao W (2017) A survey on internet of things: architecture, enabling technologies, security and privacy, and applications. *IEEE Internet Things* 4(5):1125–1139
7. Peña-Ríos A, Hagrais H, Owusu G, Gardner M (2018) Furthering service 4.0 harnessing intelligent immersive environments and systems. *IEEE Syst Man Cybern Mag* 20–31
8. Biswas AR, Giaffreda R (2014) IoT and cloud convergence: opportunities and challenges. In: *IEEE World Forum on Internet of Things*
9. Siegel JE, Kumar S, Sarma SE (2018) The future internet of things: secure, efficient, and model-based. *IEEE Internet Things J* 5(4):2386–2398

Author Index

A

Abirami, K., 857
Abishek, D., 97
Aditya, Bima, 319
Ahmed, Towfik, 867
Aishwarya, R. D., 307
Akshay Reddy, V., 495
Almeida, Genesis Dayana Pinto, 707, 789
Anantha Babu, S., 1
Anita, J. P., 179
Anita, M., 45
Anjali, C., 769
Anjum, Mohd, 751
Antariksha, K., 579
Ayyadurai, M., 903

B

Balli, Priyanka, 243
Berde, Jui Sanjay, 207
Bhaskar, S. V., 495
Bora, Aryamaan, 579

C

Choudhary, Parul, 139
Chowdhury, Afzal Un Nayeem, 867
Christopher, 189

D

Dadhwal, Shivank Kumar, 531
Dang, Thuan Minh, 841
Dao, Tuan Anh, 59
Das, Ayanmani, 579
Das, Siddhant Kumar, 579

Das, Tanaya, 85
Deepa, V. J., 307
Dehraj, Pooja, 369
Deolekar, Rugved V., 223
Dermawan, Gerry, 189
Deval, Vipin, 127
Devi, S., 15
Dhalla, Aashish, 329
Dinh, Toai Cong, 841
Divya, P., 1
Dongre, Manisha H., 385
Dongre, Snehlata, 679
Duong, Tin Tri, 841
Durgashree, J., 307

F

Fahrezi, Fazri, 657
Fatima, Ummatul, 751
Firnandes, Steven Gusda, 657

G

Gambardella, C., 479
Ganguly, Amrita, 579
Gaol, Ford Lumban, 189, 319, 441, 657,
801, 935
Gayathry Devi, M., 769
Geluvaraj, B., 495
Gilda, Shlok, 329
Goel, Vikas, 127
Gohokar, Vinaya, 461
Gomez, Shannon, 415
Gordienko, Yuri, 885
Gowrishankar, S., 273, 307, 385
Gowtham, K., 687

Guaña, Carlos Alberto Ramos, 707
 Gupta, Abhishek, 169
 Gussain, Rohit, 369

H

Hareesh Chandran, S., 97
 Hariharan, T., 687
 Harthi Al, Faiza, 641
 Hegde, Venkatesh, 273
 Hemalatha, K., 425, 687
 Hutagalung, Fonny, 657

I

Imam, Rafif, 441
 Iswardhana, Aria Muhammad, 657

J

Jacob, Rex Rony, 415
 Jaiganesh, M., 1
 Jain, Kurunandan, 607, 629, 951
 Jain, Tanvi, 329
 Joglekar, Prajakta, 461
 John Basha, M., 1
 John, Jisha, 769

K

Kadiwal, Shashank M., 273
 Kamil, Mufti Ikhsan, 657
 Kapudasu, Rohith, 951
 Karthikeyan, G., 507
 Keerthanaa, K., 243
 Khan, Nayeem Ahmad, 741
 Kulkarni, Shirish S., 207
 Kumar Chandar, S., 965
 Kumar, Krishan, 451
 Kumar, Nitin, 369
 Kumar, T., 563
 Kumar, Vineet, 353
 Kural Eniyavan, S. M., 425
 Kurniawan, Achmad Rizky, 441
 Kushwaha, Arun Kumar, 127

L

Lahkar, Rishabh, 579
 Lam, Dung Ngoc Vo, 59
 Larissa, Jessica Adine, 935
 Lokesh, S., 111
 Luong, Huong Hoang, 59, 841

M

Madhuri, Pothuka, 811
 Mageshkumar, G., 425, 593, 687, 903
 Mageshwari, S., 669
 Maha Lakshmi, G., 521
 Maher Shalal, D., 593
 Mahesh, R., 425
 Mahesh, Rohit, 415
 Maheswari, B., 45
 Manasa, Chinta Sai, 179
 Mandal, Souraneel, 85
 Mangalam, H., 669
 Mangaonkar, Nikhita, 169
 Manikandan, V., 1
 Manoji, T., 593
 Maruthi Sai Saketh, P., 111
 Mathew, Elizabeth, 769
 Matsuo, Tokuro, 189, 319, 441, 657, 801, 935
 Mayukha, S., 97
 Meishree, S., 669
 Minu, R. I., 31
 Mishra, Girish, 531
 Modi, Ritiksha, 399
 Mohan, Navya, 179
 Monisha, P. T., 769
 Mozumder, Ziaul Hasan, 867
 Muruganantham, S., 593
 Muthukumaran, N., 779

N

Nalayini, C. M., 857
 Natarajan, M., 543
 Nguyen, Hai Thanh, 59, 841
 Nguyen, Hy Khang Vuong, 59
 Nguyen, Khoi Tuan Huynh, 59
 Nguyen, Phuong Thuy Thi, 823
 Nguyen, Tong Duc, 841
 Nikkath Bushra, S., 45
 Nirmala, G., 45
 Nithyeswari, C., 507

O

Ortega, Evelyn Paulina Rovalino, 789

P

Parente, R., 479
 Pathak, Pooja, 139
 Patidar, Sanjay, 293, 451, 919
 Pavliuchenko, Ivan, 885
 Pavliuchenko, Mykhaylo, 885

Pham, Nhan Trong Van, 59
 Phan, Nghia Trong Le, 841
 Pilla, Rohit, 629
 Ponnusamy, R., 563
 Pooja, C., 385
 Pooja, P., 521
 Prabhakara Rao, Saggurthi, 425
 Prabhu, E., 811
 Pradanang, Riyan, 319
 Prajnowira, Ferdianto, 935
 Prathiksha, B., 669
 Putra, Armando Rama, 441

Q

Quinde, María Cristina Pérez, 707
 Qurashi Al, Mohammed, 741

R

Rafif, Muhammad Resnandya, 189
 Raghuraman, G., 149
 Rajalakshmi, M., 15
 Rajan, B. Pradheep T., 779
 Ramteke, Bharati, 679
 Ranjitha, A. R., 243
 Rao Deshpande, Prasanna V., 965
 Rao, Saggurthi Prabhakara, 687
 Rastogi, Anshay, 127
 Ravikumar, Aswathy, 111
 Rayhan, Adhitya, 319
 Ray, Partha Pratim, 73
 Revathi, M., 149
 Revathi, S., 243
 Robinson, Vijitha, 415
 Rubianto, Rully, 189

S

Sabitha, B., 97
 Sahabuddin, Syed, 415
 Sahana, C., 307
 Saha, Sajib, 85
 Saleema, J. S., 717
 Salma, Subia, 353
 Sandhya, T., 495
 Santhosh, K., 495
 Saran, D. V. Guru, 607
 Saranya, S., 15
 Sasikala, N., 495
 Sathyabama, A. R., 857
 Sewariya, Manish, 293
 Shahab, Sana, 751
 Sharma, Minshul, 369

Shrivathsa, N. V., 273
 Singh, Akanksha, 919
 Singh, Pravesh, 127
 Singh, Soumya, 127
 Sivaraman, M., 233
 Siva Subramanian, R., 45
 Snegha, C., 903
 Soneji, Pranav, 169
 Sony Priya, S., 31
 Sreedhara, Shreehari Harthi, 353
 Srinivasa, A. H., 273, 385
 Sriraman, Harini, 111
 Stirenko, Sergii, 885
 Suganya, P. V. Raja, 857
 Sukardi, Oggasa, 801
 Sumitha, J., 233
 Sureka, M., 903
 Suruthi, R. M., 15
 Suthagar, S., 425, 593, 687, 903

T

Ta, Hien Thu Thi, 823
 Thakkar, Nisarg, 399
 Thakkar, Pratham, 399
 Tilva, Shalin, 399
 Tomko, Maryna, 885
 Tompunu, Sherina Gloria, 935
 Torres, José David Sillagana, 789
 Touzene, Abderezak, 641
 Tran, Trung, 823

U

Umar, M. Sarosh, 751

V

Vadalia, Brijesh, 399
 Varghese, Jithy, 717
 Vásquez, Carlos Fabián Martínez, 707, 789
 Veena, A., 273, 307, 385
 Velmurugan, S., 903
 Venkatakrishnan, N., 543
 Vijayakumar, Preethi, 769
 Vindhya, J., 385
 Vinith, R. S. V. S., 521
 Visumathi, J., 149

W

Wankhade, Sunil, 223
 Wanve, Mayur, 223
 Wijaksana, Arista, 935
 Winarlle, Michael Adrian, 319

Y

Yellampalli, Siva Sankar, [521](#)

Z

Zhalifunnas, Rifqi Farhan, [441](#)