

Lecture Notes in Networks and Systems 475

G. Ranganathan
Robert Bestak
Xavier Fernando *Editors*

Pervasive Computing and Social Networking

Proceedings of ICPCSN 2022

 Springer

Lecture Notes in Networks and Systems

Volume 475

Series Editor

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

Advisory Editors

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

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

More information about this series at <https://link.springer.com/bookseries/15179>

G. Ranganathan · Robert Bestak · Xavier Fernando
Editors

Pervasive Computing and Social Networking

Proceedings of ICPCSN 2022

 Springer

Editors

G. Ranganathan
Electronics and Communication
Engineering
Gnanamani College of Technology
Namakkal, India

Robert Bestak
Czech Technical University in Prague
Prague, Czech Republic

Xavier Fernando
Ryerson Communications Lab
Toronto, ON, Canada

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-19-2839-0

ISBN 978-981-19-2840-6 (eBook)

<https://doi.org/10.1007/978-981-19-2840-6>

© 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

We are honored to dedicate the proceedings of the second edition of Pervasive Computing and Social Networking Proceedings to all the participants, organizers, and editors of ICPSN 2022.

Preface

The Second International Conference on Pervasive Computing and Social Networking (ICPCSN 2022) provides researchers, academicians, and industrialists across the globe a significant opportunity to discuss and present the research models and results in the areas of computing and network engineering. The second series of this conference welcomes both theoretical and practical research contributions by creating a considerable impact on the networking and computing domains.

ICPCSN 2022 conference is sponsored by Narasu's Sarathy Institute of Technology, Salem, India, and the proceedings of ICPCSN 2022 will be submitted to Springer publications. Henceforth, the Conference event was held at Narasu's Sarathy Institute of Technology in Salem, India, features keynote talk by outstanding researchers and diverse paper presentation sessions with six research papers on the significant and state-of-the-art topics in computing and networking paradigm. The panel has been led by the renowned research experts in computing and networking paradigm from different universities and research organizations, and the committee has also invited the conference attendees to actively participate in the discussion.

The main aim of ICPCSN 2022 conference is to address a wide range of research in computing and networking topics including theory, methodologies, tools, and applications. Each and every paper submitted to ICPCSN 2022 was peer reviewed by at least two technical program chairs, who evaluated the research works based on their novelty, research significance, technical contribution, and real-time application. In particular, all the papers were reviewed by three to four reviewers and provided their feedback for the further decision process.

For this second edition of the conference, we have totally received 272 papers of which 60 papers made to the final program. The contribution of authors to the conference has been highly appreciated, and we are pleased to invite them to continue and contribute to the future editions of ICPCSN conference.

Namakkal, India
Prague, Czech Republic
Toronto, Canada

G. Ranganathan
Robert Bestak
Xavier Fernando

Contents

| | |
|---|----|
| A Feature Extraction Based Ensemble Data Clustering for Healthcare Applications | 1 |
| D. Karthika and N. Jayashri | |
| Analysis of Axial Triradius to Detect Congenital Heart Diseases | 9 |
| Y. Mahesha and C. Nagaraju | |
| Forecasting Flash Floods with Optimized Adaptive Neuro-Fuzzy Inference System and Internet of Things | 23 |
| M. Pushpa Rani, Bashiru Aremu, and Xavier Fernando | |
| A Survey on HTTP Flooding—A Distributed Denial of Service Attack | 39 |
| Hrshikesh Khandare, Saurabh Jain, and Rajesh Doriya | |
| Tracking and Monitoring of Soldiers Using IoT and GPS | 53 |
| Adlin Sheeba, A. Vinora, P. Ananth, K. Nithya, V. Nisha Jenipher, and U. Surya | |
| Study of Vulnerabilities in the Cryptography Algorithms | 65 |
| Harjis Ahuja, Ruchita Bapna, Gargee Bhase, and Narendra Shekokar | |
| Awaaz: A Sign Language and Voice Conversion Tool for Deaf-Dumb People | 77 |
| Bharat Taralekar, Rutuja Hinge, Chaitanya Bisne, Amberish Deshmukh, and Vidya Darekar | |
| Decentralized Application (DApp) for Microfinance Using a Blockchain Network | 95 |
| C. Jyothi and M. Supriya | |

| | |
|---|-----|
| Experimental Study on LoRaWAN Technology Applied to Vehicular-to-Infrastructure (V2I) Communication | 109 |
| Samuel Alexander A. Pasia, Vince Matthew A. Rivera, Jereme Adriane D. G. Sy, Bianca Clarisse Y. Tan, Gerald P. Arada, and Elmer R. Magsino | |
| Optimal Densely Connected Networks with Pyramid Spatial Matching Scheme for Visual Place Recognition | 123 |
| P. Sasikumar and S. Sathiamoorthy | |
| Virtual Reality as a Teaching Resource in Higher Education: Professors' Assessment | 139 |
| Álvaro Antón-Sancho, Diego Vergara-Rodríguez, David G. Calatayud, and Pablo Fernández-Arias | |
| Automated Intelligent Hematology Classification System Using Image Processing and Neural Networks | 151 |
| B. G. Taralekar, Prithviraj Chauhan, Shrinath Palwankar, Celsy Phillips, and Sarang Patil | |
| Nonlinear Integrity Algorithm for Blockchain Based Supply Chain Databases | 169 |
| Mani Deep Karumanchi, J. I. Sheeba, and S. Pradeep Devaneyan | |
| Topic Identification and Prediction Using Sanskrit Hysynset | 183 |
| Prafulla B. Bafna and Jatinderkumar R. Saini | |
| Filling the Knowledge Base of the Filtering System Algorithm Development | 197 |
| Elshod Khaydarov | |
| Development of a Linear Scale Consensus Method of the Blockchain System | 205 |
| Gennady Shvachych, Boris Moroz, Maksym Khylyko, Andrii Matviichuk, Vladyslava Kozenkova, and Volodymyr Busygin | |
| Acquisition of a CNC Router for a Joinery in Brazil: An Approach from VFT, SAPEVO-M and WASPAS Methods | 219 |
| Lucas Ramon dos Santos Hermogenes, Igor Pinheiro de Araújo Costa, Marcos dos Santos, and Carlos Francisco Simões Gomes | |
| General Collective Intelligence as a Platform for Social Technology | 233 |
| Andy E. Williams | |
| Risk Analysis Related to the Corrosion of Atmospheric Storage Tanks: A Multimethodological Approach Using the FRAM and the ELECTRE-MOR Methods | 245 |
| Érick Pinto Moreira, Igor Pinheiro de Araújo Costa, Fernando Benedicto Mainer, Carlos Francisco Simões Gomes, and Marcos dos Santos | |

Implementing Braking and Acceleration Features for a Car-Following Intelligent Vehicle 259
 Johann Carlo R. Marasigan, Gian Paolo T. Mayuga, Elmer R. Magsino, and Gerald P. Arada

Homogeneous Map Partitioning Employing the Effective Regions of Movement Method 271
 Elmer R. Magsino

Demand-Based Deployment of Electric Vehicle Charging Stations Employing Empirical Mobility Dataset 285
 Camille Merlin S. Tan and Elmer R. Magsino

Design and Development of a Digital Preservation Voice Enabled Application for Cultural Heritage Towards Fishing in Vernacular Language 295
 Prasad Vadamodula, R. Cristin, and T. Daniya

Krishi Mitra: Crop and Fertilizer Recommendations System Using Machine Learning Algorithm 309
 Vijay Mane, Akash Gajbhiye, Amisha, Chinmay Deshmukh, and Kunal Gaikwad

Development of DDoS Attack Detection Approach in Software Defined Network Using Support Vector Machine Classifier 319
 Oluwashola David Adeniji, Deji Babatunde Adekeye, Sunday Adeola Ajagbe, Ademola Olusola Adesina, Yetunde Josephine Oguns, and Matthew Abiola Oladipupo

Sensor Data Fusion Methods for Driverless Vehicle System: A Review 333
 Nitheesh Kurian and K. Vadivukkarasi

A Novel Neural Network Based Model for Diabetes Prediction Using Multilayer Perceptron and Jrip Classifier 345
 B. Sreedevi, Durga Karthik, J. Glory Thephoral, M. Jeya Pandian, and G. Revathy

Development of a Neighborhood Based Adaptive Heterogeneous Oversampling Ensemble Classifier for Imbalanced Binary Class Datasets 353
 S. Santha Subbulaxmi and G. Arumugam

Real Time Facial Emotions Detection of Multiple Faces Using Deep Learning 363
 Ankita Kshirsagar, Neetesh Gupta, and Bhupendra Verma

Mobile Application for Online Pharmacy: A-Pharma App 377
 P. S. JosephNg, A. A. A. Al-Maari, K. Y. Phan, J. T. Lim, and E. H. Lim

A Convolutional Neural Network Incorporated Long Short-Term Memory with Autoencoder for Covid-19 Intensity Levels Detection 389
 J. Deepika and J. Akilandeswari

Survey on Attendance System Using Face Recognition 407
 D. Pradeep, A. Bhuvaneswari, M. Nandhini, A. Roshini Begum, and N. Swetha

A Spectral-Spatial Classification of Hyperspectral Image Using Domain Transform Interpolated Convolution Filter 421
 M. Preethi, C. Velayutham, and S. Arumugaperumal

Performance Analysis of Various Asymmetric Public-Key Cryptosystem 437
 Amogh Desai, Virang Parekh, Utsav Unadkat, and Narendra Shekokar

Design of a Web-Based Platform: Event-Venues Booking and Management System 451
 P. S. JosephNg, S. M. Al-Sofi, K. Y. Phan, J. T. Lim, and S. C. Lai

Evaluating an Automated Analysis Using Machine Learning and Natural Language Processing Approaches to Classify Computer Science Students’ Reflective Writing 463
 Huda Alrashidi, Nouf Almujaally, Methaq Kadhum, Thomas Daniel Ullmann, and Mike Joy

Secured and Lightweight Key Distribution Mechanism for Wireless Sensor Networks 479
 P. Ezhil Roja and D. S. Misbha

Detection of Exercise and Cooking Scene for Assitance of Visually Impaired People 493
 Shripad Bhatlawande, Swati Shilaskar, Anant Abhyankar, Mahesh Ahire, Ankush Chadgal, and Jyoti Madake

Electronic System for Navigation of Visually Impaired People 509
 Jyoti Madake, Bhumika Bijane, Geetai Charde, Abhijit Chine, Shripad Bhatlawande, and Swati Shilaskar

Proposal of Criteria for Selection of Oil Tank Maintenance Companies at Transpetro Through Multimethodological Approaches 521
 Eric Bremm De Carvalho, Miguel Ângelo Lellis Moreira, Adilson Vilarinho Terra, Carlos Francisco Simões Gomes, and Marcos dos Santos

A Deep Learning-Based Approach with Semi-supervised Level Set Loss for Infant Brain MRI Segmentation 533
 Minh-Nhat Trinh, Van-Truong Pham, and Thi-Thao Tran

A Reinforcement Learning Based Adaptive Traffic Signal Control for Vehicular Networks 547
S. P. Krishnendhu, Mainampati Vigneshwari Reddy, Thulunga Basumatary, and Prabu Mohandas

The Structure of the Vulnerability Detection System on Web Servers ... 563
Aziz Ibrokhimov

Comparison of Image Blending Using Cycle GAN and Traditional Approach 571
Medha Wyawahare, Ninad Ekbote, Sameer Pimperkhede, Atharva Deshpande, Pranav Bapat, and Ishan Aphale

Botnet Attacks Detection Using Embedded Feature Selection Methods for Secure IOMT Environment 585
A. Karthick Kumar, K. Vadivukkarasi, R. Dayana, and P. Malarvezhi

An Accuracy Based Comparative Study on Different Techniques and Challenges for Sentiment Analysis 601
Radha Krishna Jana and Saikat Maity

Pseudo-CT Generation from MRI Images for Bone Lesion Detection Using Deep Learning Approach 621
S. Sreeja and D. Muhammad Noorul Mubarak

Simulation of Topology Based VANET Routing Protocols Using NS3 ... 633
Mahabub Subhani Pathan and K. Annapurna

A Fast Algorithm for Hunting State-Backed Twitter Trolls 643
Shaaban Sahmoud, Abdelrahman Abdellatif, and Youssof Ragheb

A Review on Artificial Intelligence Based E-Learning System 659
U. Arun Kumar, G. Mahendran, and S. Gobhinath

Text to Image Generation Using Gan 673
Rajni Jindal, V. Sriram, Vishesh Aggarwal, and Vishesh Jain

SARS-CoV-2: Transmission Predictive Tool Based on Policy Measures Adopted by Countries Using Basic Statistics 685
Charles Roberto Telles and Archisman Roy

Performance Evaluation of an IoT Based Fetal Heart Monitoring Device 697
Olubunmi Ige, Adedotun Adetunla, Joshua Adewolu, and Adeyinka Adeoye

A Novel Approach for Handling Imbalanced Data in Breast Cancer Dataset 709
Nagateja Banothu and M. Prabu

Investigating Data Mining Trend in Cybercrime Among Youths 725
Ademola Olusola Adesina, Sunday Adeola Ajagbe,
Olakunle Sunday Afolabi, Oluwashola David Adeniji,
and Olalekan Ibrahim Ajimobi

**Concatenation of SEPIC and Matrix Converter with LCL Filter
for Vehicle to Grid Application 743**
Anushka Tripathi, S. P. Soundharya, R. Vigneshwar, and M. R. Rashmi

**Creation of SDIoT Testbed for DDoS Attack Using Mininet:
Experimental Study 759**
B. Keerthana, Mamatha Balachandra, Harishchandra Hebbar,
and Balachandra Muniyal

**Automatic Eye Disease Detection Using Machine Learning
and Deep Learning Models 773**
Nouf Badah, Amal Algefes, Ashwaq AlArjani, and Raouia Mokni

**An Automated System to Preprocess and Classify Medical Digital
X-Rays 789**
Sumera, K. Vaidehi, and R. Manivannan

PSTN Enterprise Collaboration Systems 803
P. S. JosephNg, R. C. Yeo, J. H. Wong, P. H. Tan, Y. H. Yong,
and S. L. See

Author Index 817

About the Editors

Dr. G. Ranganathan working as a Dean, Department of Electronics and Communication Engineering at Gnanamani college of technology, Namakkal, India. He has done his PhD in the Faculty of Information and Communication Engineering from Anna University, Chennai in the year 2013. His research thesis was in the area of BioMedical Signal Processing. He has a total of 30+ years of experience both in industry, teaching and research. He has guided several projects works for many UG and PG Students in the areas of BioMedical Signal Processing. He has published more than 35 research papers in International and National Journals and Conferences. He has also co-authored many books in electrical and electronics subjects. He has served as Referee for many reputed International Journals published by Elsevier, Springer, Taylor and Francis, etc. He has membership in various professional bodies like ISTE, IAENG etc., and has actively involved himself in organizing various international and national level conferences, symposiums, seminars, etc.



Dr. Robert Bestak received the PhD degree in Computer Science from ENST Paris, France (2003), and MSc degree in Telecommunications from Czech Technical University in Prague, CTU, Czech Republic (1999). Since 2004, he has been Assistant Professor at the Department of Telecommunication Engineering, Faculty of Electrical Engineering, CTU. He took part in several national, EU, and third-party research projects. He is Czech Representative in the IFIP TC6 organization and Vice-chair of working group TC6 WG6.8. He serves as Steering and Technical Program Committee Member of many IEEE/IFIP conferences (Networking, WMNC, NGMAST, etc.), and he is Member of the editorial board of several international journals (Computers & Electrical Engineering, Electronic Commerce Research Journal, etc.). His research interests include 5G networks, spectrum management, and big data in mobile networks.

Dr. Xavier Fernando is Director of Ryerson Communications Lab that has received total research funding of \$3,185,000.00 since 2008 from industry and government (including joint grants). He has (co-)authored over 200 research articles, three books (one translated to Mandarin), and several chapters and holds few patents. The present

and past members and affiliates of this laboratory can be found at this LinkedIn group and Facebook page. He was IEEE Communications Society Distinguished Lecturer and has delivered over 50 invited talks and keynote presentations all over the world. He was Member in the IEEE Communications Society (COMSOC) Education Board Working Group on Wireless Communications. He was Chair of IEEE Canada Humanitarian Initiatives Committee 2017–18. He was also Chair of the IEEE Toronto Section and IEEE Canada Central Area.

A Feature Extraction Based Ensemble Data Clustering for Healthcare Applications



D. Karthika  and N. Jayashri 

Abstract Multidimensional data sets are becoming more common in almost all research domains, and extracting insight from them necessitates complicated data analysis approaches. The increased dimensionality of data has made data mining and machine learning difficult. Early disease forecasts assist physicians to make effective decisions to save patients' survival. As a result, dimensionality reduction approaches provide a roadmap for resolving this issue, in terms of efficiency and effectiveness by reducing unnecessary, irrelevant, and noisy data, making the learning process faster in terms of computation time and accuracy. For medical data, this paper proposes a Feature Extraction based Ensemble Data Clustering (FEEDC) approach. It also includes advanced dimensionality reduction techniques and feature extraction methods, as well as MapReduce. The centroid selection is done using a support vector machine classifier during clustering. The ensemble member selection algorithm is the firefly algorithm. Finally, a clustering solution with the Normalized cut (Ncut) algorithm is used to diagnose conditions such as heart disease and breast cancer at the initial phase. The results are obtained using two UCI datasets, which achieves more accurate results.

Keywords Dimensionality reduction · High dimensional data · Feature extraction · Healthcare · Evolutionary algorithms

1 Introduction

In matters of extraordinary significance that are money-related, restorative, social, or other suggestions, it may be a fact that they commonly look for a moment assumption, sometime lately choosing, and in some cases numerous more [1]. Therefore,

D. Karthika (✉)

Department of Computer Science, SDNB Vaishnav College for Women, Chrompet, Chennai, Tamil Nadu, India
e-mail: d.karthi666@gmail.com

N. Jayashri

Faculty of Computer Applications, Dr.M.G.R Educational and Research Institute, (Deemed to Be University), Chennai, Tamil Nadu, India

numerous classification strategies apply a special classifier, whereas others create a multi-classification-based approach. In this way, procedures based on classifiers combination have acquired an extraordinary center since a long time, and in these days are recognized as an established pattern acknowledgment seed. At that point, one of the foremost dynamic zones of investigation in supervised learning has pondered strategies for building great outfits of classifiers. The most revelation is that outfits are regularly much more exact than the individual classifiers that make them up [2]. Huge information analytics is the handling in which the data units are analyzed, and the mystery data are displayed. Machine learning calculations are essentially coordinated into computers and information sources, such that information and data are extricated and encouraged into the frameworks for them to be dealt more quickly and successfully. In expansion, these machine learning strategies are moreover valuable for methodologies to minimize dimensionality. Parallel preparing information examination calculations are valuable in this approach for time imprisonment applications [3]. The key errands concerning copy information dimensionality lessening applications are times, speed, and execution. Given the anticipated increments, the methods of AI and certainty mining might not be powerful in state of mind to the tribulation of dimensionality for tall dimensional substances, and the consistency and ampleness of the investigation may fall apart quickly. This paper proposes a Feature Extraction-based Ensemble Data Clustering (FEEDC) system for curative information that gives a wide outline on progressed dimensionality diminishment strategy and feature extraction strategy, alongside MapReduce. The clustering is performed where the centroid optimal is done by a Support Vector Machine (SVM) classifier. The firefly algorithm is utilized for the outfit part determination. At last, the clustering arrangement with normalized calculation is examined to anticipate the infections like heart disease and breast cancer in the early stage. The evaluation is performed with two UCI datasets, which results in higher exactness.

2 Literature Review

The authors in [1] grasped the highlight building standards to progress the quality of machine learning demonstration to quicken the revelation of possibly curiously catalytic materials. UCI research facility information was utilized in [2] to recognize designs utilizing Artificial Neural Networks (ANN), Decision Trees (DT), SVM, and Gullible Bayes calculations. The execution of these calculations was compared with the strategy in which an ANN with 18 neurons within the covered-up layer was utilized. The strategy gave superior results compared to other strategies. The researchers in [3] utilized Convolutional Neural Networks (CNN) to foresee heart malady based on ECG signals. The procedure considered the heart cycles within the preparing stage with different beginning focuses from the ElectroCardioGram (ECG) signals. CNN can create highlights with diverse positions in the testing stage. A neural organize-based classification framework of Fetal Heart Rate (FHR) signals to decrease the mistakes caused by human examination was displayed by the creators

in [4]. The calculation was prepared on FHR information that was analyzed in genuine time. The authors have outlined a repetitive neural organize called “MKRNN” and a convolutional neural organize classification strategy called “MKNet”.

3 Proposed Methodology

The Feature Extraction based Ensemble Data Clustering (FEEDC) Architecture is being used to generate ensemble datasets, and it enables the Ensemble Member Chosen procedure, dependent on local and globally optimal solutions [5].

$$E_s = \{f(X_1, A_1), (X_2, A_2), \dots (X_{Y'}, A_{Y'})\} \text{ (where } Y' < Y \text{)}$$

The current objective function is well suitable for local exploration, focusing on the previous ensemble, E_s global search. Feature extraction is a technique for obtaining new features from an original dataset, that is highly useful, when it is necessary to reduce the number of resources required for processing, while not missing out on important characteristics. The dataset $FV = \{fv_1, \dots, fv_n\}$, are complete by biased graph $G(FV, A)$ where ‘fv’ is indicated as a usual Feature Vector. These vectors are associated with the path of the vertices. A is considered as the similarity environment by weight value a_{lm} which is connected by the boundaries amongst fv_l and fv_m (see Fig. 1).

$$a_{lm} = \begin{cases} \rho_{lm} & \text{if } fv_l \in N_c(fv_m) \\ 0 & \text{else} \end{cases} \quad (1)$$

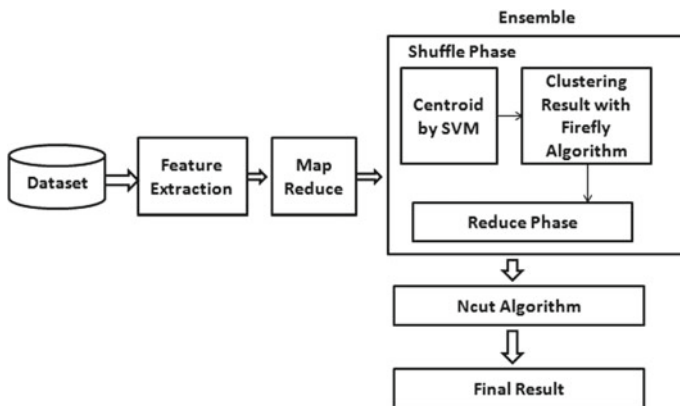


Fig. 1 Feature extraction based ensemble data clustering framework

MapReduce (MR) [7] is a well-known programming approach for dealing with big datasets [6]. The three basic phases of the MR model are map, shuffle, and reduce. The MR paradigm is founded on the concept of information creation, which is described as a key-attribute pair. This is demonstrated in the following form:

$$\text{map}(\text{key1}, \text{attribute1}) \rightarrow \{(\text{key2}, \text{attribute2}), \dots\} \quad (2)$$

Support Vector Machine (SVM) is a method of organizing estimates to deal with the centroid selected issue of clustering capabilities. SVM holds high-dimensional data in a portion-organized format [7]. An advanced SVM demonstration is created via the use of parametric direct premise work and regression estimate of P is used in centroid determination for gathering large dimensional datasets to reduce the structural potential of SVM.

$$p_k = f(a) = \langle \omega, \phi(a) \rangle + y \quad (3)$$

3.1 Fuzzy Firefly Algorithm (FFA) for Ensemble Member Chosen (EMC)

As a result, in FA, each old clustering result can be supplied a few times based on a comparison of its brightness with that of the unused ones [8]. Any unused section picked X_i , might be a firefly 'i' region throughout the display cycle. When the accuracy of the part is selected, 'i' is improved over that of another part chosen, j. The distance between the fireflies i and j is calculated using the following expression.

$$r_{ij} = \sqrt{(X_i - X_j)^2} \quad (4)$$

3.2 Consensus Matrix Function

Properties form the complete metric 'm' by calculating a variety of adjacency metrics (m_1, m_2, \dots, m_Y) as trails.

$$m = \frac{1}{Y} \sum_{y=1}^Y m^y \quad (5)$$

In this work with matrix O, a Normalized cut method (Ncut) [5] is proposed.

4 Results and Discussion

Accuracy, NMI, and ARI are used to analyze the results of two UCI datasets that are chosen for the research. The definitive value is applied to these variables after 20 runs of repeats. The outcomes of current studies are assessed using datasets, as shown in Table 1, with total data samples ‘n,’ characteristics ‘m,’ and classifications ‘y’. The entire clustering procedure is carried out using MATLAB R2014 on an Intel(R) Core-2 computer with a 1.4 GHz processor and 4 GB RAM running on a Windows-10 operating system. The entire technique is centered on repetition, up to a hundred times.

Recognized X across in clusters is a real-world consequence $X\{k_1, k_2, \dots, kn_0\}$ and the resulting X’ accomplished over clustering algorithms over k0 clusters, $X' = \{k'_1, \dots, k'_k\}$ gain the advantage of NMI, which is denoted by the following:

$$NMI(X, X') = \frac{2H_1(X, X')}{H_2(X) + H_2(X')} \tag{6}$$

$$H_1(X, X') = \sum_h \sum_l \frac{|k_h \cap k'_l|}{n} \log \frac{n|k_h \cap k'_l|}{|k_h||k'_l|} \tag{7}$$

where $h \in \{1, \dots, k\}$ and $l \in \{1, \dots, k'\}$ give appropriate information and represent the basic value of the equivalent clusters. The formula for calculating metrics is as follows (see Table 2):

Table 1 Dataset samples

| Datasets | n | m | y |
|-------------------------|-----|----|----|
| Heart disease | 360 | 30 | 15 |
| Breast cancer wisconsin | 569 | 32 | 2 |

Table 2 Metrics evaluation with ISSCE and FEEDC

| S. no. | Datasets | Metrics | ISSCE | FEEDC |
|--------|-------------------------|--------------|---------|---------|
| 1 | Heart Disease | Accuracy (%) | 74.666 | 83.8817 |
| | | ARI | 0.6768 | 0.8671 |
| | | NMI | 0.7526 | 0.9078 |
| 2 | Breast Cancer Wisconsin | Accuracy (%) | 78.3824 | 84.3065 |
| | | ARI | 0.5298 | 0.7798 |
| | | NMI (%) | 0.4591 | 0.7217 |

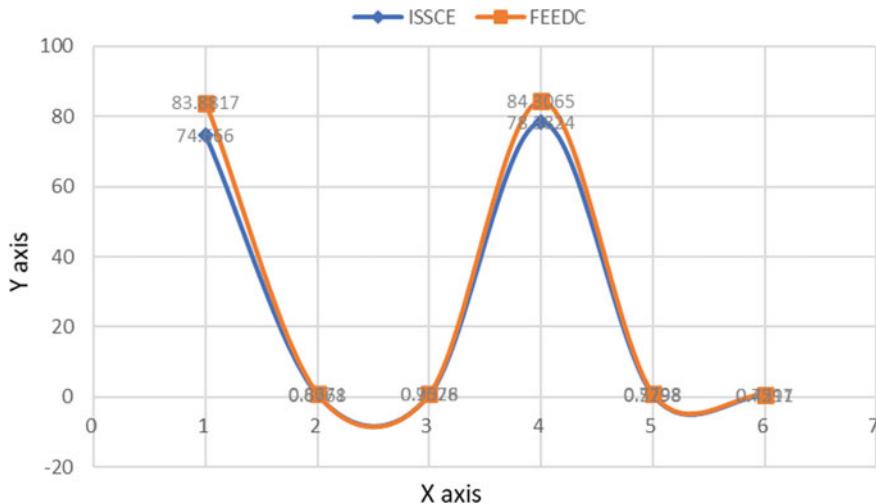


Fig. 2 Result of ISSCE and FEEDC

$$ARI(I, I') = \frac{\sum_{h=1}^k \sum_{l=1}^{k'} \binom{|k_h \cap k'_l|}{2} - p_3}{\frac{1}{2}(p_1 + p_2) - p_3} \quad (8)$$

The results show that the proposed algorithm FEEDC exceeds the previous algorithm (Figs. 2 and 3), with 83.8817% accuracy for heart disease and 84.3065% accuracy for Breast Cancer Wisconsin, correspondingly.

5 Conclusion and Future Work

The dimensionality reduction challenge is the process of decreasing the dimension of massive features and functionality, into a unified limited set that forms a huge space in n-dimensional space, as described in this paper. One of the major concerns of researchers working on disease diagnostic systems seems to be reliability. For healthcare records, this research suggests a Feature Extraction based Ensemble Data Clustering (FEEDC) technique that explores advanced feature extraction methods and feature extraction methods, as well as MapReduce. The centroid selection is done using a support vector machine classifier during clustering. The ensemble element selection method is performed by the firefly technique. Finally, the use of a clustering solution with a normalized cut method to forecast illnesses such as heart disease and breast cancer at an early stage is described. The results are obtained using two UCI datasets, which achieves more accurate results. The experimental findings suggest that accuracy is improved by using the complicated dataset. Future research will

investigate how to determine feature values based on the dataset's structure and temporal complexity.

References

1. Li Z, Ma X, Xin H (2017) 'Feature engineering of machine-learning chemisorption models for catalyst design.' *Catal Today* 280:232–238
2. Cheng C-A, Chiu H-W (2017) An artificial neural network model for the evaluation of carotid artery stenting prognosis using a national-wide database. In: Proceedings of the 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Jul 2017, pp 2566–2569
3. Zaman S, Toufiq R (2017) Codon based back propagation neural network approach to classify hypertension gene sequences. In: Proceedings of the international conference on electrical, communication, and computer engineering (ECCE), Feb 2017, pp 443–446
4. Tang H, Wang T, Li M, Yang X (Dec 2018) The design and implementation of cardiocography signals classification algorithm based on neural network. *Comput Math Methods Med* 2018. Art no 8568617
5. Zhang Y, Zhao Z (2017) Fetal state assessment based on cardiocography parameters using PCA and AdaBoost. In: Proceedings of the 10th international congress on image and signal processing, biomedical engineering information (CISP-BMEI), Oct 2017, pp 1–6.
6. Lobo Marques JA, Cortez PC, Madeiro JPDV, Fong SJ, Schlindwein FS, Albuquerque VHCD (2019) Automatic cardiocography diagnostic system based on Hilbert transform and adaptive threshold technique. *IEEE Access* 7:73085–73094
7. Dadi HS, Pillutla GKM (2016) Improved face recognition rate using HOG features and SVM classifier. *IOSR J Electron Commun Eng* 11(04):34–44
8. Cheng X, Zhao S-G, Xiao X, Chou K-C (2017) ATC-mISF: A multi-label classifier for predicting the classes of anatomical therapeutic chemicals. *Bioinformatics* 33(3):341–346

Analysis of Axial Triradius to Detect Congenital Heart Diseases



Y. Mahesha and C. Nagaraju

Abstract This paper proposes an image processing based method to conduct a study on the relationship between the angle at axial triradius on palm image and different Congenital Heart Diseases (CHDs). All the previous research works on finding the relationship between axial triradius and CHDs have considered ink palm prints, and the angle at axial triradius is found using geometrical instruments. All these researchers have taken a small sample size and provided a conclusion. Hence, further study is required with a larger number of samples for a better conclusion. Working with ink palm print and geometrical instruments on large sample size is a tedious task and hence computer-based method is required to identify triradii and to find the “atd” angle. In the present paper, axial triradius is identified using a template matching algorithm. The vector approach has been used to calculate the “atd” angle.

Keywords “atd” angle · Axial triradius · Fallot’s Tetralogy · Pattern matching · Triradii

1 Introduction

In this advanced era, people are facing different kinds of health issues. These health issues include minor, moderate and life-threatening diseases. Since heart-related diseases often lead to sudden death these diseases are considered to be life-threatening. Some heart diseases are genetically related and they are called Congenital Heart Disease (CHD). Tetralogy of Fallot (TOF), Ventricular Septal Defect (VSD), Atrial Septal Defect (ASD), and Coarctation of Aorta (COA) are some commonly known CHDs. In FT, four types of defects are associated with the heart [1]. In ASD, there exists a problem in the septum at the top chambers of the heart [2]. In VSD, there exists a problem in the septum at the bottom chambers of the

Y. Mahesha (✉) · C. Nagaraju
The National Institute of Engineering, Mysore, Karnataka, India
e-mail: maheshy029@gmail.com

C. Nagaraju
e-mail: nagaraj@nie.ac.in

heart [3]. In COA, there is a congenital narrowing of the aorta [4]. Since the CHDs are genetically related doctors have conducted an empirical study to find the association between CHDs and palm patterns. The researchers have used palm patterns such as loops, arches, whorls and triradii to explore the association between palm and CHDs.

Sanchez cascos [5] initiated a study on exploring the association between CHDs and palm print. For analysis, he has considered palm images of 150 patients and of 50 normal. Among these patients, 17 were ASD, 21 were VSD, 23 were Pulmonary valve stenosis (PS), 34 were FT, 18 were Aortic Stenosis (AS), 7 were AC, 12 were Persistent Ductus Arteriosus (PDA) and 18 were other anomalies. He found that palm prints of most of the patients suffering from Fallot's Tetralogy (FT) tends to have larger "atd" angle compared to healthy people. Mahesha and Nagaraju [6] continued the study to test the association between CHD and triradii. In their study, analysis has been made on 44 patients suffering CHD and 362 patients suffering from acquired heart disease. After analysis, they came to know that axial triradius is located at the distal position more in CHD patients than in regular heart disease patients. Desai et al. [7] continued the research considering males and females in his study. He considered 425 female patients and 375 male patients in his study. He has also considered healthy people for comparison. The result of work justifies the conclusions of previous researchers that axial triradius is in a more distal position in CHD patient than in healthy person. In recent years, Khalil and Ahmed [8] and Hashemi et al. [9] have taken up this work. For their analysis, they have considered 150 patients and 300 normals. They come up with the conclusion which was made by previous researchers and the conclusion is axial triradius is in a more distal position in CHD patients than in normal. The "atd" angle at axial triradius for CHD patients presented by the previous researchers is presented in Table 1.

The limitation of previous research works is that they used ink and paper to collect palm prints and palm patterns were recognized through eyes [10]. Required parameters to analyse CHDs were calculated through manual approach using scale and protractor. All these researchers have used a small sample size and provided a conclusion. The majority of these research works have shown that the relationship between the "atd" angle and CHDs is significant. Working with ink palm print and geometrical instruments is tedious and also it is difficult to work with a larger number of samples. Hence the computer-based method is required to identify triradii and to find the "atd" angle. This article is presenting the method which uses digital palm print and Template Matching Algorithm to identify the location of triradii and find the "atd" angle which is used to analyse CHDs.

Table 1 Results of previous research works

| Researcher | CHD | "atd" angle |
|--------------|------------|-------------|
| Sanchez | FT | >46° |
| N Pushpamala | ASD | >46° |
| Brijendra | VSD | >51° |
| Brijendra | FT and CoA | >45° |

2 Material and Methods

The steps used in the identification of triradii using palm images are explained in this section. Figure 1 depicts the steps to pass through for obtaining templates for left, right and axial triradius. A palm print image is provided as input. From this palm image, leftmost triradius, rightmost triradius and axial triradius templates are formed. The templates are in a square or rectangular shape. The mid-point of the template is the actual triradius.

The steps to pass through for locating the triradii are shown in Fig. 2. Palm image and templates have been provided as input. Template matching methods have been used to match the templates on the input palm image.

After finding the locations of triradii, the “atd” angle is evaluated using the procedure shown in Fig. 3. Provide the locations of left, right and axial triradius as input. Two vectors are formed from axial triradius. One towards the leftmost triradius and another towards the rightmost triradius. The angle at axial triradius is calculated and this angle is termed “atd” angle. “atd” angle is the parameter used to analyse the CHDs.

Images have been captured so that patterns in the palm image are clearly visible. The sample palm print image captured for the experiment is presented in Fig. 4. The size of the image taken is 2323×2917 . The templates of the leftmost triradius, rightmost triradius and axial triradius are presented in Fig. 5.

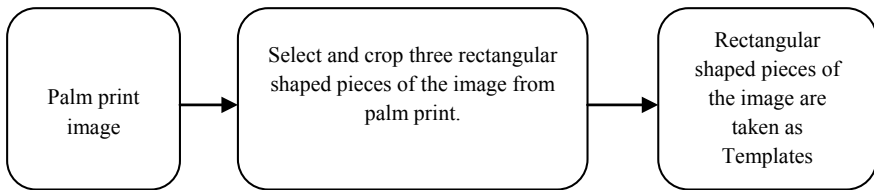


Fig.1 Template formation

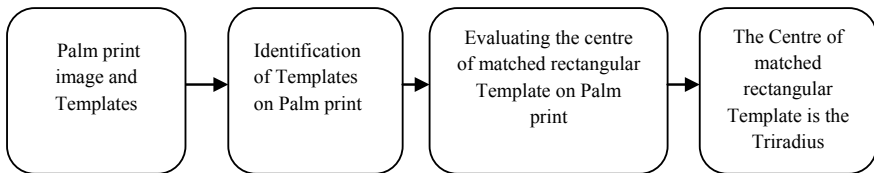


Fig. 2 Triradii identification

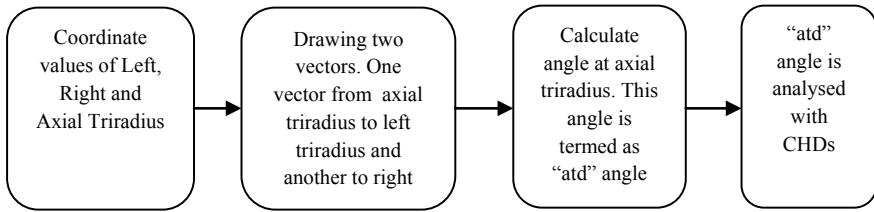


Fig. 3 Evaluation of “atd” angle



Fig. 4 Palm image

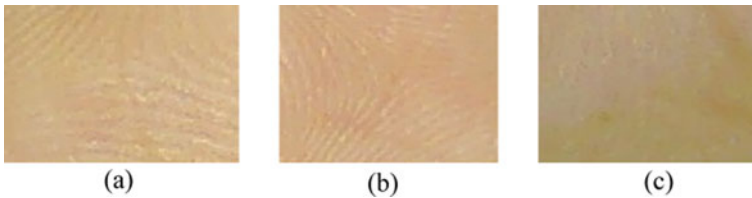


Fig. 5 Patterns of triradii: **a** Leftmost triradius **b** Rightmost triradius **c** Axial triradius

2.1 Template Matching Approach

Template matching [11] has been used to locate parts of an image that matches the template. The steps involved in the algorithm are explained below.

1. Provide palm image and templates of leftmost triradius, rightmost triradius and axial triradius as input.
2. Apply OpenCV function `matchTemplate()` with any of the six methods explained below.
3. Normalize the output.
4. Find the location which has the maximum matching value.

2.2 Template Matching Methods

The matching function passes through the input image, compares the overlaying patches of size $w \times h$ against the template using the specified method. Mathematical expressions for six different matching methods are explained below. “M” represents the input image, “N” represents the template and “R” represents the result. The summation is taken over the template and the image patch: $p' = 0 \dots w - 1$, $q' = 0 \dots h - 1$.

1. Square difference (CV_TM_SQDIFF)

Euclidean distance is calculated and which is squared. It performs subtraction between every pair of pixels and squares the difference. All the squared values are summed together. Since this method matches the squared difference the perfect match results in zero and bad matches results in a large value [12]

$$R(p, q) = \sum_p q' (M(p', q') - N(p + p', q + q'))^2 \quad (1)$$

2. Correlation (CV_TM_CCORR)

This method is based on dot products. Every pair of pixels has been considered. These pairs are multiplied and finally added together. Since this method multiplicatively matches the template the perfect match results in a large value and bad matches results in a small value or zero.

$$R(p, q) = \sum_{p', q'} M(p', q') * N(p + p', q + q') \quad (2)$$

3. Normalized Cross-Correlation (CV_TM_CCORR_NORMED)

In this method, an extreme mismatch results in a score nearer to zero [13]).

$$R(p, q) = \frac{\sum_{p', q'} (M(p', q') * N(p + p', q + q'))}{\sqrt{\sum_{p', q'} M(p', q')^2 * \sum_{p', q'} N(p + p', q + q')^2}} \quad (3)$$

4. Correlation coefficient (CV_TM_CCOEFF)

This method matches a template relative to its mean against the image relative to its mean. Hence, a perfect match results in 1 and a perfect mismatch results in -1 [14]

$$R(p, q) = \sum_{p', q'} (M'(p', q') * N'(p + p', q + q')) \quad (4)$$

where,

$$M'(p', q') = M(p', q') - 1/(w.h) * \sum_{p'', q''} N(p'', q'')$$

$$M'(p + p', q + q') = M(p + p', q + q') - 1/(w.h) * \sum_{p'', q''} N(p + p'', q + q'')$$

5. Normalized Square Difference (CV_TM_SQDIFF_NORMED)

In this method, a perfect match results in a value equal to zero.

$$R(p, q) = \frac{\sum_{p', q'} (M(p', q') - N(p + p', q + q'))}{\sqrt{\sum_{p', q'} M(p', q')^2 * \sum_{p', q'} N(p + p', q + q')^2}} \quad (5)$$

5. Normalized Correlation Coefficient (CV_TM_CCOCOEFF_NORMED)

This method returns a positive value for a match and a negative value for a mismatch [14]

$$R(p, q) = \frac{\sum_{p', q'} (M'(p', q') * N'(p + p', q + q'))}{\sqrt{\sum_{p', q'} M'(p', q')^2 * \sum_{p', q'} N'(p + p', q + q')^2}} \quad (6)$$

2.3 Finding the Position of Triradii on Palm

Let $(x1, y1)$ be the coordinate values that represent the position of the upper left corner of the matched image patch and $(x2, y2)$ be the coordinates of the lower right corner of the matched image patch. Coordinate values of the centre of the matched image patch are computed using Eqs. (7) and (8).

$$centr_x = \frac{x1 + x2}{2} \quad (7)$$

$$centr_y = \frac{y1 + y2}{2} \quad (8)$$

where $(centr_x, centr_y)$ represents x and y coordinates of the centre of the rectangle. This centre represents the position of triradius.

2.4 Finding the “atd” Angle(θ)

The angle at the axial triradius is called “atd” angle. “atd” angle is the parameter to be analysed to detect CHDs. Here, three triradii are identified. The leftmost triradius, rightmost triradius and axial triradius are represented as “a”, “d” and “t” respectively as shown in Fig. 6.

Two vectors have been drawn from axial triradius. One vector in the direction of the leftmost triradius and another in the direction of the rightmost triradius. The vectors are named “ta” and “td”. The “atd” angle which is represented as θ is obtained as follows [15].

$$\begin{aligned} \text{ta} \cdot \text{td} &= |\text{ta}| \cdot |\text{td}| \cdot \cos\theta \\ \cos\theta &= (\text{ta} \cdot \text{td}) / (|\text{ta}| \cdot |\text{td}|) \\ \theta &= \cos^{-1}((\text{ta} \cdot \text{td}) / (|\text{ta}| \cdot |\text{td}|)) \end{aligned} \quad (9)$$

3 Results

Templates for the leftmost, rightmost and axial triradii are formed from the input palm image which is taken for the study. These templates are matched against input palm image. Six template matching methods have been applied. Matched part of the image is rectangular. The mid-point of the rectangle is the triradius to be taken for analysis. After locating all the three triradii, the angle at axial triradius (“atd”) is calculated with Eq. (9). The result of six template matching methods for locating the triradius is explained below.

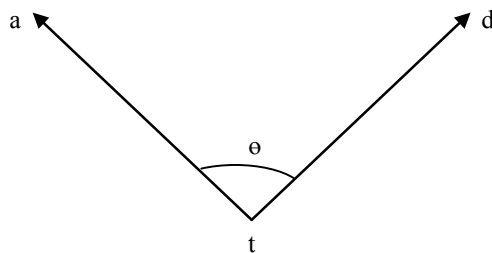


Fig. 6 “atd” angle

3.1 Identification of Left Triradius, Right Triradius and Axial Triradius on Palm Image

The left, right and axial triradii were identified using six template matching methods and the results are as shown in Fig. 7, 8 and 9 respectively. Results of Eqs. (1)–(6) are shown as SQDIFF, CCORR, CCORR_NORMED, CCOEFF, SQDIFF_NORMED and CCOEFF_NORMED respectively. Part of the image with the highest matching probability is marked in a rectangular shape. The centre of the rectangle is identified using Eqs. (7) and (8) and it is known as triradius.

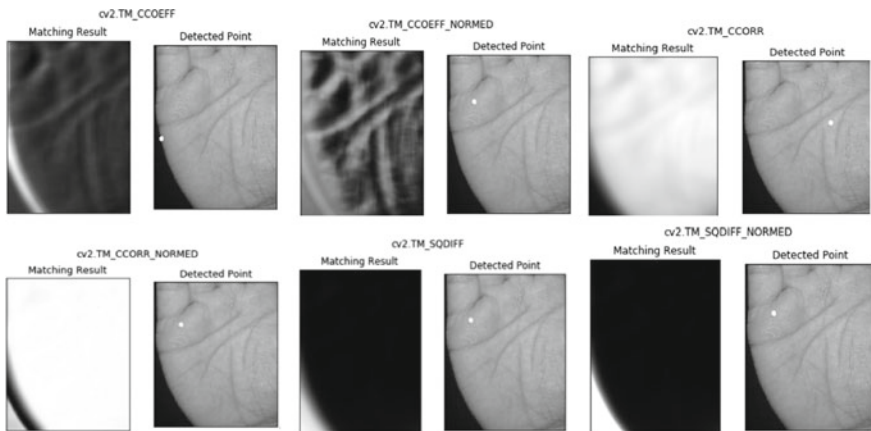


Fig. 7 Identification of left triradius on palm print using six different matching methods

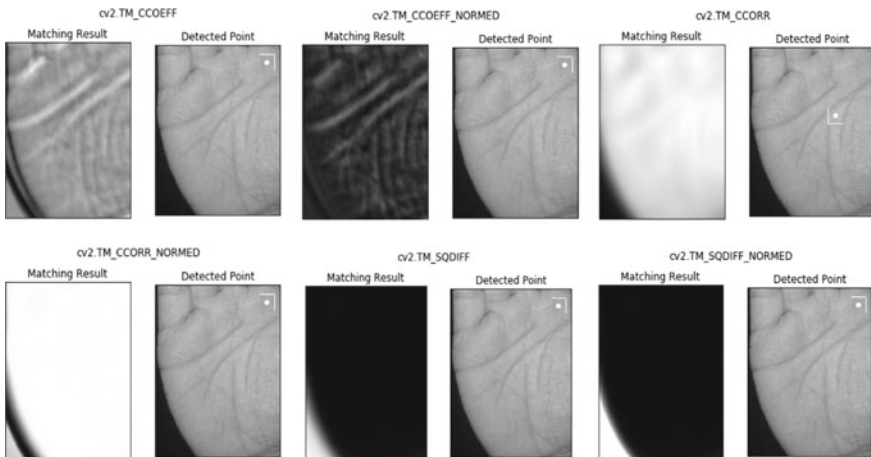


Fig. 8 Identification of right triradius on palm print using six different matching methods

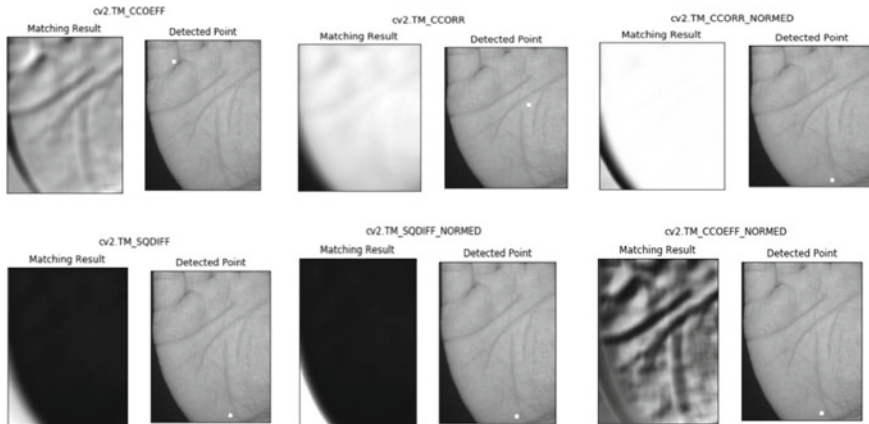


Fig. 9 Identification of axial triradius on palm print using six different matching methods

3.2 Angle at Axial Triradius

After identifying all the three triradii of the palm, the “atd” angle is calculated using Eq. (9). Locations of all the three triradii on the palm image for six template matching methods along with angle at axial triradius are as shown in Fig. 10.

The coordinate values of locations of left, right and axial triradii, and the “atd” angle resulted from six template matching methods have been presented in Table 2.

All the six template matching methods have been applied on 450 palm images. The resolution of each image is 2323×2917 . The result obtained is presented in Table 3.

The results presented in Table 3 shows that the template matching methods namely CCOEFF_NORMED, CCORR_NORMED, SQDIFF and SQDIFF_NORMED achieving the accuracy of 99.55%, 97.55%, 98.22% and 97.55% respectively.

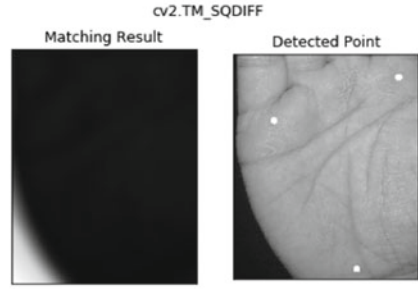
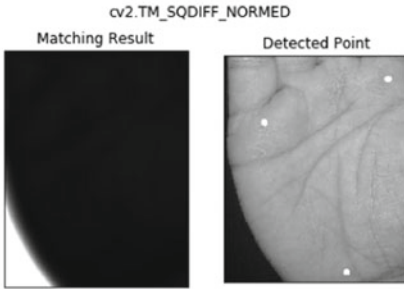
3.3 Association Between the “atd” Angle and CHDs

To test the association between the “atd” angle and CHDs, the Chi-square test has been adopted. The palm images were captured from CHD patients suffering from different CHDs and from normal. The number of palm images collected from patients and healthy people is given in Table 4. The images have been collected from people of age group between 25 and 50 years which includes both male and female.

The template-based matching method is applied to all the collected palm images to find the “atd” angle. The range of angles obtained for CHD patients and healthy people are given in Table 5.

Position of Rightmost Triradius (954, 137)
Position of Leftmost Triradius (236, 393)
Position of Axial Triradius (713, 1274)

Position of Rightmost Triradius (954, 137)
Position of Leftmost Triradius (236, 393)
Position of Axial Triradius (713, 1274)

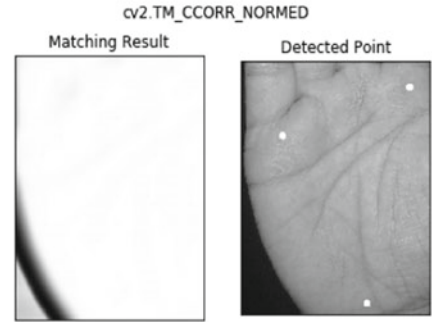
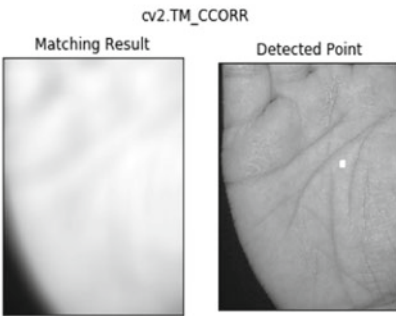


Anlgle at Axial Triradius: 40.399816632348255

Anlgle at Axial Triradius: 40.399816632348255

Position of Rightmost Triradius (734, 540)
Position of Leftmost Triradius (732, 542)
Position of Axial Triradius (733, 546)

Position of Rightmost Triradius (954, 137)
Position of Leftmost Triradius (236, 393)
Position of Axial Triradius (713, 1274)

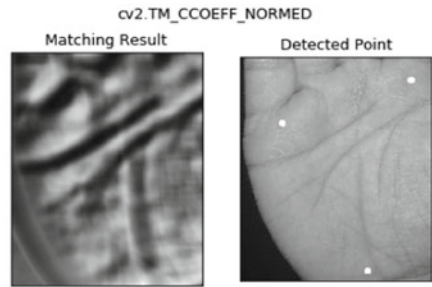
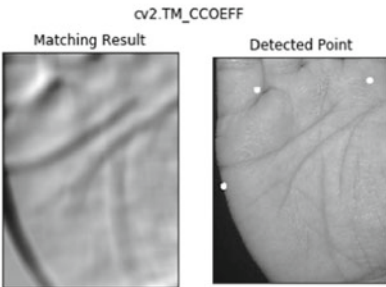


Anlgle at Axial Triradius: 23.49856567595208

Anlgle at Axial Triradius: 40.399816632348255

Position of Rightmost Triradius (954, 137)
Position of Leftmost Triradius (67, 763)
Position of Axial Triradius (271, 192)

Position of Rightmost Triradius (954, 137)
Position of Leftmost Triradius (236, 393)
Position of Axial Triradius (713, 1274)



Anlgle at Axial Triradius: 114.26410814353531

Anlgle at Axial Triradius: 40.399816632348255

Fig. 10 Position of all three triradii along with “atd” angle

Table 2 Coordinate values of position of triradii and angle at axial triradius

| Template Matching Method | Coordinate values of position of triradii | | | Angle at axial triradius |
|--------------------------|---|----------------|-----------------|--------------------------|
| | Right triradius | Left triradius | Axial triradius | |
| CCOEFF | (954,137) | (67,763) | (271,192) | 114.2641 |
| CCOEFF_NORMED | (954, 137) | (236, 393) | (713, 1274) | 40.3998 |
| CCORR | (734, 540) | (732, 542) | (733, 546) | 23.4985 |
| CCORR_NORMED | (954, 137) | (236, 393) | (713, 1274) | 40.3998 |
| TM_SQDIFF | (954, 137) | (236, 393) | (713, 1274) | 40.3998 |
| SQDIFF_NORMED | (954, 137) | (236, 393) | (713, 1274) | 40.3998 |

Table 3 Accuracy table

| Template matching method | Number of palm images | Number of palm images for which all the three triradii identified correctly | Accuracy (%) |
|--------------------------|-----------------------|---|--------------|
| CCOEFF | 450 | 28 | 6.22 |
| CCOEFF_NORMED | 450 | 448 | 99.55 |
| CCORR | 450 | 17 | 3.77 |
| CCORR_NORMED | 450 | 439 | 97.55 |
| SQDIFF | 450 | 442 | 98.22 |
| SQDIFF_NORMED | 450 | 439 | 97.55 |

Table 4 Number of palm images collected

| Type | Number of palm images |
|------|-----------------------|
| ASD | 108 |
| VSD | 134 |
| CoA | 69 |
| TOF | 213 |

From the results shown in Table 5, it is found that the majority of CHD patients have an angle at axial triradius above 45° and the majority of healthy people have an angle below 45°. Hence, the chi-square test has been applied as shown in Table 6.

The P-value obtained for the CHDs is <0.001 and it suggests that there exists a strong association between the “atd” angle and CHDs.

Table 5 Range of “atd” angle

| Range | ASD | VSD | CoA | TOF | Healthy |
|-------|-----|-----|-----|-----|---------|
| <5 | 0 | 0 | 0 | 0 | 0 |
| 5–10 | 0 | 0 | 0 | 0 | 0 |
| 11–15 | 0 | 0 | 0 | 0 | 0 |
| 16–20 | 0 | 0 | 0 | 0 | 0 |
| 21–25 | 0 | 0 | 0 | 0 | 0 |
| 26–30 | 0 | 1 | 0 | 2 | 0 |
| 31–35 | 8 | 9 | 3 | 13 | 18 |
| 36–40 | 12 | 18 | 7 | 18 | 98 |
| 41–45 | 8 | 6 | 0 | 32 | 42 |
| >45 | 80 | 100 | 59 | 148 | 9 |

Table 6 Confusion matrix

| | Number people with angle >45° | Number people with angle <45° |
|---------|-------------------------------|-------------------------------|
| ASD | 80 | 28 |
| VSD | 100 | 34 |
| CoA | 59 | 10 |
| TOF | 148 | 65 |
| Healthy | 9 | 158 |

4 Conclusion

This paper presents an image processing based method to identify the triradii and to find the “atd” angle. In the current work, six template matching methods are used for matching the templates of left triradius, right triradius and axial triradius. The six methods are Square difference, Correlation, Normalized cross correlation, Correlation coefficient, Normalized square difference and Normalized correlation coefficient. Among these methods, Square difference, Normalized cross correlation, Normalized square difference and Normalized correlation coefficient are correctly identifying the triradii on the palm. The algorithm which has been developed and implemented to evaluate “atd” angle is capable of providing the desired results as expected. Using this approach research on finding the relationship between the position of axial triradius and CHDs can be conducted on a large sample size easily and can come to a better conclusion which was not possible in the previous research works where the sample size is small and triradii are manually identified through eyes. The experiment has been conducted by taking a set of palm images from CHD patients suffering from VSD, ASD, CoA and TOF. The “atd” angle has been evaluated using the proposed method. From the results obtained, it is found that there exists a strong association between the “atd” angle and CHDs.

References

1. Cascos AS (1965) Palmprint pattern in congenital heart disease. *Brit Heart J* 27(4):599–603
2. Takashina T, Yorifuji S (1966) Palmar dermatoglyphics in heart disease. *JAMA* 197(9):689–692
3. David TJ (1981) Dermatoglyphs in congenital heart disease. *J Med Genet* 18(5):344–349
4. Pushpamala N (2015) Role of dermatoglyphics in atrial septal defect in children. *MRIMS J Health Sci* 3(3):218–222
5. Singh B, Gupta R, Agarwal D, Garg R, Katri S (2016) Dermatoglyphic's in congenital cardiac disease. *Acta Med Iran* 53(2):119–123
6. Mahesha Y, Nagaraju C (2020) A Literature review on analysis of palm patterns to detect congenital heart diseases. *Biomed Eng Appl Basis Commun* 32(2):2050012
7. Desai BK, Pandya M, Potdar DB (2013) Comparison of various template matching techniques for face recognition. *Int J Eng Res Dev* 8(10):16–18
8. Khalil M, Ahmed I (2015) Quick techniques for template matching by normalized cross-correlation method. *Brit J Math Comput Sci* 11(3):1–9
9. Hashemi NS, Aghdam RB, Ghiasi ASB, Fatemi P (2016) Template matching advances and applications in image analysis. *Am Sci Res J Eng Technol Sci* 26(3):91–108
10. Gunawan H, Neswan O, Setya-Budhi W (2005) A formula for angles between subspaces of inner product spaces. *Beitrage zur Algebra und Geometrie Contrib Algebr Geom* 46(2):311–320
11. Perveen N, Kumar D, Bhardwaj I (2013) An overview on template matching methodologies and its applications. *IJRCCT* 2(10):988
12. Khan SM, Drury NE, Stickly J, Barron DJ, Braw WJ, Jones TJ, Anderson RH, Crucean A (2019) Tetralogy of fallot: morphological variations and implications for surgical repair. *Eur J Cardiothorac Surg* 56(1):101–109
13. Celermajer DS (2018) Atrial septal defects: Even simple congenital heart diseases can be complicated. *Eur Heart J* 39(12):999–1001
14. Cox K et al (2020) The natural and unnatural history of ventricular septal defects presenting in infancy: An echocardiography-based review. *J Am Soc Echocardiogr* 33(6):763–770
15. Rogers C, Clawson RE (2019) Coarctation of the aorta. *J Am Acad Physician Assist* 32(6):46–47

Forecasting Flash Floods with Optimized Adaptive Neuro-Fuzzy Inference System and Internet of Things



M. Pushpa Rani, Bashiru Aremu, and Xavier Fernando

Abstract Over the past few decades, global warming and climate change have resulted in unpredictable floods in many regions of the world, which has the potential to cause a wide range of catastrophes. The primary objective of this work is to design a system for early flood prediction using soft computing and the Internet of Things (IoT). Predicting heavy rainfall with extreme precision is critical for saving people from flooding and minimizing property damage. There are numerous methods for predicting rainfall available today, but all of them are worthless due to drastic climate change. This study proposes a hybridized adaptive neuro-fuzzy inference system to reduce the mistakes in rainfall forecasts caused by climate change. ANFIS has been hybridized by fire fly algorithm. Weather big data was collected from the Chennai metrological region from 2010 to 2020 and analyzed using an upgraded adaptive neuro-fuzzy inference system. Additionally, IoT technology is being used to automate flood alarms and monitor flood parameters regularly. Finally, the proposed method is implemented experimentally to demonstrate the proposed early flood prediction model's accuracy.

Keywords Early flood forecasting · IoT · Adaptive neuro-fuzzy inference system · Fire fly algorithm · Disaster management

M. Pushpa Rani (✉)

Department of Computer Science, Mother Teresa Women's University, Kodaikanal, India
e-mail: drpushpa.mtwu@gmail.com

B. Aremu

Crown University Intl.Chartered Inc., Santa Cruz in Argentina, Americas, Ghana
e-mail: vc@crowintl.education

X. Fernando

Ryerson Communication Lab, Ryerson University, Toronto, Canada
e-mail: fernando@ee.ryerson.ca

1 Introduction

India is one of the world's most flood-prone countries. Floods in India are currently caused by a variety of phenomena, including severe rainfall. There were fatalities and economic losses as a result of flash floods in India's main cities of Mumbai, Chennai, and Bangalore in 2005 and 2015 [1–3]. According to the latest National Flood Commission data, over 40 million ha of land in India are at risk of flooding. Therefore, an efficient early flood warning system must be developed to address the catastrophic floods caused by climate change in many parts of the world. The Indian Ocean's warming is a strong sign of climate change and is the primary cause of cloudbursts in cities like Chennai. A 'cloudburst' is a location that receives more than 100 mm of rain in one hour. Per square kilometer, 100 mm of rainfall equates to 100,000 metric tons of rain. A perfect evidence of this is the torrential rains that fell on December 31, 2021, which the Tamil Nadu meteorological department could not forecast.

Several flood prediction models have been developed using artificial neural networks (ANNs), although they frequently fail to perform well in decision-making due to inefficient learning parameter selection, decision making and feature selection [4, 5]. Accordingly, ANFIS has been employed in this study to increase the efficiency of decision-making. Besides, to enhance the ANFIS's performance, which is optimised using the fire fly algorithm. Existing flood warning systems have been developed in the past using IoT and artificial intelligence, however the majority of their study focuses on temperature, humidity, and rainfall. Also no sensing device is used to monitor the cloud so it fails to predict floods caused by other climatic conditions like cloud burst. In this study, a cloud cover sensor was utilized to monitor the formation of rain clouds. To reduce the false negative ratio of the proposed forecast model, the model is trained through different cloud coverage levels that caused the downpour.

This article is structured as follows: Sect. 2 reviews in detail the flood warning systems developed by IoT and AI that are similar to this research. The modules and procedures of the proposed early flood warning system developed with ANFIS and IoT technologies are discussed in detail in Sect. 3. Section 4 evaluates the proposed method's practical applicability using recently developed flood warning systems. Finally, Sect. 5 summarizes the proposed study.

2 Literature Review

Dong et al. [6] recommended a strategy based on hybrid deep learning and FastGRNN-FCN for predicting urban floods using network sensor data. Three past flood occurrences provided data for training and validation of channel sensors. The flood data gathered in this manner was categorized into multivariate time series and

used as input to the model. Each input comprises six variables containing information about the channel sensor and its predecessor and successor. The precision-recall curve and F-measure are used to determine the optimal set of model parameters. After 100 iterations of experimenting with various weights and thresholds, the best model with a weight of 1 and a critical threshold of 0.59 is obtained. The technique achieved 97.8% test accuracy and 0.792% F-measure accuracy, respectively. The scheme's data was first separated into multivariate time series and then provided as input. By developing a dashboard, the data were shown to aid in the identification of high-risk locations. The technique can be improved by increasing the size of the training data collection. Additionally, infrastructural requirements might be taken into account.

Khan et al. [7] demonstrated a method for forecasting flash floods using multi-modal sensing and artificial intelligence-based decision making. Flash floods are investigated using a single device that combines passive infrared (PIR), water level sensor, ultrasonic sensor, temperature sensor, pressure sensor, and altimeter sensor. For data processing and training, a modified multi-layer feed forward neural network optimized using the particle swarm optimization algorithm. The approach is divided into two phases: data compilation and data classification. The scheme's data is derived from the Kund Malir seashore in Pakistan. The hybrid algorithm is then used to filter, classify, and optimize the collected data. After that, the data is arranged into tables to eliminate redundancy, missing entries, and repetitive values. The data is then normalized to remove any redundant information. The system's shortcomings include the requirement for additional processing resources for the collecting of real-time data. MATLAB's data acquisition procedure was unable to compute real-time data. Additionally, the data receiving rate for real-time data was quite poor.

Moishin et al. [8] provided a method for constructing and evaluating flood forecasting by merging the predictive advantages of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) Networks in a hybrid deep learning (ConvLSTM) algorithm. Developing a system capable of providing early warning of impending flooding is a significant undertaking. Additionally, early prediction requires a thorough understanding of many technologies. The study makes use of precipitation data from the Fiji Meteorological Services. The Flood Index (IF) is used to quantify the rate of water depletion over time. IF is used to determine the duration, severity, and intensity of any flood condition in the flood monitoring system. This method augments IF with antecedent and dynamic rainfall data in order to forecast the following day's IF value. Preprocessing data entails utilising the calendar mean to complete missing data points. The method's disadvantage is that it uses only two input parameters for prediction. The approach cannot be used to anticipate floods in the absence of an established IF. Another constraint necessitates the development of a user-friendly interface for implementing the technique.

Miau et al. [9] recommended a deep learning-based model for predicting the water level flood phenomenon of a river in Taiwan. CNN is used to extract features from time series data. The GRU model is used to obtain long term dependant data for time series. The model incorporates the advantages of the CNN and GRU models.

Conv-GRU connects the CNN model's output to the GRU model's input. Conv-GRU neural network is used to extract complex aspects of river water level. Conv-GRU is used to detect time series errors using open data. Pre-processing data entails resampling and filling in missing values. The wavelet transform is used to remove noise. Each model's ideal parameters were found using the size of the input and output sequences, the number of neurons, and the number of hidden layers. The model was then supplemented with training data for the aim of training and learning. Prediction was accomplished with the addition of test data for back testing. The algorithm can be extended to forecast floods in other Taiwanese rivers.

Souza et al. [10] recommended a low-cost flood warning system to prevent flood-related damages and losses in metropolitan areas. The technology monitors the river and runway water levels and sends an alarm by phone call or SMS message whenever a hazardous condition is detected. The system makes use of electronic components to monitor rivers and weather forecasting systems. Three monitoring sensors are included in the system: a pressure sensor (used to determine the water level), an accelerometer (which detects vandalism and theft), and a fluvial plution detector. ZigBee is used to transmit data between nodes. ZigBee is used to transmit data and communicate between network nodes. The results are kept in the cloud. The reference to the water level is determined using an ultrasonic sensor. Enhancing the system necessitates safeguarding it against damage and theft. Additionally, the system can provide alternate routes to the populace from flooded areas.

3 Proposed Methodology

Figure 1 illustrates the overall architecture of the proposed early flood warning system based on IoT and adaptive neuro-fuzzy inference.

This is a three-tier architecture; the first module of the framework makes use of IoT weather monitoring devices, whereas the second module integrates a methodology for weather big data analysis (early weather prediction). A number of sensors for monitoring the weather in the IoT module have been installed at various points across Chennai, including the bank of the river, the bank of the dam, and at various locations throughout the city. Communication between these IoT module and early weather prediction module established by the use of Wifi and cloud computing technology. Additionally, high security AMQP (advanced message queue protocol) is used for Internet of Things (IoT) message exchanges, which accelerates data exchange between the IoT module and data analysis module. The data analysis module plays a very important role in this early flood warning system. To enhance early flood warnings, adaptive neuro-fuzzy inference has been applied. The data analysis module conducts a variety of actions to help prevent catastrophic flooding losses. To begin, this module receives weather data through Wifi and AMQP from various weather sensors deployed around Chennai's flood-prone areas. This gathered data is fed into the data analysis module as input data. Additionally, the cloud

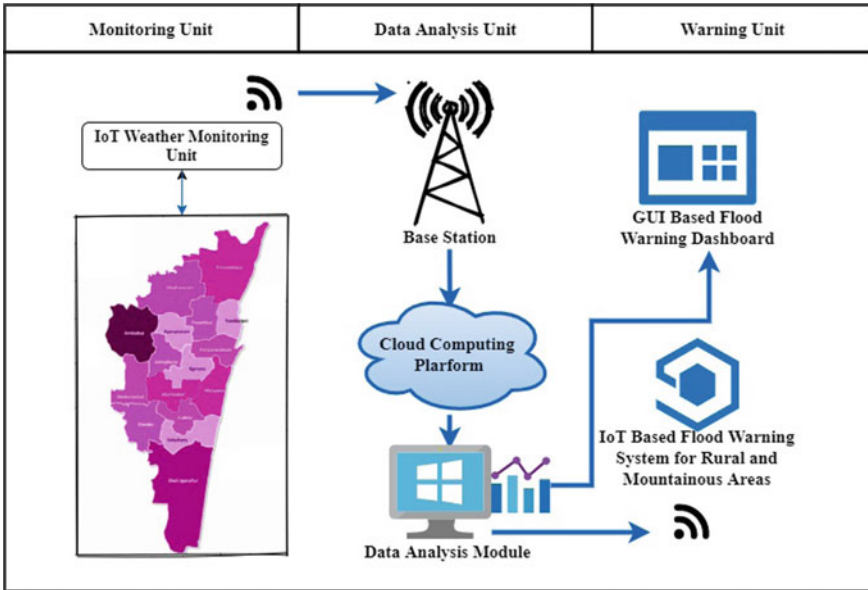


Fig. 1 General components of the proposed early flood warning system

platform is employed to store this data analysis module and the associated wealth data. Figure 2 depicts the data analysis module’s class diagram.

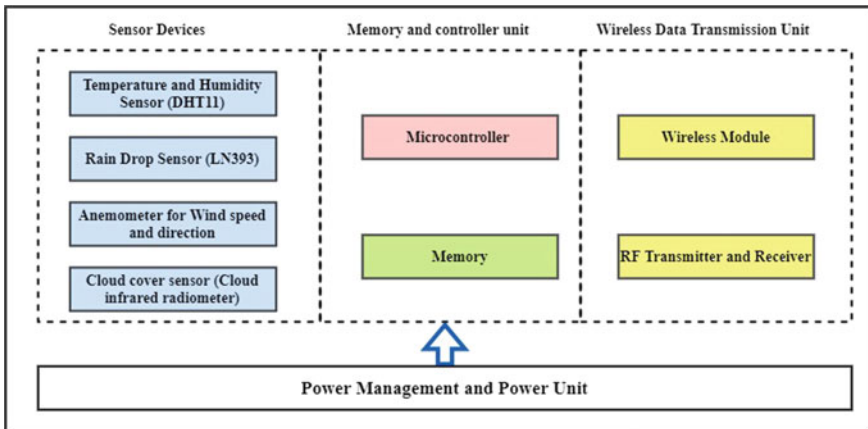


Fig. 2 The IoT units’s device and data communication details

Table 1: Weather Parameters details.

| Attribute name | Attribute description | Attribute data type |
|-----------------------|--|---------------------|
| Year | This attribute indicates the year in which the climatological variable was gathered. | Numeric |
| Month | This element specifies the month during which the climatological data was collected. | Numeric |
| Date and time | It specifies the date and time of the metrological parameter collection. | Numeric |
| Maximum temperature | The highest temperature observed during the day is represented by this attribute. | Numeric |
| Minimum temperature | The lowest temperature observed during the day is represented by this attribute. | Numeric |
| Maximum humidity | The highest humidity observed during the day is represented by this attribute. | Numeric |
| Minimum humidity | The lowest humidity observed during the day is represented by this attribute. | Numeric |
| Rain fall | This is the amount of daily precipitation that occurs. | Nominal |
| Cloud coerage details | This attribute is used to describe the details of a rain cloud. | Nominal |

3.1 Study area and dataset details

The Chennai metro area, which covers an area of 1.18 km² and has a population of 8,653,531, is India's fourth-most populous city, behind New Delhi and Mumbai. As a result of sudden climatic changes, Chennai is vulnerable to excessive flooding. Since this, Chennai was chosen as the study area for this research. This dataset contains data on the weather in Chennai from 2014 to 2020. All of this weather information is sourced from the Regional Meteorological Center Nungambakkam in Chennai. The dataset's attributes details are summarized in Table 1 [16].

3.2 IoT Unit and Monitoring Platform

Figure 2 illustrates the IoT weather monitoring module in greater detail. The Internet of Things unit is composed of four major components: the IoT weather sensing unit, the memory and microcontroller unit, the data transmission unit, and the power unit. The IoT weather sensing unit monitors five essential aspects that contribute to flooding: temperature and humidity sensors (DHT11), rain drop sensors (LN393), an anemometer for wind speed and direction, cloud cover sensors (cloud infrared radiometer), and river dam level monitoring (Ultrasonic Sensor). The Raspberry Pi was employed as the microcontroller in this study. The specifics of the weather to be observed are stored in the device's memory (monitoring time intervals, weather

parameter threshold levels etc.) by the Arduino C software, which is written in embedded C program. In this study, the ESP8266 module is used for wireless data communication. This IoT weather monitoring system’s power supply is provided by a rechargeable battery.

3.3 Weather Analysis Platform for Early Flood Prediction

This section will discuss the proposed ANFIS model for early flood prediction. ANFIS is an artificial intelligence decision-making system based on fuzzy inference and artificial neural networks (ANN) [11, 12]. ANFIS’s architecture is primarily composed of five levels, as illustrated in Fig. 3. The ANFIS model forecasts weather utilizing IF-THEN fuzzy rules, and their performance is based on these rules. The number of rules varies depending on the neurons numbers that are used as input. Maximum and minimum temperature, maximum and minimum humidity, wind speed and direction, cloudiness, rain level, dam water level, and river water level are given as input to the ANFIS. The network architecture of ANFIS for flood prediction is depicted in Fig. 3, which includes five levels as specified in the architecture. The first layer contains the values for each weather parameter’s inputs and weights. The second layer is composed of the membership functions for the seven weather parameters stated previously. The third layer, which comprises if-then fuzzy rules, is used to make decisions. The fourth layer comprises the fuzzy rules’ output. The final layer plots the flood forecast results. The flood prediction results are displayed by final layer; if the likelihood of a flood is high, the output is 1, if it is extremely low,

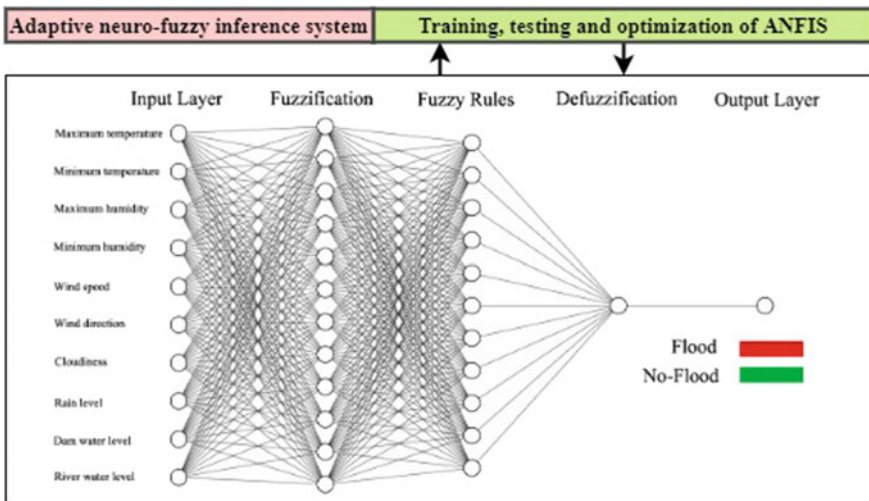


Fig. 3 ANFIS’s communication processes on training, validation, and optimization

the output value of an ANFIS is 0, and values between 0 and 1 indicate a moderate probability of a flood.

Layer 1 performs the fuzzification process by utilising the membership function and input data from IoT wireless sensors. The Gaussian membership function (GMF) employed in this study is computed using the formula below.

$$L_i^1 = \mu_{A_i} = \exp\left[-\frac{(x - c)^2}{\sigma^2}\right] \quad (1)$$

where μ is the Gaussian membership function (GMF), signifies the GMF's width, and c is the GMF's middle value.

Layer 2 is made up of if-then-else fuzzy rules. It defines the rules for deciding the size of the input parameters. It defines the rules for determining what to do when the input parameters are changed.

$$L_i^2 = \omega_i = \mu_{A_i}(x_1) \times \mu_{B_i}(x_2), \forall i = 1, 2, 3, 5 \dots n \quad (2)$$

The fuzzy rules are normalised using Layer 3, which employs Formula 1.

$$L_i^3 = \bar{\omega}_i = \frac{\omega_i}{\sum_{j=1}^1 \omega_j} \quad (3)$$

The decision making process takes place through Layer 4. The following formula is used to make the decision.

$$L_i^4 = \omega_i f_i = \bar{\omega}_i (P_i x + Q_i y + r_i) \quad (4)$$

De-Fuzzification is performed in layer 5 using the output from the fourth layer. This layer establishes the likelihood of floods.

3.4 ANFIS Optimization

In the field of optimization, the fire fly algorithm is the most extensively utilized technique [13, 14]. Yang invented this optimization approach in 2008. It was designed in accordance with the mating habits of fireflies. It has been enhanced in numerous ways over the last decades, including improved versions of the fire fly algorithm,

the standard fire fly algorithm, and the hybrid fire fly algorithm. According to the firefly algorithm, the light intensity emitted by the fire fly can be seen by using the following formula to the fire fly at a distance of k .

$$L(k) = \frac{L_0}{K^2} \quad (5)$$

where L_0 is the luminance emitted by the light source (fireflies). If γ is the medium's absorption coefficient, then Formula (6) gives the light intensity, I , at distance K .

$$I = L_0 \exp(-\gamma K^2) \quad (6)$$

where K denotes the range between both the illumination and the illumination measurement point. This luminance can be related to the attractions of fireflies within the FA, as defined by Formula (7).

$$\delta = \delta_0 \exp(-\gamma K^n) (n \geq 1) \quad (7)$$

where δ_0 denotes the attractiveness of the fireflies at a distance of $K = 0$. The Euclidean distance between the two fireflies f_i and f_j is defined by Formula (8).

$$K_{ij} = |f_i - f_j| = \sqrt{\sum_m^d (f_i, m - f_j, m)^2} \quad (8)$$

The location of the firefly can be altered for each iteration using the following Formula (9).

$$F_i = f_i + \delta_0 \exp(-\gamma K_{ij}^2) (f_i - f_j) + \varphi \omega \quad (9)$$

where φ denotes a random integer produced using the Distribution function and ω denotes the randomization value. Algorithm 1 is used to optimize the ANFIS process.

Algorithm1: ANFIS's optimal learning parameters selection using Fire-Fly algorithm

Step1: Initialize the ANFIS learning parameters (learning rate, network weight etc).

Step2: Calculate the learning value and total network weight of the ANFIS (Fitness function should be used to evaluate the firefly).

Step3: Determine the distance between each ANFIS learning parameters using the best-fitting equation. (Determine the distance between each firefly using the best-fitting equation).

Step4: Equation 9 is used to update the position of the ANFIS learning parameters (Equation 9 is used to update the position of the fireflies).

Step5: Determine the fitness function's value.

Step6: Rank and sort the ANFIS learning parameters according to the results of the fitness function (Rank and sort the fireflies according to the results of the fitness function).

If (get the optimal ANFIS learning parameters)

{
Return to the beginning of the loop
}

Stop the iteration after selecting the optimal ANFIS learning parameters.

END

4 Results and Discussion

The Weather data set is classified into four categories based on the amount of rain that falls: light rain (when the rainfall intensity is <2.6 mm/h), moderate rain (when the rainfall intensity is between 2.6 and 7.5 mm/h), heavy rain (when the rainfall intensity is >7.5 mm/h), and severe rain (when the rainfall intensity is >50 mm/h). The presented FA-ANFIS is trained, tested, and optimized using the K fold cross validation approach. In this study, the weather data set is separated into k subsets based on class attribute, for example, $K = 40$. FA-ANFIS is trained using K-10 samples and tested using the remaining ten. The optimization of FA-ANFIS is conducted several times to verify that all of the data samples were used in the training and testing of FA-ANFIS.

Additionally, in this research, ANFIS's critical training variables, such as learning rate and procedure for updating ANN weights, are replaced with FA optimization rather than backpropagation. In this study, the most vital AI efficiency validating metrics used to demonstrate the proposed FA-ANFIS's training efficiency including

Mean Absolute Error (MAE), Mean Square Error (MSE), and Root Mean Square Error (RMSE) [15].

Mean Absolute Error returns AI model training-related errors with error values ranging from 0 to 1. If the error values returned by MAE are close to zero, the prediction model is highly efficient. MAE helps to measure the difference between actual and predicted results generated by AI-based prediction models. MAE is calculated using the formula below. In Formula 10, (\widehat{y}_i) denotes the predicted value by FA-ANFIS, while y_i denotes the actual probability of flooding.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \widehat{y}_i| \quad (10)$$

MSE denotes the mean square deviation between both the proposed FA- ANFIS's forecasting result and the actual flood risk. The Eq. 11 is used to calculate the MSE values for the recommended FA-ANFIS and existing AI-based flood forecasting methods.

$$MSE = \frac{1}{N} \sum_{i=1}^N |y_i - \widehat{y}_i|^2 \quad (11)$$

The root mean square error (RMSE) is defined as the square root of all MSE values between the risk prediction result of the proposed FA-ANFIS-based flood forecasting system and the real risk of flooding. The RMSE values for the recommended FA-ANFIS and existing AI-based flood forecasting methods are calculated using the Formula 12.

$$RSME = \sqrt{\frac{1}{N} \sum_{i=1}^N |y_i - \widehat{y}_i|^2} \quad (12)$$

To train the FA-ANFIS models, the input meteorological data set is split into three combinations: 60, 70, and 80%. The train and test datasets are chosen at randomly from the meteorological database.

The MAE, MSE, and RMSE values of the suggested FA-ANFIS-based flood prediction model are all close to zero, indicating its effectiveness, but the MAE, MSE, and RMSE values of the traditional AI based flood prediction models are all quite high. Comparative analysis demonstrates that the FA-ANFIS model outperforms standard AI-based flood prediction approaches in terms of MAE, MSE, and RMSE (Figs. 4, 5 and 6). There are two main reasons for the decline in MAE, MSE, and RMSE in this research: first, this study examined the cloud coverage details that contributed to the sudden flood, and then used the FA to optimise the learning variables rather than the more typical back propagation. This gives a good improvement in prediction accuracy also which is discussed in the following section. The

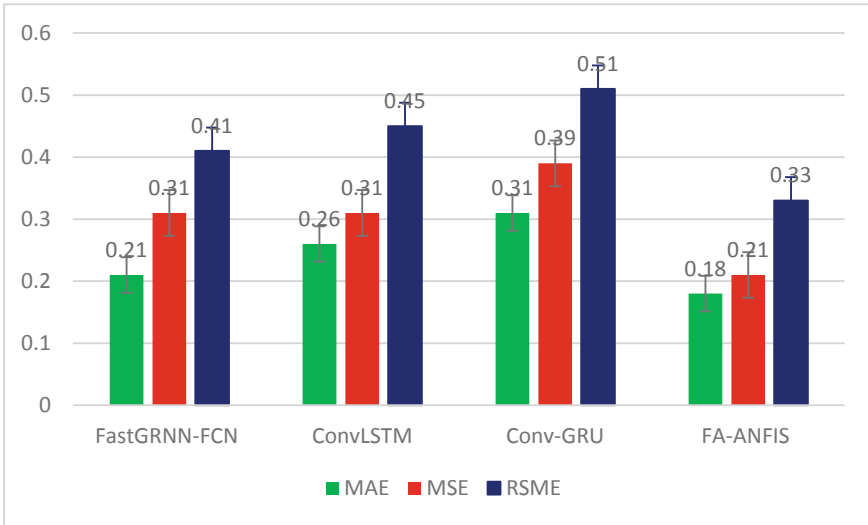


Fig. 4 Training errors comparison results (MAE, MSE, RMSE) with 60% of training data

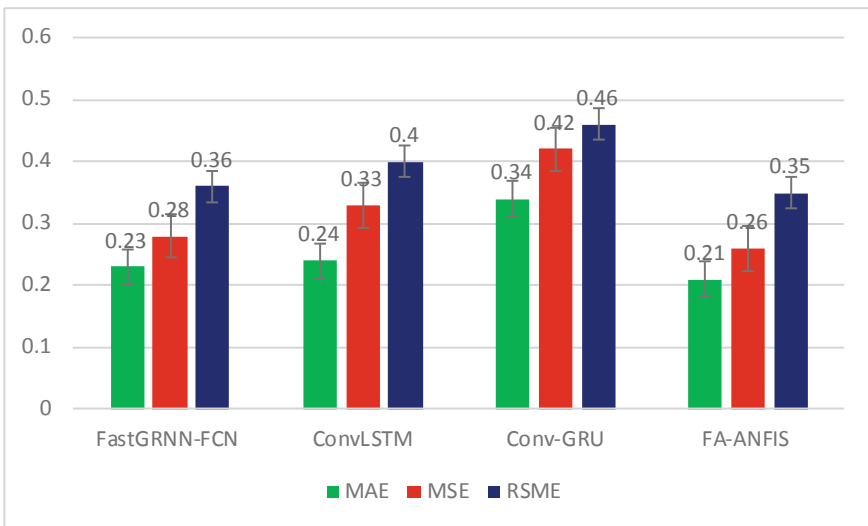


Fig. 5 Training errors comparison results (MAE, MSE, RMSE) with 70% of training data

experimental results demonstrate that, by optimising the learning rate and training procedure, better prediction outcomes can be attained.

This section discusses the experimental findings associated with the proposed FA-ANFIS-based flood prediction model. Additionally, the Accuracy, Precision, Recall and F1-score are employed to assess the proposed FA-ANFIS-based flood prediction

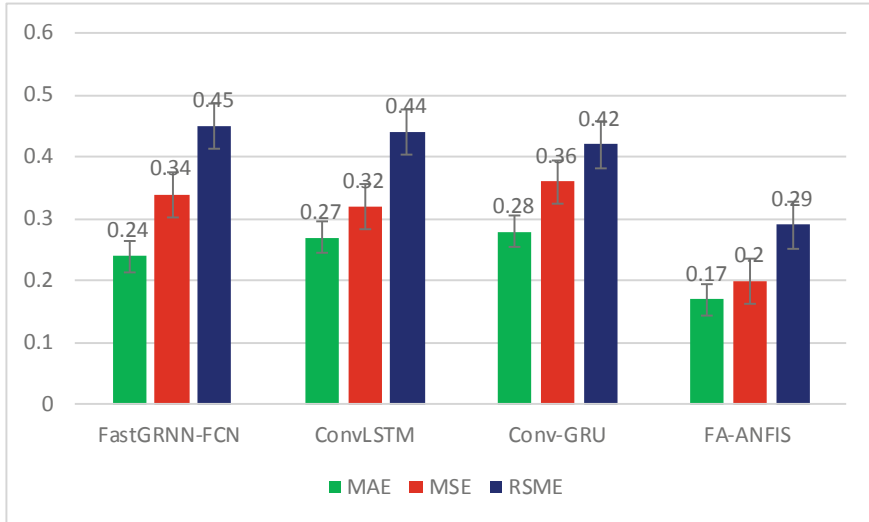


Fig. 6 Training errors comparison results (MAE, MSE, RMSE) with 80% of training data

model's predictive reliability. The suggested FA-ANFIS approach's sudden flood prediction efficacy is determined using Eqs. 13–16.

$$Accuracy = \frac{TP + TN}{TTP + TN + FP + FN} \quad (13)$$

$$Precision = \frac{TP}{TP + FP} \quad (14)$$

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

$$F1 = \frac{2(Precision \times Recall)}{Recall + Precision} \quad (16)$$

- True Positive flood forecasting: True positive flood forecasting occurs when the proposed method forecasts the probability of flooding utilizing trained FA-ANFIS and IoT-monitored input weather data.
- True negative flood forecasting: True negative flood forecasting occurs when the proposed FA-ANFIS-based flood prediction model accurately forecasts non-flooded events.
- False positive flood forecasting: False positive flood forecasting happens whenever the proposed FA-ANFIS-based flood prediction model mistakenly forecast the flooded events.

- False negative flood forecasting: False negative flood forecasting occurs when the proposed FA-ANFIS-based flood forecasting model forecasts non-flooded events incorrectly.

Figure 7 compares the proposed flood forecasting strategy to existing AI-based methodologies in terms of accuracy, precision, recall and F1-score. The suggested flood forecast model has a 96% accuracy, a 97% precision, a 95.3% recall, and a 94.9% F1-score. Training time for AI models is a critical performance metric. In general, the amount of time required to train an AI model varies according to the configuration of the computer being trained, the efficiency of the optimization technique, and the network complexity of the AI model. In general, highly effective AI models consume less time and processing power to train. In this study, the training time for AI models are calculated in minutes. Figure 8 illustrates the results of a training comparison between FA-ANFIS and existing flood prediction models. According to the results of the comparison, the FA-ANFIS required less training time.

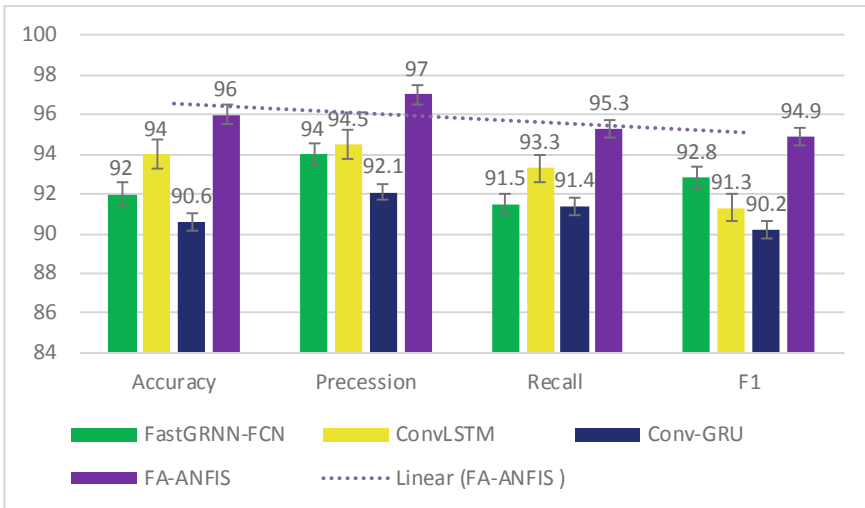


Fig. 7 Accuracy comparison results

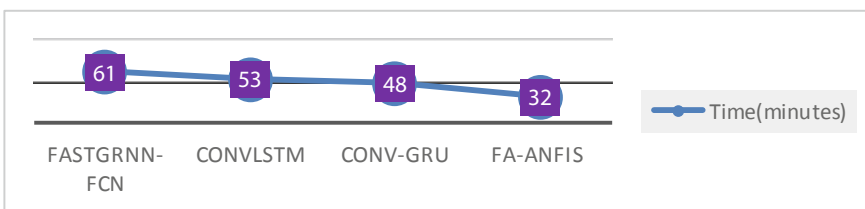


Fig. 8 Proposed and existing flood forecasting frameworks' training times

5 Conclusion

Numerous methods for flood forecasting have been developed by metrological researchers that fail to predict sudden flood (cloud burst) caused by climate change due to inefficient learning parameter selection, decision making and feature selection. In this research, ANFIS was hybridised with the fire fly algorithm to overcome those shortcomings. Additionally, a cloud coverage monitoring device was used in this study to enhance forecast efficacy. FA-ANFIS has a low training error and a good accuracy, as proven by experimental data.

References

1. Vasantha Kumaran T, Murali OM, Rani Senthamarai S (2020) Chennai floods 2005, 2015: vulnerability, risk and climate change. In: Singh R, Srinagesh B, Anand S (eds) Urban health risk and resilience in Asian cities. Advances in geographical and environmental sciences. Springer, Singapore.
2. Kuiry S (2019). Impact of urban sprawl on future flooding in Chennai city, India. *J Hydrol* 574:486–496. <https://doi.org/10.1016/j.jhydrol.2019.04.041>
3. Sundaram S, Devaraj S, Yarrakula K (2021) Modeling, mapping and analysis of urban floods in India—a review on geospatial methodologies. *Environ Sci Pollut Res Int* 28(48):67940–67956. <https://doi.org/10.1007/s11356-021-16747-5>. Epub 2021 Oct 9. PMID: 34626336
4. Elsafi SH (2014) Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Eng J* 53(3):655–662
5. Dtsisibe FY, Ari AAA, Titouna C et al (2020) Flood forecasting based on an artificial neural network scheme. *Nat Hazards* 104:1211–1237
6. Dong S, Yu T, Farahmand H, Mostafavi A (2021) A hybrid deep learning model for predictive flood warning and situation awareness using channel network sensors data. *Comput Aided Civ Inf.* 36:402–420
7. Khan TA, Alam MM, Shahid Z, Su'ud MM (2020) Prior recognition of flash floods: concrete optimal neural network configuration analysis for multi-resolution sensing. *IEEE Access* 8:21006–21022. <https://doi.org/10.1109/ACCESS.2020.3038812>
8. Moishin M, Deo RC, Prasad R, Raj N, Abdulla S (2021) Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm. *IEEE Access* 9:50982–50993. <https://doi.org/10.1109/ACCESS.2021.3065939>
9. Miao S, Hung W-H (2020) River flooding forecasting and anomaly detection based on deep learning. *IEEE Access* 8:198384–198402. <https://doi.org/10.1109/ACCESS.2020.3034875>
10. Souza AS, de Lima Curvello AM, dos Santos de Souza FL, da Silva HJ (2017) A flood warning system to critical region. *Procedia Comput Sci* 109:1104–1109
11. Pavithra D, Jayanthi AN (2021) An improved adaptive neuro fuzzy interference system for the detection of autism spectrum disorder. *J Ambient Intell Human Comput* 12:6885–6897
12. Harifi S, Khalilian M, Mohammadzadeh J, Ebrahimnejad S (2020) Optimizing a neuro-fuzzy system based on nature-inspired emperor penguins colony optimization algorithm. *IEEE Trans Fuzzy Syst* 28(6):1110–1124. <https://doi.org/10.1109/TFUZZ.2020.2984201>
13. Yang XS. (2009) Firefly Algorithms for Multimodal Optimization. In: Watanabe O., Zeugmann T. (eds) Stochastic Algorithms: Foundations and Applications. SAGA 2009. Lecture Notes in Computer Science, vol 5792. Springer, Berlin, Heidelberg.

14. Łukasik S, Żak S (2009) Firefly algorithm for continuous constrained optimization tasks. In: Nguyen NT, Kowalczyk R, Chen SM (eds) Computational collective intelligence. Semantic web, social networks and multiagent systems. ICCCI 2009. Lecture notes in computer science, vol 5796. Springer, Berlin, Heidelberg
15. Mopuri R, Kakarla SG, Mutheneni SR et al (2020) Climate based malaria forecasting system for Andhra Pradesh, India. *J Parasit Dis* 44:497–510
16. Antony Sylvia JM, Pushpa Rani M, Aremu B (2021) Analysis of IoT big weather data for early flood forecasting system. In: 2021 fourth international conference on electrical, computer and communication technologies (ICECCT), pp 1–6. <https://doi.org/10.1109/ICECCT52121.2021.9616941>

A Survey on HTTP Flooding—A Distributed Denial of Service Attack



Hrishikesh Khandare, Saurabh Jain, and Rajesh Doriya

Abstract A Distributed Denial of Service (DDOS) attack is a particular type of Denial of Service (DOS) attack in which, a web server is sent multiple requests in huge amount, simultaneously from unauthorized users in such a way that, the legitimate users are denied access the services of the server. HTTP Flooding attack which exploits the application layer, is amongst some of the very popular DDOS attacks. This attack uses HTTP requests which are indistinguishable from valid user requests. In this paper, the basics of HTTP Flooding attacks have been discussed along with its main types i.e., HTTP GET flood and HTTP POST flood attack, along with the tools and techniques used for the attack. Further, a study of existing methodologies for detection of HTTP Flooding attack in cloud and IoT (Internet of Things) environment has been described. Mitigation of these attacks is also an important aspect. Hence, a study of some existing techniques for the mitigation of application layer HTTP Flooding attack is also shown in later part of this paper.

Keywords Distributed denial of service · HTTP flooding · Application layer · Low and slow attack

1 Introduction

The usage of internet is growing day by day at a very rapid rate. One of the main causes behind this growth is the mobile and web-based applications which get almost any type of tasks done at a click or two without any inconvenience. Most of these internet-based applications are based upon a web server which serves the requests of

H. Khandare (✉) · S. Jain · R. Doriya
Department of Information Technology, National Institute of Technology, Raipur 492010,
Chhattisgarh, India
e-mail: hrishikesh0408@gmail.com

S. Jain
e-mail: sjain.phd2017.it@nitrr.ac.in

R. Doriya
e-mail: rajeshdoriya.it@nitrr.ac.in

different clients from different platforms. As a result, an attack on the web servers acting as a backbone for these applications can cause inconvenience to customers as well as owners of many firms. Hence, security of these web servers against attacks is becoming a point of concern nowadays.

A variety of active and passive are attacks available, which can be used for compromising the communication over internet. But a Denial of Service (DOS) comes into picture, where an attack is needed to be performed on a server in limited number of resources. In a DOS attack, an attacker incapacitates a server, by sending bulk quantity requests which it isn't capable of handling and hence causing denial of services to the intended users [1, 2]. A Distributed Denial of Service attack is a variant of DOS attack which uses a group of well-coordinated nodes to perform DOS attack [3]. This attack is even more dangerous. Since, in these attacks, less resources are required and a server can be incapacitated by using a single point of attack, these attacks are widely adopted across the world.

One of the most dangerous attacks that can disrupt working of a web server is a HTTP flooding attack [4]. Since, in this attack the method to perform attack is by bombarding humongous amount of HTTP requests to the victim and HTTP being a layer 7 protocol, this attack aims to attack the application layer. This is a DDOS attack in which a targeted server is forced to deny genuine user requests. A huge load of requests are sent towards the target during the attack, consuming server's resources such as compute power, bandwidth, memory, and so on. Also, a huge is incurred due to the unavailability of the server. Sometimes the loss can be in the financial, sometimes it can be in terms of customer good will.

HTTP Flooding attacks are majorly of two sorts depending upon the types of requests sent: HTTP GET and HTTP POST attack. The first sort of attack is HTTP GET flooding attack. Numerous computers or devices interact in this attack to send multiple requests to a certain targeted web server for photos, files, or other entities. As a result of running out of resources and being overburdened by this scenario, the target server is forced to refuse services requested by genuine users. HTTP POST attack is the second sort of attack. When a user submits data via a form on a website, the web server should accept request. Then, later sends the data to the database (DB). In comparison to computer and network resources consumed to submit the POST request, the procedure for managing the form data and running the database operations that are required is comparatively intensive. The gap between the number of resources required in both the scenarios is exploited by this attack, by bombarding multiple number of requests to the targeted server and hence resulting in denial of service (see Fig. 1).

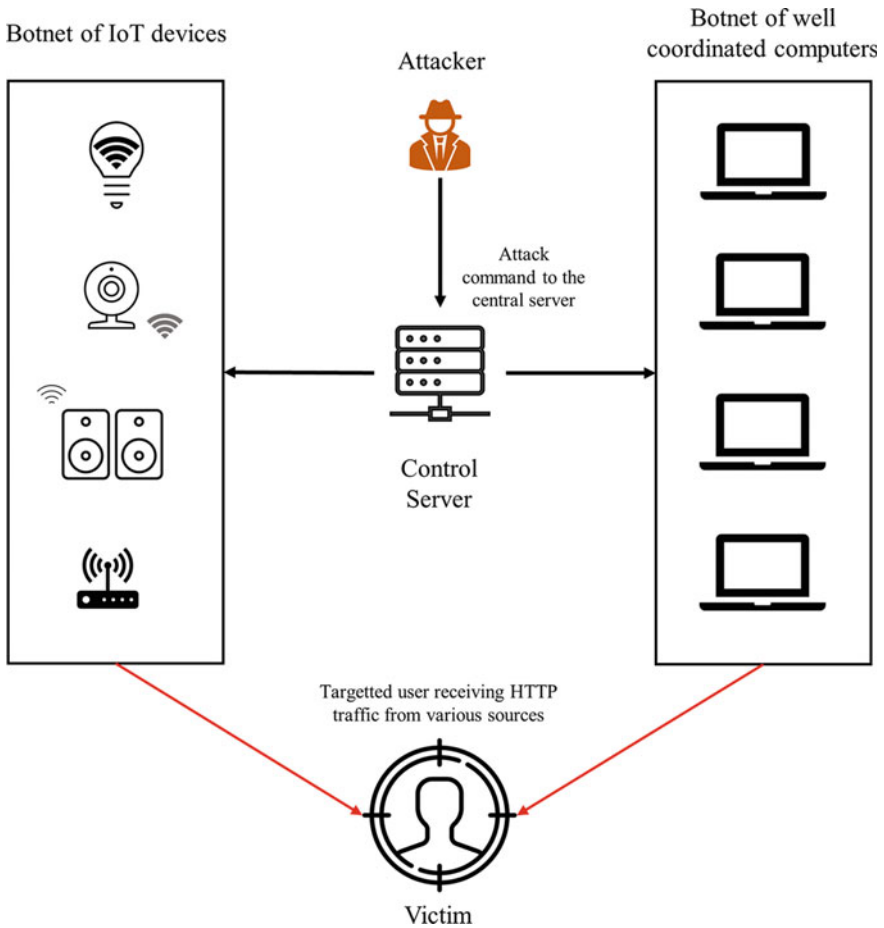


Fig. 1 Image illustrating HTTP flooding attack scenario with the help of Botnets

2 Background of HTTP Flooding Attacks

This section describes the background of HTTP Flooding attacks which are classified broadly into two types. Namely, HTTP GET Flooding attack and HTTP POST Flooding attack. Short descriptions about the tools and techniques used to execute these attacks are also discussed in this section.

2.1 HTTP GET Flood Attack

Flooding attack on the application layer results in unwanted use of bandwidth and other system resources such as compute power and memory. Which in turn prohibits a web page server to work under normal routine. Many free to use tools are available on the world wide web for performing Denial of Service attacks. These tools enable attackers to set the attributes of an attack easily and execute it as well. In [5] Takeshi Yatagai et al. have mainly divided HTTP-GET flooding attacks as follows:

- F5 attack: Many browsers have F5 key as the preset shortcut key for redownloading web pages. In this attack, server is overused using the excessive use of F5 key. F5 attack is very simple to perform and lets us decide the parameters for the attack as well.
- Botnet [6, 7]: Botnet is akin to a computer virus in its behavior. After infecting an individual's personal computer, it executes a HTTP GET flood attack towards a server. In contrast to computer viruses, attack parameters for Botnet can be changed at any time, making it more adaptable.
- Computer Virus: A computer virus infects a machine, causing a web server to be flooded with HTTP GET requests. Such an attack's attributes include a particular universal resource locator of the victim host, time period between attacks, etc. These properties, however, must be established before a machine is infected with a virus.

In [8], R. Sanjeetha et al. have suggested a method of mitigating HTTP GET flood attacks using Software Defined Network (SDN). Defense mechanisms against HTTP GET Flood attacks in traditional networks are generally deployed at the server. However, it might end up exhausting resources which otherwise could have been used to serve the requests of clients. SDN has a different type of architecture that can help deal with this issue [9]. R. Sangeetha et al. demonstrated an HTTP GET Flood DDoS attack on a server in the SDN scenario in [8] and suggested a strategy for mitigation. HTTP GET flood attacks can be prevented at the server level by implementing a captcha test or a firewall that can filter malicious traffic. Both of these methods, however, are computationally demanding. Since network control and forwarding are separated in SDN, it may be programmed directly and changed to meet the network's changing needs. SDN is capable to manage HTTP GET flood attacks of this nature. As the SDN has a controller which can collect per-flow statistics from the switch and then discover the number of requests sent to the primary server making it easy to mitigate the attack. Figure 2 shows the HTTP GET attack scenario.

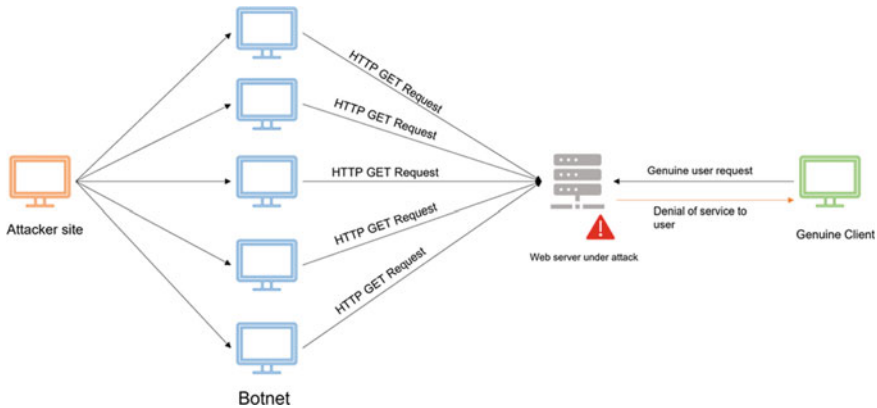


Fig. 2 An image depicting HTTP GET Flooding attack Scenario

2.2 HTTP POST Flood Attack

In this attack, the form fields in a website are analyzed and legitimate HTTP POST requests are sent to the server in bulk quantity. These requests are hard to distinguish from normal legit requests. Also, these requests contain large amount of form data many times and while server is busy incorporating these requests, denial of service for other users is observed. Literature about these types of attack is a bit rare to find. However, there are some tools available for performing these attacks on the internet. Some of the tools are:

- R-U-Dead-Yet (RUDY) [10]: RUDY can be called a “Low and slow” attack. This tool enables attackers to send lengthy form fields in a website. This attack is performed by establishing multiple HTTP-POST connections with target server. However, the number of connections made to the server are not hefty in amount. The number of connections is relatively less than and the sessions are kept alive as long as possible.

During the attack, small packets of data are sent to the server with significant times in between the adjacent two packets. In other words, first the connections are made, but sending over the body of the requests is done at a very low rate. Hence, resources of the server are busy handling these requests and denial of service occurs. This tool has an intuitive interface and is available on the internet.

- GoldenEye [11]: this tool is freely available at GitHub that can be used for testing purposes. In this tool, it is possible to make multiple connections simultaneously to a server in order to check if it can withstand to a flooding attack or not. In this tool multiple modes can be used i.e., it can perform HTTP GET as well as HTTP POST Flood attacks. Also, GET and POST requests can be sent in random order. Using this tool, even a single machine can also incapacitate a server by using genuine HTTP traffic.



Fig. 3 HTTP POST Flooding attack using RUDY tool

- Torshammer [12]: Torshammer is a tool that can perform slow paced HTTP POST in a similar fashion to that of RUDY. However, this tool has an added advantage over RUDY. Since this tool can perform attacks on Tor network, it becomes tedious to reveal the true identity of the attacker. Also, there are various parameters such as number of threads, target address can be set for the attack (see Fig. 3).

3 Detection Techniques for HTTP Flood Attacks

As we have seen, these HTTP flooding attacks are severe in nature and can take down web servers immediately. So, in order to defend against these attacks, or take countermeasures against them, it is important to know how to detect them in first place. Some of the detection techniques in various environments have been discussed below:

3.1 Detection of HTTP Flooding Attack in Cloud Environment

Choi et al. [13] used the entropy of the parameters: volume of data traffic, Internet Protocol address of source, Internet Protocol address of destination, port number of source and port number of destination, to detect HTTP GET flooding attacks. This method determines the parameter thresholds based on the HTTP response of the traffic. This approach has an 88 percent detection rate and has FPs. The limitation of this approach is that the detection rate is not that great and has FPs in addition. The authors in [14] proposed a signature-based Intrusion Detection System approach for identifying genuine and malicious traffic in virtual machines. The network traffic is intercepted and processed by Snort intrusion detection system to determine the attack.

In [15], Authors have used the network log of the virtual machine along with the access log of the web server. From this data, related attributes like number of GET requests in one second, status of virtual machine, time of arrival between consecutive requests, request time, repetition in a given period, alikeness in request string and port number are picked up. These are then provided as an input to a preprocessor which in turn gives normalized features. These are later fed to a detector module. It

its benefit. Random forest [18], decision trees [19], support vector machines [20], K-nearest neighbors [21], and neural networks [22] are among the classifiers we examine for attack detection.

In [23], P. Bhatt and A. Morais have proposed a hybrid network anomaly detection system. Proposed system implements algorithms capable of detecting anomaly with the help of machine learning in order to identify malicious traffic data. Common connection events such as opening of a new TCP connection are generated by the system after monitoring the on-device traffic using a packet sniffing tool in a non-intrusive manner. Further, the feature extraction [24] module analyzes the events over the network and finds out respective connection and traffic characteristics of protocol being considered.

The table which follows, describes some more methods of detecting of HTTP flooding attack in various scenarios (see Table 1).

Table 1 Summary of some existing HTTP flooding attack detection techniques

| Author | Year | Used Methodology/Approach | Limitations |
|-----------------------------|------------|--|--|
| Lu and Yu [25] | 2006 | Detection using browser behavior based upon Hidden Semi-Markov Model | The proposed approach is ineffective for pulse attacks |
| Das et al. [26] | 2011 | DSB: A Query based projected clustering technique | Use of relatively old datasets and normalization of threshold values δ , σ and β is required |
| Choi et al. [27] | 2012 | AIGG Threshold Based Detection | A modern DDOS attack which makes use of small number of requests can bypass the system |
| Wang et al. [28] | 2014 | HTTP SoLDiER detection scheme using large deviation principle | Ineffective against single URL attack |
| Jin Wang et al. [29, 30] | 2013, 2015 | HTTP sCAN: anomaly-based HTTP-flooding detection scheme | Only two surfing attributes are used for attack detection |
| Munivara Prasad et al. [31] | 2017 | Bio-Inspired Anomaly based detection | This approach is only tested with homogenous dataset, its performance is unknown for heterogeneous dataset |
| Sreeram et al. [32] | 2019 | Bio-Inspired bat algorithm for fast and early anomaly detection | The model lacks ability to deal with legitimate traffic that is accompanied by malicious traffic |

4 Mitigation Techniques for HTTP Flooding Attacks

In [33], S. Umarani et al. proposed a method for detection of HTTP flooding with the help of traces of traffic flow. An access matrix is created in this technique. It makes use of PCA to cut short the detection attributes. After that, Naïve Bayes and KNN are used to separate regular and abnormal traffic. The efficiency of the system is quantified based upon the detection and FP rate.

In [34], C. Tang et al. have demonstrated a method to prevent HTTP-GET flood attacks using an approach based upon the meta-data. Here, suspicious requests are identified with the help of BDA (big data analysis) It is claimed that the proposed methodology can provide uninterrupted service to legit users in spite of the attack line rate being nine Gigabits/sec. Firstly, an intelligent probe is used to obtain the meta-data regarding an HTTP connection. It can contain IP as well as the URL. After that, real-time big-data analysis is done on the meta-data, which can yield those IP addresses whose HTTP requests tend to break the norms again and again. A blacklist of these IP addresses further enables the inline devices such as Firewall, load balancer to deploy rate limiting conditions to prevent the attack.

In [35], the author has proposed a comber approach known as filtering tree for security services. This filtering tree consists of 5 filters to detect and mitigate the HTTP based distributed denial of service attack. This approach is helpful for XML DDOS attacks as well. The primary goal of this paper is detection and prevention of HTTP and XML based distributed denial of service attack attacks in cloud computing. Inside the system architecture, there exists 3 modules namely Embed SOAP message, IP Traceback and Cloud defender (see Fig. 5).

In embed soap message module, double signature, IP marking and client puzzle mechanisms are used to keep DDOS attacks in check. In the given architecture, IP Traceback is used to store the IPaddresses provided by cloud defender. When, a client message has to pass through IP Traceback, it first verifies the source IP of message with the IP addresses already stored with it. If the IP matches with on of the existing IP addresses, then the request message is deleted. Else, it is sent to cloud defender.

Finally, Cloud defender has 5 phases to filter the attacks namely Sensor filter, Hop Count Filter, IP Frequency divergence filter, Puzzle resolver filter, and double

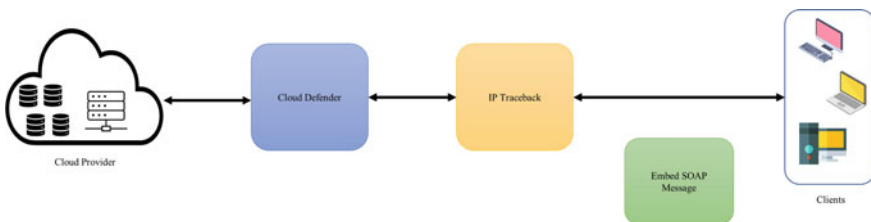


Fig. 5 Proposed architecture for a comber approach for protection of cloud computing against HTTP Distributed Denial of Service attack [35]

signature filter. HTTP DDOS are detected in first four stages and last is for XML attacks.

In [36], H. Mirvaziri has proposed a new methodology to reduce the effects of a HTTP GET flooding attack. In this paper, author have implemented a HGFMS method, which stands for HTTP GET Flood Mitigation System. According to the author, this system sits in between the firewall and web server. In this method, a new criterion server status is introduced by using arbitrary mechanisms for time delay and by using a trap link. Hence, without more performance can be obtained without maintaining the history.

A server is assigned states based on the load upon the server. If server is running on less than 40% power, then it is referred as normal state whereas server running on more than 40% power represents abnormal state. If user traffic is suspicious, user is taken to a page, which asks user to wait for a specific period of time after which a button is activated, then after the press of button, user is redirected to the desired page. A history list and blacklist are also maintained for suspicious for the IPs requesting pages. The IP addresses in the blacklist are not allowed to access the server. The author has also compared the proposed method with IOSEC module and claims to be more effective than the same (see Fig. 6).

In [37], M. R. A. Ahmed et al. have proposed an enhanced hybrid intrusion detection and prevention system for various flooding attacks using a decision tree. This method is claimed to have high detection rate while also reducing the false positives with the help of a decision tree-based method. This decision tree method is based upon the C4.5 algorithm [38, 39]. The authors claim that this proposed technique performs better in terms of detection rate and false positives (FP) when compared to earlier hybrid systems that used signature and anomaly-based detection.

In the proposed system, the traffic observed traffic data is being passed to both signature-based detector module & anomaly-based detector module in parallel. Signature based detector matches the traffic data with existing malicious traffic signatures and checks if the traffic is suspicious. On the other hand, anomaly-based detector with the help of a decision tree-based C4.5 algorithm classifies the data as normal or abnormal traffic. Later, the output of both the module is given input to a decision combiner which is going to generate the decision. If the traffic is suspicious, the packets are dropped from that process. The authors have also compared this technique with some other existing techniques and claim that the proposed algorithm has 99.98% accuracy.

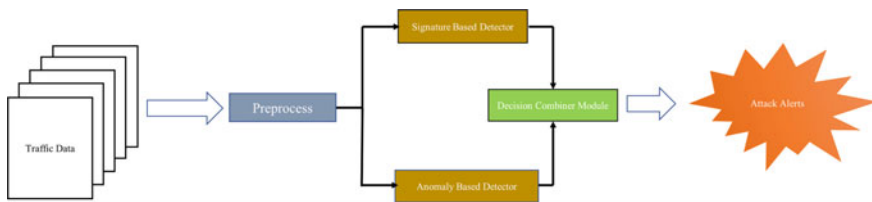


Fig. 6 Architecture of hybrid intrusion detection system [37]

Table 2 Summary of some HTTP flooding attack mitigation techniques

| Author | Year | Used methodology/approach |
|---------------------|------|---|
| Ak et al. [40] | 2012 | Threshold based kernel level HTTP Filter |
| Iyengar et al. [41] | 2014 | Fuzzy Logic based defense mechanism for DDOS attacks |
| Arafat et al. [42] | 2015 | Open-source web proxy: NGINX based mitigation approach. Click or tap here to enter text |
| Vidal et al. [43] | 2018 | Adaptive artificial immune networks-based approach for mitigating DOS flooding |
| Viet et al. [44] | 2017 | Mitigation using NetFPGA based OpenFlow switch in SDN |
| Vanitha et al. [45] | 2018 | Defense module based on Packet filtering using IP tables |

Table 2 shows some more existing mitigation techniques which can be used against HTTP Flooding attacks.

5 Conclusion

In this paper, we have briefly discussed about the HTTP flooding attacks, their broad classification, and the nature of threat proposed by these attacks as well as their severity. We have also discussed various types of existing detection of HTTP flooding attacks in cloud and IoT environment and also studied some existing mitigation techniques. We can conclude that both the types of HTTP flooding attacks have the capability of incapacitating a web server within a short period of time and using a limited pool of resources and can cause huge loss to the victim. Also, these attacks are able to escape the security mechanisms many a times. The existing methodologies can offer protection against HTTP flooding attacks using various approaches. However, there is still scope for some research on this topic which can give a concrete solution to the problem.

References

1. Estevez-Tapiador JM, Garcia-Teodoro P, Diaz-Verdejo JE (2005) Detection of Web-based attacks through Markovian protocol parsing. In: 10th IEEE symposium on computers and communications (ISCC'05), pp 457–462. <https://doi.org/10.1109/ISCC.2005.51>.
2. Byers S, Rubin AD, Kormann D (2004) Defending against an internet-based attack on the physical world. *ACM Trans Internet Technol* 4(3):239–254. <https://doi.org/10.1145/1013202.1013203>
3. Douligeris C, Mitrokotsa A (2004) DDoS attacks and defense mechanisms: classification and state-of-the-art. *Comput Netw* 44(5):643–666. <https://doi.org/10.1016/J.COMNET.2003.10.003>

4. Verma A, Kumar Xaxa D (2016) A Survey on HTTP flooding attack detection and mitigating methodologies. *Xaxa Int J Innov Adv Comput Sci*. Accessed Dec 02, 2021. <https://www.researchgate.net/publication/309385762>
5. Yatagai T, Isohara T, Sasase I (2007) Detection of HTTP-GET flood attack based on analysis of page access behavior. In: 2007 IEEE pacific rim conference on communications, computers and signal processing, pp 232–235. <https://doi.org/10.1109/PACRIM.2007.4313218>
6. Lee J-S, Jeong H, Park J-H, Kim M, Noh B-N (2008) The activity analysis of malicious HTTP-based botnets using degree of periodic repeatability. In: 2008 international conference on security technology, pp 83–86. <https://doi.org/10.1109/SecTech.2008.52>
7. Suen HY, Lau WC, Yue O (2010) Detecting anomalous web browsing via diffusion wavelets. In: 2010 IEEE international conference on communications, pp 1–6. <https://doi.org/10.1109/ICC.2010.5502089>
8. Sanjeetha R, Shastry KNA, Chetan HR, Kanavalli A (2020) Mitigating HTTP GET FLOOD DDoS attack using an SDN controller. In: 2020 international conference on recent trends on electronics, information, communication technology (RTEICT), pp 6–10. <https://doi.org/10.1109/RTEICT49044.2020.9315608>
9. Goransson P, Black C, Culver T (2016) Software defined networks: a comprehensive approach. Morgan Kaufmann
10. Damon E, Dale J, Laron E, Mache J, Land N, Weiss R (2012) Hands-on denial of service lab exercises using Slowloris and RUDY. In: Proceedings of the 2012 information security curriculum development conference, InfoSec CD 2012, pp 21–29. <https://doi.org/10.1145/2390317.2390321>
11. GitHub-jseidl/GoldenEye: GoldenEye Layer 7 (KeepAlive+NoCache) DoS Test Tool. <https://github.com/jseidl/GoldenEye>. Accessed Dec 02, 2021
12. Torshammer download | SourceForge.net. <https://sourceforge.net/projects/torshammer/>. Accessed Dec 02, 2021
13. Choi J, Choi C, Ko B, Kim P (2014) A method of DDoS attack detection using HTTP packet pattern and rule engine in cloud computing environment. *Soft Comput* 18(9):1697–1703
14. Bakshi A, Dujodwala YB (2010) Securing cloud from DDOS attacks using intrusion detection system in virtual machine. In: 2010 second international conference on communication software and networks, pp 260–264. <https://doi.org/10.1109/ICCSN.2010.56>
15. Raja Sree T, Mary Saira Bhanu S (2020) Detection of HTTP flooding attacks in cloud using fuzzy bat clustering. *Neural Comput Appl* 32(13):9603–9619. <https://doi.org/10.1007/s00521-019-04473-6>
16. Sree TR, Bhanu SMS (2018) Detection of HTTP flooding attacks in cloud using dynamic entropy method. *Arab J Sci Eng* 43(12):6995–7014. <https://doi.org/10.1007/s13369-017-2939-7>
17. Doshi R, Apthorpe N, Feamster N (2018) Machine learning DDoS detection for consumer internet of things devices. In: 2018 IEEE security and privacy workshops (SPW), pp 29–35. <https://doi.org/10.1109/SPW.2018.00013>
18. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
19. Safavian SR, Landgrebe D (1991) A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern* 21(3):660–674. <https://doi.org/10.1109/21.97458>
20. Pisner DA, Schnyer DM (Jan 2020) Support vector machine. In: Machine learning: methods and applications to brain disorders, pp 101–121. <https://doi.org/10.1016/B978-0-12-815739-8.00006-7>
21. Bhatia N, Vandana (Jul 2010) Survey of nearest neighbor techniques. (IJCSIS) *Int J Comput Sci Inf Secur* 8(2). <https://arxiv.org/abs/1007.0085v1>. Accessed Dec 02, 2021
22. Wang S-C (2003) Artificial neural network. In: Interdisciplinary computing in java programming, pp 81–100. https://doi.org/10.1007/978-1-4615-0377-4_5
23. Bhatt P, Morais A (2018) HADS: hybrid anomaly detection system for IoT environments. In: 2018 international conference on internet of things, embedded systems and communications (IINTEC), pp 191–196. <https://doi.org/10.1109/IINTEC.2018.8695303>

24. Levine MD (1969) Feature extraction: a survey. *Proc IEEE* 57(8):1391–1407. <https://doi.org/10.1109/PROC.1969.7277>
25. Lu WZ, Yu SZ (2006) An HTTP flooding detection method based on browser behavior. 2006 international conference on computational intelligence and security, ICCIAS 2006, vol 2, pp 1151–1154. <https://doi.org/10.1109/ICCIAS.2006.295444>
26. Das D, Sharma U, Bhattacharyya DK (2011) Detection of http flooding attacks in multiple scenarios. In: ACM international conference proceeding series, pp 517–522. <https://doi.org/10.1145/1947940.1948047>
27. Choi YS, Kim IK, Oh JT, Jang JS (2012) AIGG threshold based HTTP GET flooding attack detection. *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol 7690, pp 270–284. https://doi.org/10.1007/978-3-642-35416-8_19
28. Wang J, Xiaolong Y, Zhang M, Keping L, Jie X (2014) HTTP-SoLDiER: an HTTP-flooding attack detection scheme with the large deviation principle. *15 Sci China Inf Sci*, 57(15):102301. <https://doi.org/10.1007/s11432-013-5015-2>
29. Wang J, Zhang M, Yang X, Long K, Xu J (2015) HTTP-sCAN: detecting HTTP-flooding attack by modeling multi-features of web browsing behavior from noisy web-logs. *China Commun* 12(2):118–128. <https://doi.org/10.1109/CC.2015.7084407>
30. Wang J, Zhang M, Yang X, Long K, Zhou C (2013) HTTP-sCAN: detecting HTTP-flooding attack by modeling multi-features of web browsing behavior from noisy dataset. In: 2013 19th Asia-Pacific conference on communications, APCC 2013, pp 677–682. <https://doi.org/10.1109/APCC.2013.6766035>
31. Munivara Prasad K, Rama Mohan Reddy A, Venugopal Rao K (2017) BIFAD: bio-inspired anomaly based HTTP-flood attack detection. *Wirel Pers Commun* 97(1):281–308. <https://doi.org/10.1007/S11277-017-4505-8/TABLES/7>
32. Sreeram I, Vuppala VPK (2019) HTTP flood attack detection in application layer using machine learning metrics and bio inspired bat algorithm. *Appl Comput Inf* 15(1):59–66. <https://doi.org/10.1016/J.ACI.2017.10.003>
33. Umarani S, Sharmila D (2014) Predicting application layer DDOS attacks using machine learning algorithms. *Int J Comput Control Quantum Inform Eng* 8(10):1780–1785
34. Tang C, Tang A, Lee E, Tao L (2015) Mitigating HTTP flooding attacks with meta-data analysis. In: 2015 IEEE 17th international conference on high performance computing and communications, 2015 IEEE 7th international symposium on cyberspace safety and security, and 2015 IEEE 12th international conference on embedded software and systems, pp 1406–1411. <https://doi.org/10.1109/HPCC-CSS-ICCESS.2015.203>
35. Karnwal T, Sivakumar T, Aghila G (2012) A comber approach to protect cloud computing against XML DDoS and HTTP DDoS attack. In: 2012 IEEE students' conference on electrical, electronics and computer science: innovation for humanity, SCEECS 2012. <https://doi.org/10.1109/SCEECS.2012.6184829>
36. Mirvaziri H (Nov 2017) A new method to reduce the effects of Http-get flood attack. *Futur Comput Inform J* 2. <https://doi.org/10.1016/j.fcij.2017.07.003>
37. Ahmed MRAG, Ali FMA (Sep 2019) Enhancing hybrid intrusion detection and prevention system for flooding attacks using decision tree. In: Proceedings of the international conference on computer, control, electrical, and electronics engineering 2019, ICCCEEE 2019. <https://doi.org/10.1109/ICCCEEE46830.2019.9071191>
38. Archana S, Elangovan K (2014) Survey of classification techniques in data mining. *Int J Comput Sci Mob Appl* 2:65–71. www.ijcsma.com. Accessed Dec 02, 2021
39. Brijain M, Patel R, Kushik M, Rana K (2021) A survey on decision tree algorithm for classification. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.673.2797>. Accessed Dec 02, 2021
40. Ak MI, George L, Govind K, Selvakumar S (2012) Threshold based kernel level http filter (tbhf) for ddos mitigation. *Int J Comput Netw Inform Secur* 4(12):31
41. Iyengar NCSN, Banerjee A, Ganapathy G (2014) A fuzzy logic based defense mechanism against distributed denial of service attack in cloud computing environment. *Int J Commun Netw Inform Secur (IJCNIS)* 6(3)

42. Arafat MY, Alam MM, Alam MF (2015) A practical approach and mitigation techniques on application layer DDoS attack in web server. *Int J Comput Appl* 131(1):975–8887. www.tesserver.com. Accessed Dec 03, 2021
43. Vidal JM, Orozco ALS, Villalba LJG (2018) Adaptive artificial immune networks for mitigating DoS flooding attacks. *Swarm Evol Comput* 38:94–108. <https://doi.org/10.1016/J.SWEVO.2017.07.002>
44. Viet AN, Van LP, Minh HAN, Xuan HD, Ngoc NP, Huu TN (Nov 2017) Mitigating HTTP GET flooding attacks in SDN using NetFPGA-based OpenFlow switch. *ECTI-CON 2017-2017 14th international conference on electrical engineering/electronics, computer, telecommunications and information technology*, pp 660–663. <https://doi.org/10.1109/ECTICON.2017.8096324>
45. Vanitha KS, Uma SV, Mahidhar SK (Jun 2018) Distributed denial of service: attack techniques and mitigation. In: *2nd international conference on circuits, controls, and communications, CCUBE 2017-proceedings*, pp 226–231. <https://doi.org/10.1109/CCUBE.2017.8394146.K>

Tracking and Monitoring of Soldiers Using IoT and GPS



Adlin Sheeba, A. Vinora, P. Ananth, K. Nithya, V. Nisha Jenipher,
and U. Surya

Abstract The article describes an Internet of Things-based military surveillance and tracking system. The system's goal is to follow every soldier's movement on the battlefield, to monitor them, and to issue either instructions or commands. The system includes a humidity sensor, a gyro sensor, a fire sensor, and a pulse sensor, among other sensors. Using a GPS module, this technology allows the control unit to track the precise location of soldiers. The present position and location of the soldier in his side are displayed on an LCD monitor. The sensor data is transferred via an Arduino module. In order to avoid the risk, the army must have a secured medium of communication, so that we can improve the defense system of the armed forces. The serial monitor is utilised in the control room to track and monitor each warrior's activity and provide essential instructions for advancements in the fighting field. They can also be used to determine who requires medical treatment, such as first aid, or immediate attention due to cannon attacks, as well as who requires additional weaponry, such as rifles, bullets, and cannons. The device not only aids in the monitoring of soldier movements, but it also facilitates communication between commanders and soldiers.

Keywords IoT · Tracking system · Monitoring · GPS · Sensors · Soldiers · Arduino

A. Sheeba (✉) · A. Vinora · V. Nisha Jenipher
Department of Computer Science and Engineering, St. Joseph's Institute of Technology, Chennai,
India
e-mail: adlinsheeba78@gmail.com

P. Ananth
Tata Consultancy Services, Chennai, India

K. Nithya · U. Surya
Department of Information Technology, Dr. M.G.R Educational and Research Institute, Chennai,
India
e-mail: nithya.it@drmgrdu.ac.in

U. Surya
e-mail: surya07ananthi@gmail.com

1 Introduction

In today's world, the terrorists sometimes invade our country and try to dominate our country even though our army is trying to secure it. In some cases we renounce our warriors to secure our land in battle. Due to surveillance issues during the civil war, we undoubtedly lose to our formidable enemy and usually tend to surrender to him voluntarily.

The main task of military communications is to ensure that commanders and staffs at all levels are able to maintain constant control over subordinate forces in all conditions and provide the armed forces with timely signals about the threat of enemy attack and the conduct of battles to submit readiness. The common main requirements for military communications are timeliness of intellectual establishment, sufficient reliability of necessary operations, possible speed of appropriate measures and strict secrecy of transmitted information.

The reliability of military communications is achieved through the combined use of various types of communications depending on the combat situation. Military communications are usually organized on the intellectual basis of the Supreme Commander's final decision, the correct instructions from the Senior Commander-in-Chief and the necessary orders from the higher headquarters, usually depending on the constant availability and ideal state of the communications forces and sophisticated equipment.

The security of the military Communication typically remains a key issue. Therefore, we usually need a sophisticated device that usually improves official communication as well as ensures operational security.

There are typically many systems developed for the armed soldiers, especially for the medical care of the sufficient soldiers. In the realm of sensors, there has been a lot of progress, and many wearable body sensors (WBS) have been developed. In [8], the system of the WBS is mounted on the soldier's jacket and the data of the WBS is transmitted to the unmanned aerial vehicle (UAV) and to explain the strong communication, the data is also transmitted to the military truck.

Arduino Uno and Node MCU are used in designing soldier surveillance system to monitor physiological information [9]. The soldier's orders and medical information can be relayed to the base station in real time, allowing the base station to take the appropriate actions.

We are approaching the use of the Arduino in the recommended arrangement. The practical purpose of the correct system is to carefully track and monitor the active movement of each responsible soldier on the battlefield and issue either instructions or orders. In addition, for those needs all medical assistance like first aid, we can use instant attention due to the cannon attacks. The device not only aids in the monitoring of army movements, but it also aids in communication between commanders and soldiers.

2 Literature Review

For Soldier safety, the Body Sensor Network (BSN) system is mounted on the Soldier's body, consisting of various types of small physiological sensors, biomedical sensors, transmission modules and processing functions, enabling a cost-effective, portable, unobtrusive health monitoring solution can [1].

This article develops a real-time cardiac monitoring system considering cost, ease of use, accuracy, and data security. The technology is designed to permit the health practitioner and the affected person to communicate. The main purpose of this study is to facilitate access to the latest healthcare services for distant cardiac patients [2].

The system will be useful for soldiers involved in missions or in special operations and in different environmental conditions where the temperature is maximum or minimum. This device allows these soldiers and patients to be tracked using GPS (Global Positioning Systems) [3].

The study considers a calamity that strikes a medium-sized smart city in an Alliance country. To provide disaster aid, a small international army has been dispatched. Situational awareness (SA) is critical so that resources, such as persons and materials, can be prioritised to support those who are most in need. Information from Internet of Things (IoT) devices might considerably improve this SA, especially in a smart city context [4].

A device with biosensors is utilised for health care in this research. A GPS is also utilised to track the soldier's location if necessary. A GSM modem is also utilised to make the system wireless-capable [5].

This technology contains a Wi-Fi module that establishes a network with the base station, and the base station receives critical information about the soldier's health parameters and whereabouts. GPS tracks the soldier's position anywhere in the world and the health system also monitors the vital health parameters of the soldiers, providing security and protection for the soldiers [6].

The project details an IoT-based system for troop health monitoring and tracking. Biomedical sensors provide each soldier's body temperature and environmental parameters to the control room. This technology can be helpful in determining the exact location of missing soldiers in critical condition and overcoming the disadvantage of soldiers going missing in the field. The addressing system is also helpful for improving communication between soldiers in emergency situations and for proper navigation to the control room [10].

Wireless technology (My Rio) is used in LabVIEW software for effective communication. The GPS is connected to the soldier unit (mobile unit), which immediately transmits the geographic position of the armed soldier on the battlefield to the base station unit via the My Rio module's transmitter. The signal is received affirmatively by the base station receiver, which tracks the location [11].

The precise location and health parameters of the soldier can be relayed to the base station in real time using this document, allowing for the right response in the event of a crisis. GPS is used to log longitude and latitude so direction is easily known. We utilise the body temperature sensor to detect the soldier's body temperature and the heart rate sensor to assess the soldier's heart rate to determine the soldier's health state [7].

Using GPS modules and wireless body area sensor networks, the Army Control Unit can follow soldiers' locations and check their health status [12].

The proposed technology, according to paper [13], can be affixed to the soldier's body and used to track his health and current location using GPS. The proposed system consists of tiny wearable physiological devices, sensors and transmission modules.

In [14], the soldier site measures temperatures, blood pressure and heartbeat and the server site accesses the soldier site data using IoT. And, using GPS, determines the soldier's current location.

The paper in [15] contains a feature that allows soldiers who become lost or wounded on the front line to chart their positions and record their safety in real time. The data from the sensors and GPS module will be relayed wirelessly using the LoRa kit carried by the soldiers. Army officials may maintain track of every soldier after getting their health condition.

The study [16] focuses on the fundamental elements of ECG telemetry and develops an internet of things (IoT)-based application to monitor patient health in both an indoor and outdoor setting. In the experimental portion, data management parameters are assessed alongside medical terminology to emphasise the planned work performance.

Based on the above papers, the continuous monitoring and tracking is usually carried out to carefully collect the valuable information and various sensors such as temperature detector, humidity sensor, independent GPS module are typically used for tracking purposes and the LCD display is used to show the current readings to the fighters. Arduino UNO was typically used for data transmission from the soldier to the console in all the articles surveyed.

The current system to be implemented typically consists of a sophisticated gyro sensor, active humidity sensor, fire sensor, GPS module and heart rate sensors and Arduino is used to transmit the required data from the soldier to the control unit.

3 Methodology

The system typically consists of sensing elements for pulse, specific location, suspicious movement, excessive humidity and so on as shown in Fig. 1. The data is carefully collected by the Arduino which acts as a microcontroller in the system and it is properly attached with the soldier's war suit as a kit. When compared to other microcontrollers, it uses the least amount of power and has a simple to programme interface. It comes in a small package and allows for simple analogue circuit interface.

The sensor device carefully collects the data from various sensing elements, processes it and the data is promptly sent to the commander-in-chief. He carefully gathers the soldier's data through the monitor console and typically performs the tracking and monitoring.

The Gyro sensor observe the rotational motion of the soldier and his change of orientation, GPS module uses the constellation of satellites and ground stations to calculate accurate position wherever it is located. The weather is monitored using humidity sensors.

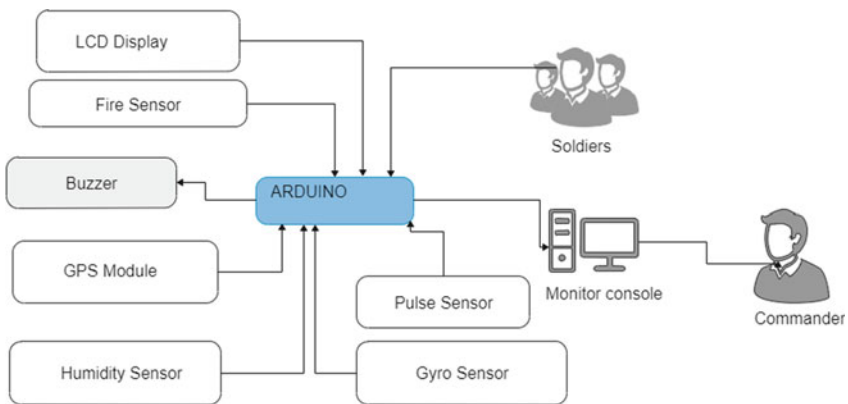


Fig. 1 Tracking and monitoring soldiers

Pulse sensors measure the heartbeat rate of the warrior and decides if he requires any medical assistance as well as confirms he is alive or dead. LCD display shows the current position and the location of the soldier in his side itself. They are employed in the control room to clearly track and carefully monitor the functional activities of each successful warrior and provide the necessary instruction for the intellectual advancements in the war field.

4 Results and Discussions

The sensors like Humidity, Gyro, Fire, Pulse and GPS have been interfaced with the Arduino as shown in Fig. 2. An external supply has been adequately provided to the Arduino by typically using OTG cable and it is additionally employed as the direct link for programming the Arduino. The sensing and data collection is typically performed by the Arduino and is intimately connected to the power supply for performing the task. For official communication between the soldier and the commander-in-chief, Arduino is utilised. The Arduino helps us to gather facts from soldiers.

The additional components like potentiometer, 1 K resistor, PN Junction Diode, 2N222 Transistor and 12 V buzzer are being utilized. As depicted in Fig. 3, the fire sensor is mainly used for finding and responding to the occurrence of a fire or flame.

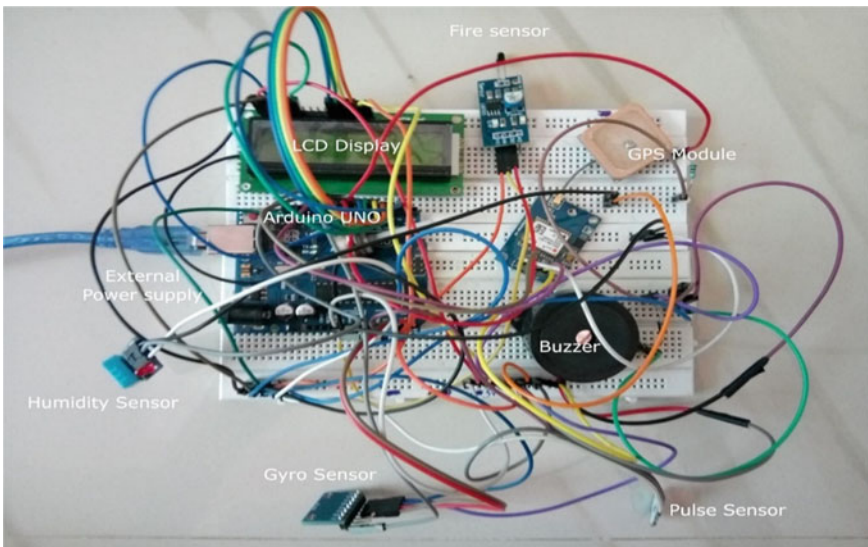


Fig. 2 Hardware Implementation of tracking system



Fig. 3 Output on LCD when fire sensor detects fire nearby

A gyro sensor measure the tilt and lateral orientation of the soldier. The gyro sensor readings with no soldier movement and with soldier movement are shown in Fig. 4. and Fig. 5 respectively.

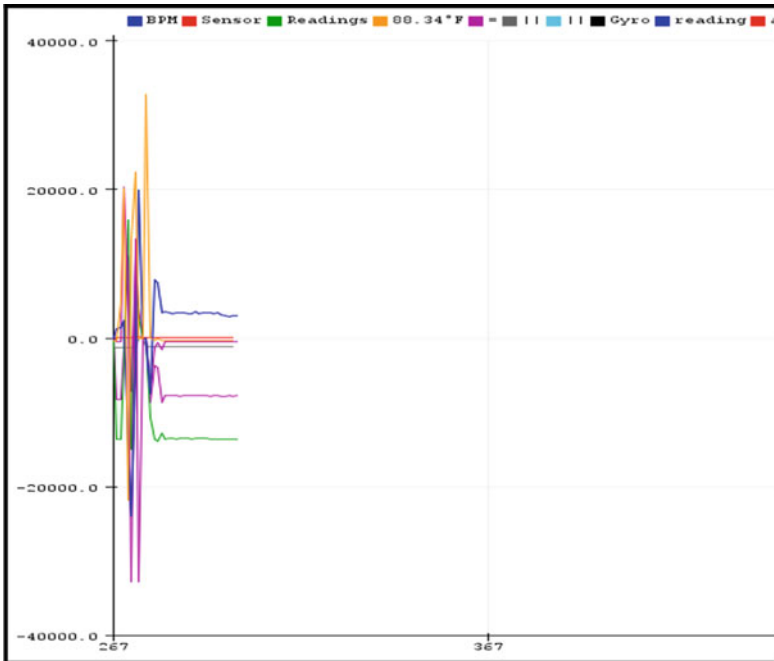


Fig. 4 Gyro sensor readings with no soldier movement (straight lines)

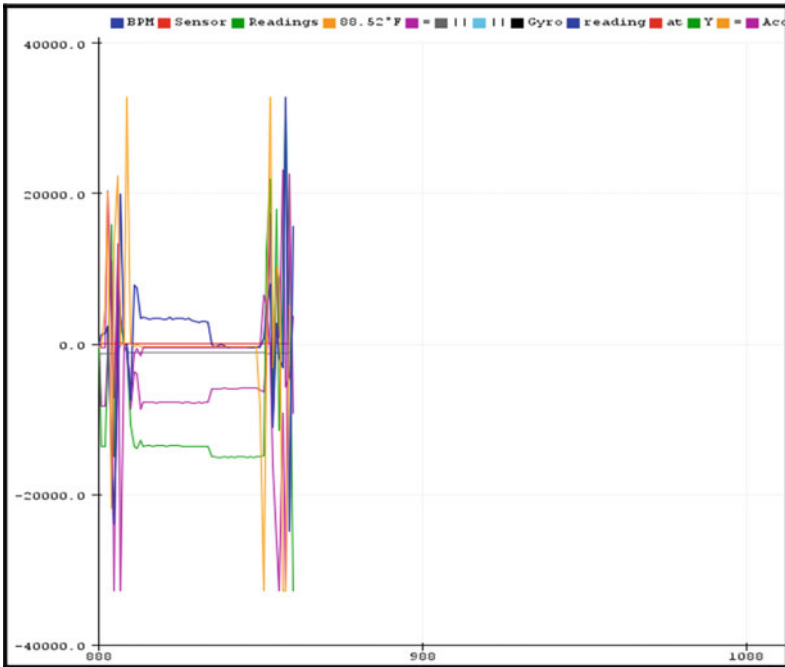


Fig. 5 Gyro sensor readings with soldier movement

The soldier’s heart rate is measured by the Pulse Sensor, which is shown in Fig. 6 as a red line.

Humidity refers to the quantity of water in the air. Human health is affected by the moisture content of the air. Humidity Sensors are very important devices that help in measuring the environmental humidity as shown in Fig. 7.

The readings from the pulse, fire, humidity and gyro sensors are shown in Fig. 8.

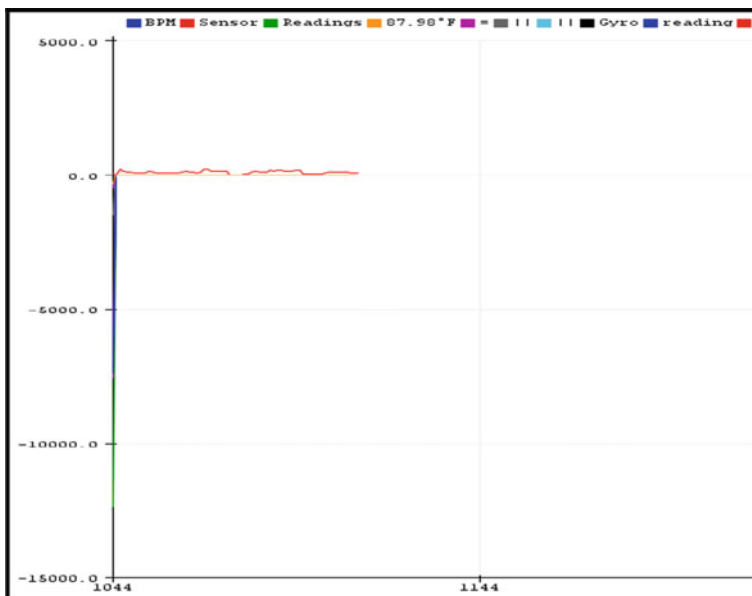


Fig. 6 Pulse sensor reading of the soldier (red color line) (Color figure online)

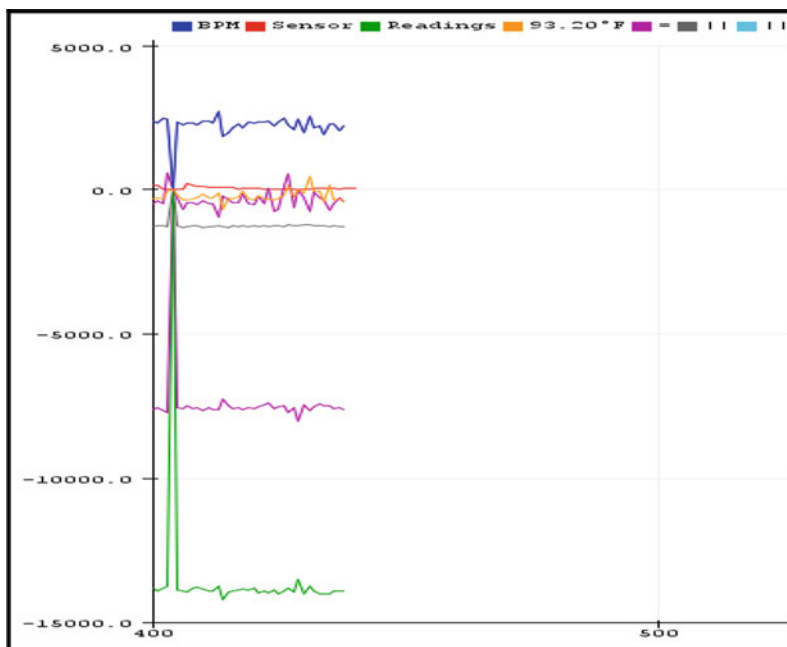


Fig. 7 Humidity sensor readings of the soldier (orange and violet lines) (Color figure online)

```

COM4

Gyro Sensor Readings:

Acceleration at X = 6268 || Acceleration at Y = -6632 || Acceleration at Z = 6892
Gyro reading at X = -1216 || Gyro reading at Y = 9247 || Gyro reading at Z = 6838

Humidity Readings:

Humidity: 86.00%
Temperature: 32.60°C ~ 90.68°F

Pulse Sensor Readings:

BPM: 0

Fire Sensor readings:
No worries

GPS module readings:

Gyro Sensor Readings:

Acceleration at X = -7880 || Acceleration at Y = -13936 || Acceleration at Z = 624
Gyro reading at X = -1232 || Gyro reading at Y = -549 || Gyro reading at Z = -380

Humidity Readings:

Humidity: 87.00%
Temperature: 32.20°C ~ 89.96°F

Pulse Sensor Readings:

BPM: 178

```

Fig. 8 Overall readings sensed by the Arduino from the interfaced sensors

5 Conclusion

The system represents a prototype for the tracking and monitoring purposes of soldiers. Various sensors like Humidity, Pulse, Fire, GPS and Gyro were used for carefully monitoring the movements of soldiers and their current health status instantly. The system is capable to carefully track each and every soldier by typically using Arduino. The Arduino acts as a microcontroller which carefully collects all the sensor data and transfers it to the commander. The system's components are designed to be exceedingly user-friendly. By using these components, commander-in-chief can efficiently monitor and properly maintain the centralized state of the army. By utilizing it, the soldiers also can communicate anywhere to the armed camp within the battle field itself. The device might be used to keep real-time track on soldiers' health.

References

1. Arputha J, Guruchitra M, Shifana Fathima S, Sindhuja M, Sivanesskumar S (2019) Effective monitoring of health status of soldiers using IoT. *Int J Inform Comput Sci* 6(3):104–108
2. Priyanka Kakria NK, Tripathi PK (2015) A real-time health monitoring system for remote cardiac patients using smartphone and wearable sensors. *Int J Telemed Appl* 2015:11 p
3. Chandratre PP, Gahiwade AR, Shaharkar DA (2019) GPS and IOT based soldiers & patient, health indication system. *J Adv Res Electro Engi Tech* 6(1 and 2):5–7
4. Vasilache B, Fuchs C, Johnsen FT, Furtak J, Wrona K, Pradhan M, Krzysztoń M, Dyk M, Marks M, Suri N, Pellegrini V, Zieliński Z (2018) Application of IoT in military operations in a smart city. In: *IEEE international conference on military communications and information systems (ICMCIS)*, pp 1–8
5. Pulimamidia B, Naresh Kumar D, Prasad SVS (2020) Soldier tracking and health monitoring system. In: *AIP conference proceedings*, vol 2269, no 1
6. Anuja G. Asole, P. B. Domkondwar, Monika V. Bhivarkar, Sakshi S. Budhlani, Simran S. Budhlani.: IOT Based Soldier Navigation & Health Indication System. *International Journal of Science Technology & Engineering*. Vol. 4, Issue 7, pp. 90–94 (2018)
7. Bhivarkar MV, Asole AG, Domkondwar PB (2018) IOT and GPS based soldier position tracking and health monitoring system. *Int J Emerg Technol Eng Res* 6(1): 47–50
8. Hari J, Pramoth R, Anoint J, Abhishek S, Prabhu T (2021) Soldiers health monitoring and tracking system using IOT and UAV. *Int Res J Adv Sci Hub* 78–81
9. Archana Padikar A, Cinmayee CK, Chaithra E, Chethan, Puneeth Kumar DN (2020) Health monitoring and soldier tracking system using IOT. *Int J Eng Res Technol* 8(14)
10. Kulkarni P, Kulkarni T (2019) Secure health monitoring of soldiers with tracking system using IoT: a survey. *Int J Trend Sci Res Dev* 3(4):693–696
11. Leela Venkata M, Prudvi S, Eliyaz M, Pavan M, Pragnya Reddy G (2020) Soldier tracking and health monitoring system using LabVIEW. *Int J Emerg Trends Eng Res* 8(5)
12. Gondalia A, Dixit D, Parashar S, Raghava V, Sengupta A, Raja Sarobin V (2018) IoT-based healthcare monitoring system for war soldiers using machine learning. *Procedia Comput Sci* 133:1005–1013
13. Patii IB (2017) Health monitoring and tracking system for soldiers using Internet of Things (IoT). In: *2017 international conference on computing, communication and automation (ICCCA)*, pp 1347–1352
14. Kumbhare KK, Umate S, Khadake S, Wanjari R, Shivdharkar M (2019) Survey paper on soldiers health tracking system using Internet of Things (IOT). *Int J Innov Sci Res Technol* 4(3):261–265
15. Rahimunnisa K (2021) LoRa-IoT focused system of defense for equipped troops [LIFE]. *J Ubiquit Comput Commun Technol* 2(3):153–177
16. Balasubramaniam V (2020) IoT based biotelemetry for smart health care monitoring system. *J Inform Technol Digit World* 2(3):183–190

Study of Vulnerabilities in the Cryptography Algorithms



Harjis Ahuja, Ruchita Bapna, Gargee Bhase, and Narendra Shekocar

Abstract The authors demonstrate the vulnerability of particular algorithms using the principles of timing attacks. RSA and Diffie-Hellman are the algorithms used and comparative study was done on them. In each approach, the private keys are extracted by analysing time differences in various scenarios while modifying specific parameters of the private key. The authors have identified crucial points of difference that can be exploited to decipher the private keys. The authors also attempt to exploit the Elliptical Curve Cryptography technique through a timing attack.

Keywords RSA · Diffie-Hellman · Timing attacks · Exponentiation · Private keys · Cryptography

1 Introduction

Side Channel Attacks are attacks that are based on information leaked through various channels. That information is auxiliary information obtained through the encryption device [1].

In the earlier days, attacks were mainly focused on breaking the ciphertext (ciphertext only attacks) or the plain text (plaintext only attacks). There was also a category of attacks which concentrated on viewing the output of encryption on a select few plaintexts and then breaking the cipher (known plaintext attacks). This was because the operation of an encryption device was perceived to be limited to receiving plaintext and producing ciphertext. Today it is known that they also have additional inputs and outputs. They produce information about the time that various operations take, the variation in the power consumed during different operations, radiation information. They also have unintended outputs. They are voltage fluctuations. Side channel attacks use these auxiliary information to break the cipher and recover the key.

H. Ahuja (✉) · R. Bapna · G. Bhase · N. Shekocar
Dwarkanadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: harjisahuja@gmail.com

N. Shekocar
e-mail: narendra.shekocar@djsce.ac.in

Timing attacks state that it may be possible to find the secret keys by accurately measuring timing differences between instances of an operation. Variation of timing calculations can occur as a result of various factors. Some of the reasons could be branch operation, conditional statements, caching and instructions that are not carried out in fixed time. The concept of Timing attacks was first introduced by Paul Kocher [2].

The Diffie-Hellman Key exchange algorithm uses two private numbers which are never shared to derive a common key via exponentiation. It is an encryption mechanism and produces decryption keys. However, breaking the Diffie-Hellman algorithm is very difficult for a malicious entity.

While implementing the Diffie-Hellman Key Exchange algorithm, two users, let's term them Alice and Bob, choose two natural numbers p and q . P is a prime number, and q is a prime number generator. Q , when raised to numbers not greater than p always produces a different result for two numbers. q is generally smaller than p .

After mutually deciding on p and q , Alice and Bob agree on whole number keys a and b . They are individual to both Alice and Bob and not divulged to each other. In an ideal scenario, they are not written down anywhere nor stored anywhere. After this step is completed, Alice and Bob compute keys a^* and b^* based on the formulae

$$a^* = q^a \text{ mod } p \quad (1)$$

and

$$b^* = q^b \text{ mod } p \quad (2)$$

After the previous step is accomplished, the two of them can share the keys a^* and b^* . They can even be shared on a network which is known to be insecure. The Internet and a Wide Area Network are examples of such networks. Any person using the system can produce a number x by making use of their private keys and these public keys.

The initial letter of its authors' surnames, Ron Rivest, Adi Shamir, and Leonard Adleman, is RSA. It is asymmetric cryptography algorithm. It's a message encryption and decryption algorithm. The encryption key in this method is public, whereas the decryption key is not.

While using this method, the user will create the public key using the larger of the two prime numbers along with the additional values. The prime number is kept a secret. Encryption can be accomplished by anyone. Any person using the system can encrypt. But, provided the size of the public key is adequate, only a person in possession of the private key can carry out decryption. The RSA Algorithm is the name of this approach.

The Elliptic Curve Cryptography is a cryptography algorithm which is used for encryption and is faster, and provides better keys. The representation of the curve is not in an oval form, but in the manner of a line which connects two axes. ECC is derived from a set of equations created from the group where lines intersect axes.

A group is a mathematical entity where operation carried out on two constituents produce a third component. Multiplication of a point and a number produces a separate point, but retrieving the original point is difficult. Elliptical equations possess the single most important property related to encryption; a one-way function. To compute scalar multiplication, double and add algorithm is used on rational points. Understanding the algorithm and applying statistical methods we can obtain the original point, which indirectly conveys about the secret parameters.

1.1 Related Works

Valuable information is leaked out through various sources when computer operate. The acoustic channel is one of the sources. These can give out information about the software running on the computer as well as the calculation about various security keys. Acoustic attack can be carried out by capturing the adversary's laptop's acoustic emanations by using a spying app, a spying bug, compromising mobile devices, faraday cages and air gapes. On an older version of the CASCADE smart card, an attack is initiated. This was the first practical use of the timing attack, and it served as the foundation for much of subsequent research [2]. The use of Montgomery's technique and the Chinese Remainder Theorem to provide a new type of timing attack that permits factorization of an RSA-modulus when the secret exponent is exponentiated [3]. Timing attack against the OpenSSL library has also been implemented. One way in which this can be implemented is using Virtual Machines. While using OpenSSL, the web server can extract private keys from the VM containing the keys.

Some other works where cryptography is extensively used include applying multi-factor authentication in optimistic fair exchange by firstly creating a partial signature and upon verification of the Arbitrator and Signaller details which are used to get trajectory information, creating the complete signature [4]. In the medical field secure transfer of images is crucial and this has been achieved in several ways applying cryptography. One such method involves fusion of Tent, Henon and Sine known as Fusion chaotic map which provides enhanced security [5]. Another technique is combining Deniable authentication encryption and blockchain to develop an encryption algorithm for image sharing [6].

1.2 Organization

Section 1 has Introduction and Related Works, Sect. 2 the Motivation, Sect. 3 the Proposed Model and Future Scope, Sect. 4 the Model, Sect. 5 the Results, Sect. 6 the Prevention and Sect. 7 the Conclusion.

2 Motivation

Side channel attacks are easily mounted attacks which convey information about the secret key of the computers running the software. Consider a situation where after obtaining timing information, an intruder can obtain the secret keys of the attacked computer. These secret keys enable an attacker to know about the running processes and reveal secured information. This hampers the confidentiality and integrity of the system. Side channel attacks are of huge concern since it enables one to access the secret keys. These attacks can be established quickly and do not require extensive resources in terms of libraries required, CPU processing and cost and providing with great results (required secret keys) which in turn help them to seek other information as well. Thus, making it necessary to address this issue by providing mitigation solution to this attack.

3 Proposed Model

Using Timing Attack, we intend to extract the RSA Key, during decryption process. RSA decryption is a modular exponentiation

$$m = c^d \text{ mod } N \text{ where } N = p * q \quad (3)$$

This is the RSA modulus, c is the ciphertext being decrypted, and d is the private decryption exponent. Square and Multiply is the simplest algorithm for determining

$$g * d \text{ mod } q \quad (4)$$

The algorithm squares g about $\log_2 d$ times and multiplies g by $\log_2 d/2$ times. The product is lowered modulo q after each step. Montgomery Multiplication, often known as Montgomery Reduction, is another technique (Fig. 1).

The Montgomery reduction transforms a reduction modulo q into a reduction modulo a power of two, which is denoted by R . This algorithm is faster than other alternatives for modular exponential calculation. A reduction modulo a power of 2 results in timing variation in modular multiplication, increasing the size of timing characteristics. Similarly, Diffie-Hellman Algorithm can be used to extract the secret key. Here we have established client server communication and evaluated the time for exponential operations using above algorithms and observed timing differences in keys. These timing differences occur because of conditional branches as every key has different bits which causes different conditions and the branch taken depends on the condition. Since attacks are feasible their prevention is must and one way is to do away with the variations in execution times of operations. That is, essentially make the operation execute for a fixed amount of time. All operations must take at least as much time as the slowest operation to mask the effect of timing attacks.

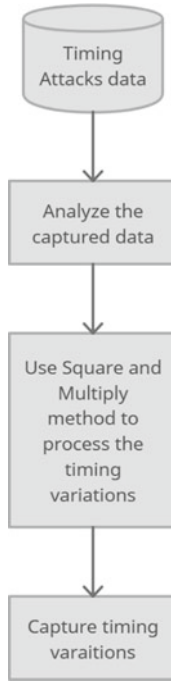


Fig. 1 Block schematic of proposed work

Another method is to introduce random delays. There is a better method to prevent such attacks. Blinding signatures can be added so that attackers are not able to guess the inputs to the exponentiation function.

4 Our Model

4.1 Exponentiation Model

This model can be used to analyse and attack all algorithms which use exponentiation routine. Both RSA and Diffie-Hellman use this, hence it can be used in the analysis of both. Attack Algorithm:

```

x = m
for i = n - 2 till 0
  x = x2
  if (kj == 1) then
    x = x*m
  endfor
return

```

Square and multiplication algorithm is used in which square operation is performed if the private key bit is 1 else normal multiplication is done. Since squaring increases the computation time and also requires modular operation based on these characteristics we can identify the bit to be 1 else 0.

4.2 ECC Model

Timing attack on ECC: The equation of Ellipse E(K) is given as:

$$y^3 = x^3 + ax + b \quad (5)$$

The basic operation performed in Elliptic Curve Cryptography (ECC) is scalar multiplication. Consider a point P on E(K). Now to compute Q, we use the equation $Q = k*P$, where $k*P$ represents the addition of P to itself k times. The discrete logarithm problem consists in finding the value of k from the values of P and $Q = k*P$. Now to attack the algorithm, we need to find the value of k. For example given a set of inputs $S_m = M_1, M_2, M_n$ and set of keys $S_k = K_1, K_2, K_d$, here d is the number of possible keys. The timing distribution is defined as $P_i(t) = F(S_m, k)$ which is different for each key. The attacked pc timing information is measured to form a timing distribution P(t) from the given set of input values S_m . To attack the system usual detection technique is used in which the value of K_i is to be identified from the known values $P_i(t)$ and P(t).

```

Let Q=P
for j=W-2 to 0 do
  Q = 2Q
  if kj=1 then Q=Q+P
return Q

```

To find the value of key K_i , we try to identify each bits. The jth bit of secret key k is denoted by k_j . Total running time for point P_i which is input is given as

$$T_i = e_i + (D_{i,j} + k_j A_{i,j}) \quad (6)$$

This formula is evaluated from $j = 0$ to $w-1$, where: $A_{i,j}$ represents the time required for performing addition operation of bit j . $D_{i,j}$ represents doubling operation of bit j . E_i represents noise.

The value of $A_{i,j}$ and $D_{i,j}$ are influenced by the values of point P_i and j . Let $0 \leq j < r < w$, and assume that we already know the upper bits k_{w-1}, \dots, k_{r+1} of the binary representation of the secret key k . Here the purpose is to determine values of $D_{i,r}$ and $A_{i,r}$ using the above information. The value of point Q is identified using point P_i at the beginning of iteration $j = r$. As soon as we know the points P and Q and using the decryption device we can identify $D_{i,r}$ and $A_{i,r}$ (assuming that single point operation running time can be identified). Using $D_{i,r}$ and $A_{i,r}$ probable value of k_r is determined using statistical methods such as standardization/normalization.

4.3 RSA Model

In this model, we assume the most significant bit to be 1 and then we further extract the bits, one at a time, proceeding towards the least significant bit. RSA basically consists of two large prime numbers. Let's assume these numbers as p and q .

Now, let n be a number which is the product of p and q .

$$d \times e = 1 \pmod{\phi(n)} \tag{7}$$

D is the private key. We attempt to extract the private key using the timing variances. To further attack the exponentiation mechanism of RSA, we use the square and multiply algorithm.

$$m = c^d \pmod{n} \tag{8}$$

4.4 Diffie-Hellman Model

The major goal of Diffie-Hellman key exchange algorithm is computation of

$$R = y^x \pmod{n} \tag{9}$$

In this case, everyone knows n , and an adversary can derive y . The purpose of the attacker is to identify x , the secret key. During the attack, the attacker captures the operation time of $y^x \pmod{n}$ for several different values.

We begin with establishing a communication between Client and Server, here the Client tries to eavesdrop the private of the Server. Since it is required that computation must be performed for several values of y we make the same Client act as the several

Clients by changing its private key continuously and asking to interact with the Server in order to establish a secret communication key.

The attacker should be aware of the computation time in this case, i.e., the client should be aware of the computation time. Generally, most of the timing variation existing in a modular multiplication operation are caused to the modular reduction steps. We use the Montgomery multiplication algorithm which helps to stem the mod n reduction steps. Thus, as a result of this, the size of timing characteristics increases.

These timing measurements are observed by the client for several values of y , on which statistical measures are carried (Normalization or Standard Deviation) to get the private key exponents of the server.

4.5 Experimental Settings

For this experiment a computer with RAM 4 GB and a Processor of Pentium 5 and above is required. Here we have applied Diffie Hellman key exchange between two users (Client–Server model) multiple times by changing the value of Public Key. We are observing the time difference during each communication (for all bits except first one) to get information about the Private key (used by server to perform such calculation).

5 Results

After establishing connection with client several iterations are done where the time difference is calculated every time for the secret key being evaluated (Fig. 2).

After establishing connection with sever several iterations are done where different private key with same public key is used to communicate with the server and evaluating the required secret key for symmetric (Fig. 3).

The output obtained from the execution of code which is displaying the time taken by server's private key bit by bit to perform for various values of Client's private key; for each iteration we observe that the second bit (Bit 1) of Private key of server is taking significantly less time to execute in comparison with the third bit (Bit 2) of the Private key, indicating that extra operation is performed for third bit and its value is 1 and second bit's value is 0. In the above example the server's private key is 5 which is represented as 101 in bit form, assuming the first bit to be 1, the remaining bits values are evaluated based on the timing difference observed. Hence, implementing timing attack for Diffie-Hellman model (Fig. 4 and Table 1).

Side Channel Attacks are of various kind. Acoustic Cryptanalysis is an attack in which the high-pitched noise emitted by computers is captured, which can give us information about the secret keys. For Acoustic Attacks lab-grade setup is required which is expensive and not portable, these setups provide best measurement capabilities. Another option is a portable measurement setup, which offers somewhat

```
C:\Windows\System32\cmd.exe
C:\Users\1HP_PC\Downloads>java GreetingServer.java
C:\Users\1HP_PC\Downloads>
waiting for client on port 8088...
Just connected to /192.168.56.1:56263
From Server : Private Key = 5
From Client : P = 23.0
From Client : G = 9.0
From Client : Public Key = 8.0
Time diff: 10201
Time diff: 78500
Secret Key to perform Symmetric Encryption = 16
From Server : P = 23.0
From Client : G = 9.0
From Client : Public Key = 3.0
Time diff: 95600
Time diff: 38000
Secret Key to perform Symmetric Encryption = 13
From Server : P = 23.0
From Client : G = 9.0
From Client : Public Key = 4.0
Time diff: 16400
Time diff: 72700
Secret Key to perform Symmetric Encryption = 12
From Server : P = 23.0
From Client : G = 9.0
From Client : Public Key = 13.0
Time diff: 15300
Time diff: 87701
Secret Key to perform Symmetric Encryption = 4
From Server : P = 23.0
From Client : G = 9.0
From Client : Public Key = 2.0
Time diff: 14000
Time diff: 27800
Secret Key to perform Symmetric Encryption = 9
From Server : P = 23.0
From Client : G = 9.0
From Client : Public Key = 10.0
Time diff: 15500
Time diff: 25300
Secret Key to perform Symmetric Encryption = 3
```

Fig. 2 Server side results with timing variations

```
C:\Windows\System32\cmd.exe
Just connected to /192.168.56.1:8088
a5
From Client : Private Key = 5
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 16.0
a6
From Client : Private Key = 6
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 13.0
a7
From Client : Private Key = 7
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 12.0
a8
From Client : Private Key = 8
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 4.0
a9
From Client : Private Key = 9
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 9.0
a10
From Client : Private Key = 10
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 3.0
a11
From Client : Private Key = 11
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 1.0
a12
From Client : Private Key = 12
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 8.0
a13
From Client : Private Key = 13
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 10.0
a14
From Client : Private Key = 14
From Server : Public Key = 8.0
Secret Key to perform Symmetric Encryption = 6.0
```

Fig. 3 Results on client side

lower measurement capabilities, but is battery-operated and fits in a briefcase. Also, to record the acoustic emanations we require specific laptops which can provide us with distinguishable results. Timing attacks are popular. In this attack secret keys may be found out by measuring the time difference between various private key operations. The attack can be made possible for any routine that has some variable execution components. Timing attack against the OpenSSL library has also been implemented. One way in which this can be implemented is using Virtual Machines. A Virtual Machine Monitor is used to enforce separation between two

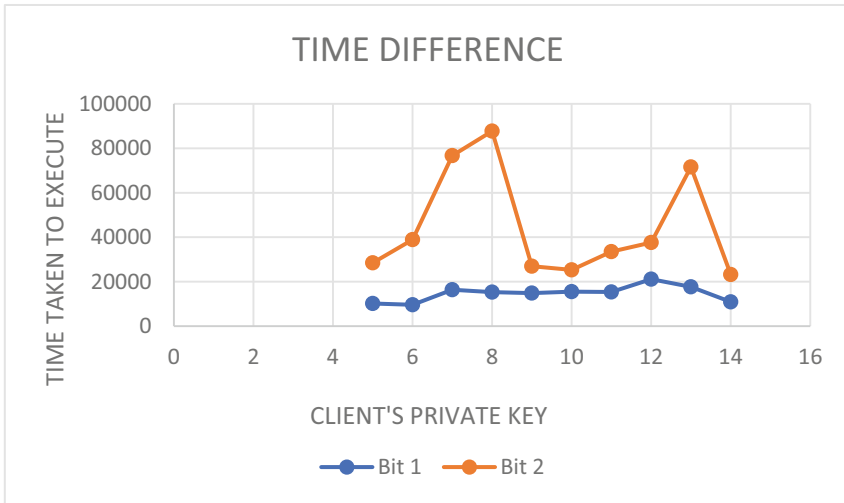


Fig. 4 Time difference observed for Bit 1 and Bit 2

Table 1 Comparison table of existing research work

| Serial no | Name | Strength | Weakness |
|-----------|--|--|---|
| 1 | RSA Key Extraction Via Low-Bandwidth Acoustic Cryptanalysis | RSA Key Extraction using acoustic attacks | Success of the attack depends on age and configuration of the machine and several other factors |
| 2 | Timing attacks on implementation on Diffie-Hellman, RSA, DSS and other systems | Timing attacks proposed | Implementation not proposed |
| 3 | Remote timing attacks are practical | Actual timing attacks are implemented on OpenSSL | Attack is a bit complex and specific to OpenSSL |

Virtual Machines. While using OpenSSL, the web server can extract private keys from the VM containing the keys.

6 Prevention

Timing attack prevention techniques include variations in execution time of operations. That is, essentially make the operation execute for a fixed amount of time. However, there is difficult to do practically. There are many factors which add to

the difficulty, a few major ones being platform dependencies, compiler optimizations, RAM cache hits and instruction timings. Also, there is a major disadvantage of fixed time executions being noticeably slower than their variable counterparts. A chain is only as strong as its weakest link, and the same rule applies here. All operations must take at least as much time as the slowest operation to mask the effect of timing attacks. Another method is to introduce random delays. Use a decryption routine with an input-independent number of operations. Execute the largest number of Montgomery additional reductions possible, even if they aren't required. It may be impossible to restore existing systems without changing their entire decryption method. Moreover, 'overly-optimizing' compilers that eliminate non-functional code that is there to cover a time leak should be avoided. Blinding signatures can be added so that attackers are not able to guess the inputs to the exponentiation function. For blinding, we implement three steps. First, we calculate $x = \text{reg} \bmod N$ before each decryption for a different random r . Then decrypt in the usual way. Finally, unblind: divide by $r \bmod N$ to get the ciphertext g 's decryption. Because r is random, x should be random as well – and input g should have a low association with total decryption time.

7 Conclusion and Future Scope

Thus, we have analysed the effect of Timing Attacks on Diffie-Hellman algorithm and have been able to observe timing differences in keys which are associated with the conditional branch taken due to differences in bits.

Future Scope: The time attack might easily be expanded to some square and multiply algorithm versions. Also processing the bits two by two, performing two square followed by a multiplication by 1, m , m^2 or m^3 , depending on whether the bits are 00, 01, 10 or 11. The attack could quite easily be adapted. Another possible work is devising an attack on ECC Algorithm with modification on Square and Multiply Algorithm.

References

1. Shamir A, Tromer E, Genkin D (214) RSA key extraction via low bandwidth acoustic cryptanalysis, *Advances in Cryptology CRYPTO 2014 Lecture Notes in Computer Science*, pp 444–461
2. Leroux P-A, Willems J-L, Mestr P, Quisquater J-J, Dhem J-F, Koeune F (200) A practical implementation of the timing attack, *Lecture Notes in Computer Science Smart Card Research and Applications*, pp 167–182
3. Schindler W (2000) A timing attack against RSA with the Chinese remainder theorem, *Cryptographic Hardware and Embedded Systems CHES 2000 Lecture Notes in Computer Science*, pp 109–124 (2000)
4. Satheesh M, Deepika M (2020) Implementation of multifactor authentication using optimistic fair exchange. *J Ubiquitous Comput Commun Technol (UCCT)* 2(02):70–78

5. Doraipandian M, Rajendran S (2021) Design of medical image cryptosystem triggered by fusional chaotic map. In: Intelligent data communication technologies and internet of things: proceedings of ICICI 2020, pp 395–409
6. Joe, Vijesh C, Raj JS (2021) Deniable authentication encryption for privacy protection using blockchain. *J Artif Intell Capsule Netw* 3(3):259–271

Awaaz: A Sign Language and Voice Conversion Tool for Deaf-Dumb People



Bharat Taralekar, Rutuja Hinge, Chaitanya Bisne, Amberish Deshmukh, and Vidya Darekar

Abstract The silent group is confronted with issues all around the world. There are numerous connection issues. Ordinary people can speak with each other through typical oral communication. People who are unable to speak or listen can use sign language as a substitute, but they notice many things when speaking with typical people. As a result, communication is always a possibility. The fact that a speech-impaired individual interacts with others through ineffective gestures is a clear impediment to this communication barrier. The paper's purpose is to bridge the gap between deaf-dumb persons and the rest of society. To convert gesture language to voice, the overall paper is built on machine learning and computer vision. The proposed work is carried out using a large self-made labeled dataset of about 15,000 images. We proposed a deep convolutional neural network for sign and gesture recognition. The model is being deployed using Tkinter GUI for creating Desktop based user application.

Keywords Sign language · Deaf-dumb · Machine learning · Computer vision · Voice conversion · Tkinter GUI

B. Taralekar · R. Hinge (✉) · C. Bisne · A. Deshmukh · V. Darekar
Department of Electronics and Telecommunication, Vishwakarma Institute of Technology, Pune, India
e-mail: rutuja.hinge19@vit.edu

B. Taralekar
e-mail: bharat.taralekar@vit.edu

C. Bisne
e-mail: chaitanya.bisne191@vit.edu

A. Deshmukh
e-mail: amberish.deshmukh19@vit.edu

V. Darekar
e-mail: vidya.darekar19@vit.edu

1 Introduction

Sign languages are still very much alive. The most widely used sign languages in the country are ASL (American Sign Language), ISL (International Sign Language), ISL (Indian Sign Language), and MSL (Malaysian Sign Language). These languages were designed and created with a lot of hard work and realistic study for people who are deaf or deafeningly deaf. First, a phrase and a definition. This language employs sign and action. The phrase here will not help people understand the meaning of a symbol. We can't teach them certain terms because they were born deaf. It enables structures to continuously learn and build on their knowledge without the need for explicit coding. The term "artificial intelligence" refers to computer programmes that can analyze personal data. The goal of gesture identification is to recognise and exploit individual human gestures for the transmission of knowledge. A camera records and transmits data about body motions to a device that employs gestures as input to monitor devices or apps. The goal of creating an interaction is to extract data through the use of hand motion recognition.

Because we need to extract relevant sign gesture features from hand gestures, there are enormous irrelevant regions in every frame, making it difficult to extract representative required sign gesture characteristics for hand motions. In order for traditional classifiers to work successfully, the extracted features need to have strong descriptors. These descriptors encode enough information to account for the temporal dependency of the inter-frames, as well as the hand position, shape, and orientation in each frame. The computed characteristics should be able to reduce the impact of various scenarios such as background clutter and geometric distortion. As a result, we used deep learning as a potential option in our work.

2 Background and Related Work

Using 2D video footage and 3D motion data from 14 anomalous and 10 normal movements, this research presents a novel method for generating a hand gesture database. Whereas, based on Indian Sign Language, we have created our own data set (ISL) [1]. Examine the feasibility of categorizing hand motions based on the difference in impedance between two antennas using transfer learning. Because of near-field disturbance, when a hand gesture is done near an antenna, the impedance varies over time. Recognize the impedance fluctuation pattern to determine the gesture. To measure signatures at diverse places, we recommended using two 2.4 GHz monopole antennas. The use of two antennas has increased the system's cost and complexity [2]. To make the detector more robust, AdaBoost-based hand-pose detectors are trained with a smaller collection of Haar-like features. The proposed solution, which is based on context-free grammar, provides effective real-time performance. Rectangles are causing issues, therefore they've implemented an alternative representation approach for the identical gestures, namely fingertip detection using the convex

hull algorithm. While the HSV colour model purports to segregate Hue, Saturation, and Brightness/Value, changes in Hue can result in substantial variations in perceived brightness [3]. Image processing is the process of performing operations on an image in order to improve it or extract relevant information from it. The statistical approach of principal component analysis (PCA) is used to minimize data dimensions or to data decorrelation. While PCA aids in dimensionality reduction, it also results in information loss [4]. The system consists of a computer-connected camera that captures photos of hand motions. Picture segmentation (skin color detection and region segmentation are performed during the image segmentation step) and a feature extraction method are used to recognize the signer's hand motions. A pre soundtrack will be played in response to identified hand gestures. The segmentation section is incapable of coping with difficult operational conditions [5]. This research looks into using micro-Doppler signatures and a DCNN to classify human hand motions. Doppler radar was used to examine the spectrograms of ten different hand gestures. The DCNN was utilized to characterize the micro Doppler properties of the hand motions. Obtaining the angle or posture of the human body's joints in FMCW radars is difficult, resulting in poor performance and a complex system [6]. This study presents a unique feature vector for capturing static hand gestures, as well as a viable technique for distinguishing dynamic hand motions using just a Leap Motion controller (LMC). There has been no mention of them in any other publications. The feature vector containing depth information is produced and fed into the Hidden Conditional Neural Field (HCNF) classifier to recognise dynamic hand gestures. The two basic operations in the system are feature extraction and classification using the HCNF classifier. Leap Motion controller, on the other hand, is unable to "look through the fingers," as when one finger covers the other. Fingers that are close together can also cause problems for the cameras, since they may not be detected individually [7]. A flex sensor-based gesture recognition module is being developed in this project to identify English alphabets and a few words, as well as an HMM-based Text-to-Speech synthesizer to convert the relevant text. Flex sensors raises project costs. Due to the change in flexibility of the sensors generate errors after a lengthy period of time [8]. Gesture recognition systems based on images are divided into three stages. For the captured image, image pre-processing includes color to binary conversion and noise filtering. Using the Convexity hull algorithm, tracking is mostly used to track a hand gesture from a captured image. Finally, tracking features such as convex hull and contour flaws are used to identify the object. For image analysis, blob detection and contour detection are utilized. For finger point detection and number recognition, the convexity hull technique is used. Sentence formation, on the other hand, is not implemented in this system, although it is in ours [9]. The user's gesture is captured by a single camera, and the image is used as input to the suggested algorithm. Segmentation, orientation detection, feature extraction, and classification are the four primary processes in the entire algorithm. It does not necessitate any sample data training. It is inefficient in terms of cost and requires frequent maintenance [10]. This paper discusses an ultrasonic 3D gesture recognition system that detects the location of a target, such as a hand, using a proprietary transducer chip and an ASIC. The downsides of optical 3D pictures for gesture recognition are their size

and power. In a real-time situation, the user has to carry a load of gloves that is tough to manage. Also this system necessitates additional maintenance [11]. To achieve marker less hand extraction, Kinect's depth and skeletal information are used. Super pixels represent the hand shapes, associated textures, and depths, while retaining the overall shapes and colors of the gestures to be identified. Based on this, the super pixel earth mover's distance (SP-EMD) is proposed as a new distance metric for measuring the dissimilarity of hand motions. This approach is sensitive to skin colour differences. The intricate and unpredictable hand traits, which pose enormous obstacles to its reliability, have also been presented as an appearance-based detection framework. The depth threshold is also chosen empirically and is prone to inaccuracies [12]. A frequency modulated continuous wave (FMCW) radar sensor detects dynamic hand movements, which are subsequently classified using a recurrent three-dimensional convolutional neural network. The network is trained to predict class labels from in-progress gestures in unsegmented input streams using the Connectionist Temporal Classification (CTC) technique, which enhances processing performance. Incorporating FMCW radar raises the price. The utilization of a wide range of frequencies and low peak power for the transmitter element of the FMCW radar also causes interference from other adjacent radio systems [13]. Gestures which are made by the user are detected using flex sensors, and corresponding messages are sent via the GSM module to the user's android device, which has the corresponding application which converts text messages into speech. Because GSM uses pulse-based burst transmission technology, it conflicts with several electronic devices. In addition, GSM has a limited data rate capability [14]. To detect gestures, a Microsoft Kinect device is employed. The user's skeleton data is tracked using 20 joints and their locations. Messages are typed and sent to the Text to Speech Converter. It's not suitable for outdoor use [15]. The flex sensors and the MPU6050 module are utilized in the hand glove, and the output is sent to the microcontroller, which uses an analogue-to digital converter to convert it to appropriate digital values. These results are then compared to database values and displayed on the mobile app. Voice output is produced by TTS [16]. The comparison of different sign language techniques is done on the basis of feature extraction, datasets and techniques for gesture recognition [17]. In this paper, morphological points on hand motions are noted, and then PNN and KNN (K-Nearest Neighbor) are used as classifiers to get the required outcomes. However, the accuracy of the KNN and PNN approaches is 87.04% and 90.74%, respectively. Which is a lot less than our method, which has a 99.85% accuracy rate [18]. For extracting features from sign language videos, a simple feature extraction method based on centroid and Discrete Cosine Transform (DCT) is proposed. This project considers ten different sign language videos. For the recognition of ten sign languages, two neural network (ANN) models have been built. When compared to CNN, which is included in our system, the ANN is less powerful. As a result, the recognition rate of the ANN-based system is only 80%, which is lower than our proposed approach [19]. The suggested framework employs an approach that begins with the collection of video sequences and ends with the extraction of temporal and spatial information. "Convolutional Neural Networks" were used to analyze spatial features, and "Recurrent Neural Networks" were used to train on temporal features. In comparison

to CNN, RNN has a lower feature compatibility [20]. A smart glove records hand motions and converts them to text. This text can be delivered wirelessly to a smart phone or shown on an integrated LCD display. It is bulkier than the suggested one and requires more frequent maintenance [21]. Face identification from a distance has been a common challenge; to combat this, high resolution cameras are utilized. However, the images from such cameras have a high resolution, which slows down face detection. As a result, eye localization is used in this paper to find the face and eyes. During face recognition, a Gaussian derivative filter is also utilized to reduce feature size [27]. Pruning ELM, Incremental ELM, Symmetric ELM, and other ELM (Extreme Learning Machine) versions for various classification tasks are addressed in this paper. Because ELM is used to solve supervised learning problems, it is limited in a variety of ways, making it unable to perform some of the more difficult tasks. As opposed to unsupervised learning, supervised learning cannot extract unknown information from training data [28]. Deep neural networks are discussed in terms of speech recognition, image recognition, natural language processing, fraud detection, and disease diagnosis. Although the DNN is a useful tool, it also necessitates enormous data sets for training and a lot of computing resources. While cluster computing or cloud computing is addressed in conjunction with deep neural networks, it also demands a significant amount of training time and power consumption [29]. This research provides an approach for producing a less weighted collection of models using bilinear convolution neural networks, which leads in superior visual identification in remote sensing photos using fine-grained models. However, the model is quite massive and complex in comparison to our system [30] (Table 1).

3 Proposed System

Figure 1 depicts the Recognition of Hand Sign Gesture through Live Webcam architecture for detecting hand gestures using a live camera. We propose a system which makes advantage of ML algorithms, specifically CNN. Our proposed model will be trained on approximately 17,000–19,000 photos and with increasing epochs to improve accuracy. The reason we used CNN would be that it contains numerous layers, which enables us to easily train the system. We are using a 2D CNN model. Simultaneously, we used Open Computer Vision Technology to interface with the camera and obtain live input from it. We created and defined various signs using photos, which are educated using an algorithm. The individual will be allowed to accomplish a sign in front of the camera. The sign will be identified after receiving live input from the camera. Recognized signs will generate text, which will first be converted to audio sound. The overall trained model is then connected with the tkinter GUI. The GUI is completely based on python programming.

Table 1 Various techniques used in gesture recognition system

| Sr.no | Title of paper | Signs used | Segmentation and feature extraction techniques | Classification and recognition method |
|-------|--|---|---|---------------------------------------|
| 1 | Machine learning techniques for Indian sign language recognition [26] | 25 Indian sign language alphabets | PCA for dimensionality reduction | KNN classifier |
| 2 | A signer independent sign language recognition with coarticulation elimination from live videos: an Indian scenario [24] | 26 Indian sign language and 11 words | Extraction of region of interest using a 6-dimensional feature vector with coarticulation removal | SVM classifier |
| 3 | A novel mathematical modeling and parameterization for sign language classification [22] | Pakistan sign language alphabets and 0 to 9 numbers | Curve fitting technique and centroid estimation | KNN classifier |
| 4 | Conversion of Sign Language into Text [25] | 26 Indian sign language alphabets | Otsu segmentation algorithm with morphological filtering | LDA algorithm |
| 5 | American sign language recognition using multi-dimensional hidden markov models [23] | 26 American sign language alphabets and 36 handshapes | The Linde-Buzo Gray (LBG) technique was used to quantize a 21 dimensional feature vector | Multi-dimensional Hidden Markov model |

3.1 Database Creation

Everyone contributed to the creation of the database. We separated our data into five groups. As shown in Fig. 2 we built the dataset from A to Z and a null set of photos for the null value. During the database's construction, the limitation was that each image's background had to be white and black. For each set, we must give the number of samples. The database has approximately 17,000 to 19,000 photos, each of which contains 600 to 700 photographs of 27 samples.

3.2 Database Preprocessing

Because we took photographs of the hand move, we first clipped the necessary images from the original images (as shown in Fig. 3). As shown in Fig. 4, we used the

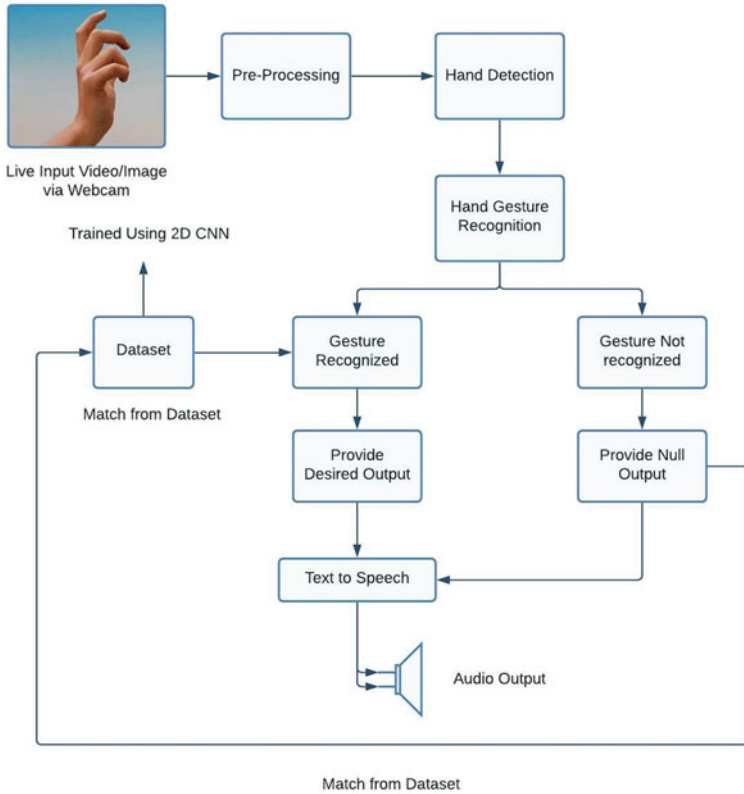


Fig. 1 Recognition of hand sign gesture through live webcam

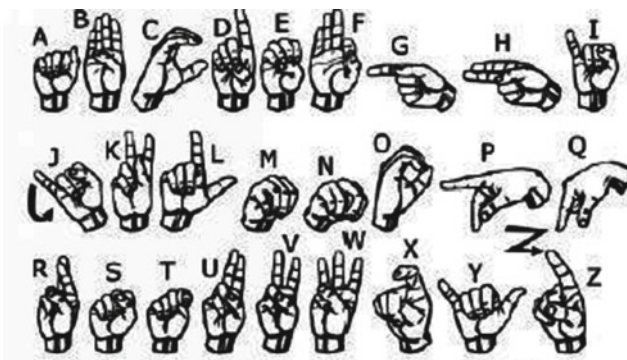


Fig. 2 Dataset



Fig. 3 Cropped image



Fig. 4 Pre-processed image

Gaussian filter since we needed the photos to be black and white. The Gaussian filter is used as a pre-processing technique to smooth the image and remove all irrelevant noise. During filtering, we looked at Intensity and Non-Maximum suppression, which are used to remove erroneous edges. After applying the Gaussian filter, we used double thresholding to get a firm edge on each image for improved pre-processing. As a result, all weak edges are eventually deleted, and only the strong edges are examined in the subsequent phases.

Figure 3 depicts a pre-processed image with feature extraction. This image was then sent to be classified.

3.3 Database Splitting

Classification of Dataset

To begin, the dataset is set up using alphabets from A to Z and a null value to provide space between two words or letters. As illustrated in Fig. 5, the dataset is divided into two parts: training data set and testing data set. The testing dataset has 100–150 photos for each of the 27 classes, totaling 3,780–4,000 images, whereas the training dataset

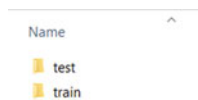


Fig. 5 Classification of dataset


```
{'0': 0, 'A': 1, 'B': 2, 'C': 3, 'D': 4, 'E': 5,
```

Fig. 6 Labeling of dataset

had 400–500 images for each of the 27 classes, totaling 12,000–15,000 images. We used somewhere between 17,000 and 19,000 photos in all.

Labeling the dataset

Python programming is being used to label the dataset. The application is given the dictionary location where the dataset is kept. Our dataset has been labeled as a class. The data for A pictures is labeled as 1, and all 27 samples are labeled from 0 to 26 (Fig. 6).

3.4 Training of Dataset

Figure 7 depicts 2D convolutional neural network structure. CNN (convolutional neural network) is a deep learning neural network type. Consider CNN to be a machine learning system that can take an input image, give importance (learnable weights and biases) to various aspects/objects in the image, and differentiate one from the other. CNN works by pulling information from videos. A CNN is composed of the following elements:

- The input layer is a grayscale image.
- The output layer is made up of binary or multi-class labels.
- The hidden layers are made up of convolution layers, ReLU (rectified linear unit) layers, pooling layers, and a fully connected Neural Network.

As shown in Fig. 8, The model type we employed is Sequential. Layer by layer, we construct our model. We used the ‘add()’ function to add our models. To construct our model, we employed three layers. Our three layers are Conv2D layers. These layers deal with our 2-dimensional matrices as input images. Each layer has 128 nodes in the first layer, 96 in the second layer, and 64 in the third layer. The Rectified Linear Activation function was utilized for the first three levels (ReLU). ReLU is a linear function that returns the value of the input if it is positive (i.e., >0), else it

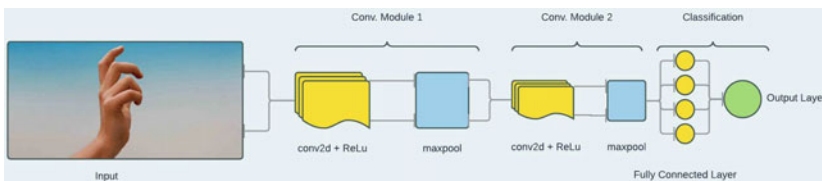


Fig. 7 2D Convolutional neural network based structure

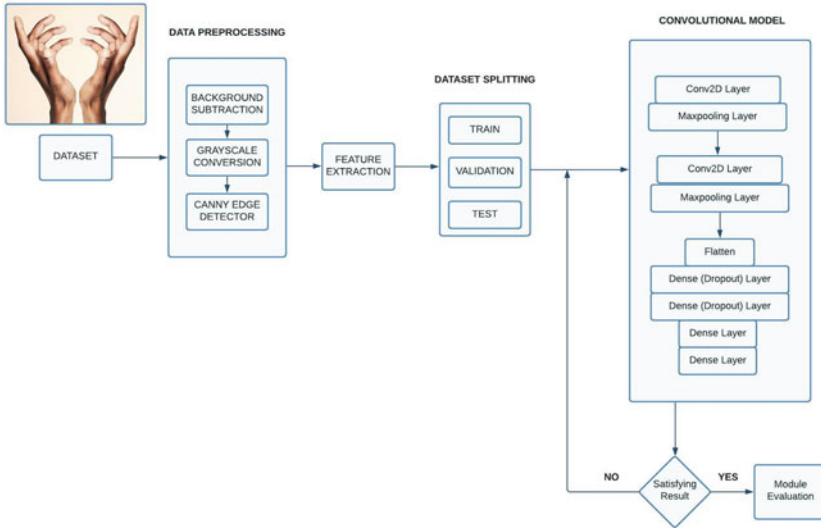


Fig. 8 Dataset training architecture

returns zero.

$$\text{ReLU}(x) = \max(0, x) \tag{1}$$

A two-dimensional convolutional neural network is used to train the dataset. The prior data is sent to the softmax layer for classification and training of the dataset. Equation defines the softmax function (1). The class with the highest probability will be the expected output.

$$\text{SoftMax}(xi) = \frac{e^{xi}}{\sum_{j=1}^k e^{xi}} \tag{2}$$

xi = corresponding class

k = number of classes

The categorized data is then made available for epoch modeling. For more precision, we have used a total of 20 epoch numbers. The data provided is being trained for 20 epochs. The classifier model is saved in.h5 format.

The trained model is 99.85% accurate.

3.5 Hand Gesture Recognition

The final step is to offer live webcam video to validate the trained model’s accuracy. The software is then given the classifier model. The following steps are taken to recognise sign gestures:

- i. Images from a live webcam are supplied.
- ii. A window is created to display hand gestures.
- iii. When a hand gesture is identified, it is compared to a similar image in the collection.
- iv. When the image is recognised, the text that relates to that image is displayed.

3.6 Text to Voice Conversion

The identified photos will be used to generate text from the given input. The text will include words from the A to Z alphabet that will form a letter or a sentence. The statement is subsequently passed to Google Text to Speech (gtts API), which converts it from text to speech.

3.7 Tkinter Desktop GUI

Python provides numerous possibilities for designing graphical user interfaces (GUIs) (Graphical User Interface). Tkinter is the most commonly used of all the GUI techniques. It is a standard Python interface to the Python-supplied Tk GUI toolkit. Python with tkinter is the quickest and easiest approach to construct graphical user interface (GUI) applications (Fig. 9).

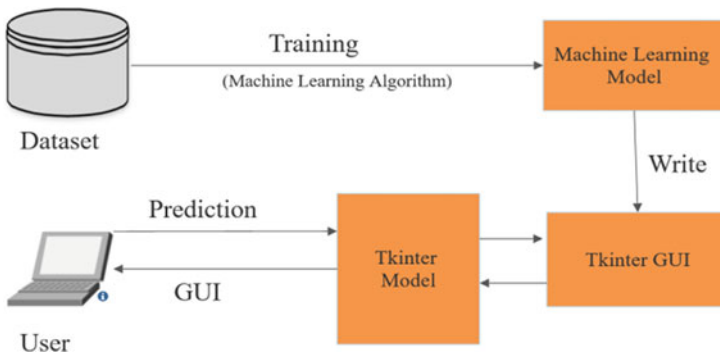


Fig. 9 Tkinter GUI model

We have already completed the training of our dataset and model for hand gesture detection using 2D CNN. For generating a user-friendly desktop application, our model is deployed using the tkinter Graphical User Interface (GUI). The following are the steps for designing a GUI with Tkinter:

- i. Installing necessary prerequisites
- ii. Creating the main window with the necessary buttons and functions
- iii. Thread-based Model Deployment

4 Result and Analysis

4.1 Comparison Between 2D CNN, KNN, SVM and Naive Bayes Classifier Algorithms

Using our dataset, we examined and contrasted various modern technologies with our 2D CNN model. We developed the algorithms K-nearest Neighbor (KNN), Support Vector Machine (SVM), and Naive Bayes. By examining the confusion matrices which are shown in Figs. 10 and 11, we calculated the accuracy, precision, f1-score, and recall of the aforementioned techniques and compared them to our 2D CNN model (as shown in Table 3). As a result, the overall accuracy of various algorithms is shown in Table 2. As can be observed, the accuracy for 2D CNN model is 99.85%, for the KNN is 97.80%, for SVM is 98.76%, and 75.50% for Naive Bayes Algorithm. For 2D CNN it has the highest accuracy when compared to other models, at 99.85%.

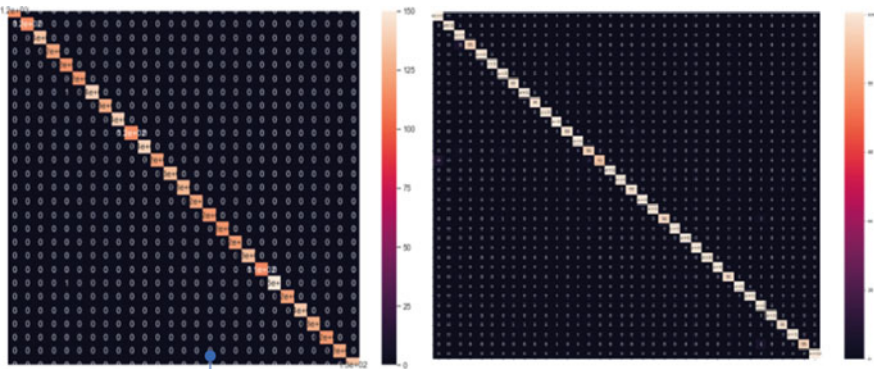


Fig. 10 Heat map representation of Confusion Matrix for CNN and KNN model on Indian Sign Language Dataset

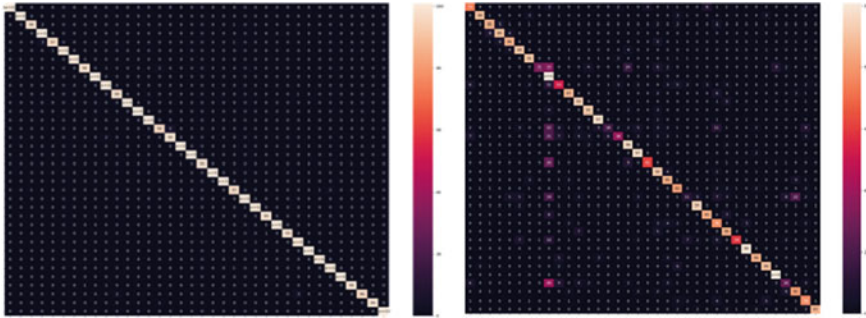


Fig. 11 Heat map representation of Confusion Matrix for SVM model and Naive Bayes model on Indian Sign Language Dataset

Table 2 Comparison between CNN, KNN, SVM, and Naive Bayes

| The proposed method | Accuracy (%) | Sign language | The sign |
|---------------------|--------------|----------------------|---|
| 2D CNN | 99.85 | Indian Sign Language | A to Z and a null value for providing space between letters |
| KNN | 97.80 | Indian Sign Language | A to Z |
| SVM | 98.76 | Indian Sign Language | A to Z |
| Naive Bayes | 75.50 | Indian Sign Language | A to Z |

Table 3 Accuracy, precision, F1-score, recall values for classification algorithm

| Classification Algorithm | Accuracy (%) | Precision | F1-Score | Recall |
|--------------------------|--------------|-----------|----------|--------|
| 2D CNN | 99.85 | 0.99 | 1.00 | 1.00 |
| KNN | 97.80 | 0.98 | 1.00 | 0.99 |
| SVM | 98.76 | 0.99 | 0.99 | 0.99 |
| Naive Bayes | 75.50 | 0.92 | 0.89 | 0.86 |

4.2 Loss and Accuracy of 2D CNN Model

We have plotted our model accuracy and training loss using matplotlib library as shown in Fig. 12.

4.3 Sign Gesture Recognition

Figure 13 depicts the detection and recognition of Gesture for letter C via the machine’s webcam. The model will detect the hand gestures only in the green box.

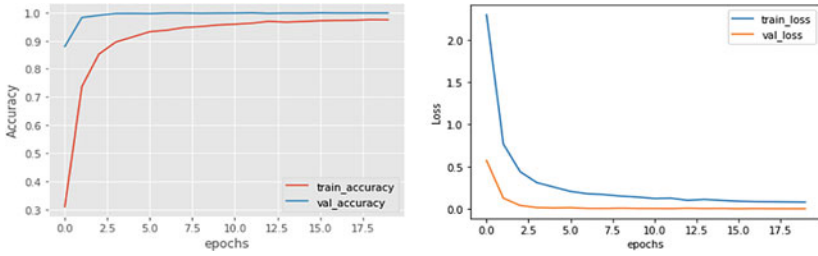


Fig. 12 2D CNN based structuring training loss and accuracy on dataset

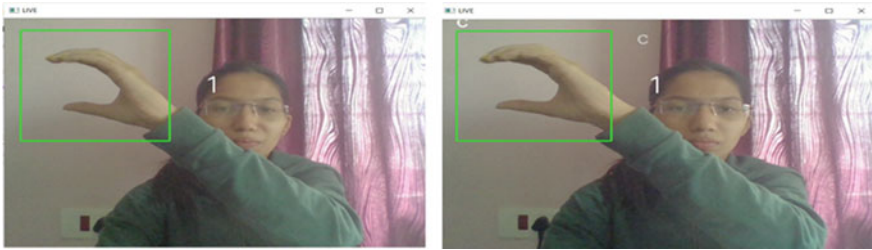


Fig. 13 Detecting and recognition of gesture for letter C via live webcam

As shown in Fig. 14, the model can discriminate between words and blank spaces. As we can see, the system recognises the letter A. To place a gap between two words, we must wait 2 s without making any hand gestures to the camera; after this time, the system will automatically offer a space.

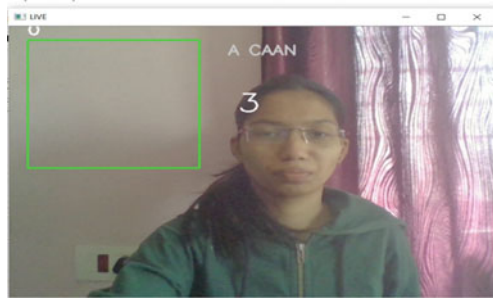


Fig. 14 Detecting a word and a space



Fig. 15 Main page of GUI

4.4 Tkinter GUI Desktop Application

Figure 15 depicts the user interface’s main page. The following buttons are available on the page:

- i. System Start: This button is used to boot up the system. When the button is touched, the live webcam begins recording the user’s hand gesture.
- ii. Result: The result icon displays the results of the hand gesture.
- iii. Crop Image: In this window, the hand gesture will be shown in black and white.
- iv. Play Sound: After completing the hand movements, pressing the play sound button will transform the relevant text result into sound and allow the user to listen to the system’s output voice.
- v. Quit: Pressing this button will cause the system to exit the main loop.

Figure 16 shows the method of making a word with hand motions. Only the green box recognises the gesture. The gesture has been transformed into black and white format, as shown in the figure. After that, the transformed image is compared to the photos contained in the dataset. If the image is matched, the result will be displayed

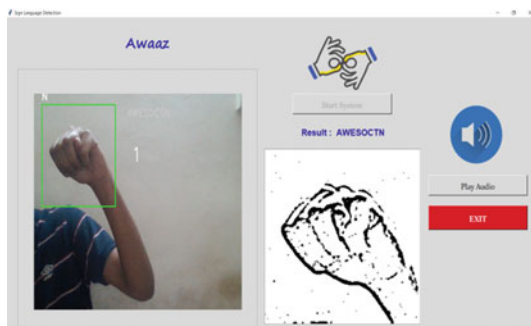


Fig. 16 Hand gesture recognition using desktop GUI

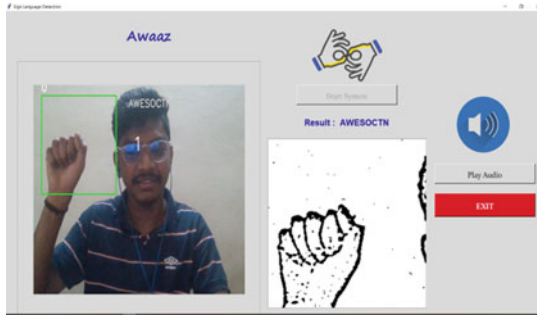


Fig. 17 Voice conversion result

in text format; if the image is not matched, the system will display the null value, i.e. space.

After we have finished the word, we may listen to it by pressing the “Play Audio” button in the GUI, as seen in Fig. 17. When you press the play audio button, the system will convert text to speech using Google’s Text to Speech (gTTS) model.

5 Conclusion

Communication between deaf and mute persons and hearing people is especially challenging if the latter are not trained in sign language. As a result, they are unable to communicate effectively. As a solution to this problem, a sign language to voice conversion system is provided, which employs the CNN algorithm, which has a 99.85% accuracy rate. The text is sent through Google’s text-to-speech synthesizer, which outputs voice. Deaf dumb persons can utilise the technique to effortlessly communicate and share their feelings with others.

6 Future Scope

In the future, the research will concentrate on the interpretation of American Sign Language and International Sign Language as text or voice.

Acknowledgements We would like to convey our heartfelt gratitude to our Head of Department Prof. Shripad Bhatlawande, and Prof. Medha Wyawahare for their encouragement to work on this fantastic topic.

References

1. Hwang B, Lee S, Kim S (2006) A Full-body gesture database for automatic gesture recognition. In: 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG), University of Southampton, UK, 2006, pp 243–248
2. Alnujaim I, Alali H, Khan F, Kim Y, Hand gesture recognition using input impedance variation of two antennas with transfer learning. *IEEE Sens J*, 1–1. <https://doi.org/10.1109/JSEN.2018.2820000>
3. Gurav R, Kadbe P (2015) Real time finger tracking and contour detection for gesture recognition using OpenCV. 2015 International conference on industrial instrumentation and control, ICIC 2015. 974–977. <https://doi.org/10.1109/IIC.2015.7150886>
4. Prajapati R, Pandey V, Jamindar N, Yadav N, Phadnis N, Hand gesture recognition and voice conversion for deaf and dumb
5. Itkarkar R, Nandi A (2013) Hand gesture to speech conversion using Matlab. 2013 4th international conference on computing, communications and networking technologies, ICCCNT 2013, 1–4. <https://doi.org/10.1109/ICCCNT.2013.6726505>
6. Kim Y, Toomajian B, Hand gesture recognition using micro-doppler signatures with a convolutional neural network. *IEEE Access* 4:1–1. <https://doi.org/10.1109/ACCESS.2016.2617282>
7. Lu W, Tong Z, Chu J, Dynamic hand gesture recognition with leap motion controller. *IEEE Signal Process Lett* 23:1–1. <https://doi.org/10.1109/LSP.2016.2590470>
8. Vijayalakshmi P, Aarthi M (2016) Sign language to speech conversion. 2016 international conference on recent trends in information technology (ICRTIT), pp 1–6. <https://doi.org/10.1109/ICRTIT.2016.7569545>
9. Nikam A, Ambekar A, Sign language recognition using image based hand gesture recognition techniques, 1–5. <https://doi.org/10.1109/GET.2016.7916786>
10. Haria A, Subramanian A, Asokkumar N, Poddar S, Nayak J, Hand gesture recognition for human computer interaction. *Procedia Comput Sci* 115:367–374. <https://doi.org/10.1016/j.procs.2017.09.092>
11. Przybyla R, Tang H-Y, Shelton S, Horsley D, Boser B, 3D ultrasonic gesture recognition. digest of technical papers. *IEEE international solid-state circuits conference*, 57. <https://doi.org/10.1109/ISSCC.2014.6757403>
12. Wang C, Liu Z, Chan S-C, Superpixel-based hand gesture recognition with Kinect depth camera. *Multimedia IEEE Trans* 17:29–39. <https://doi.org/10.1109/TMM.2014.2374357>
13. Zhang Z, Tian Z, Zhou M, Latern: dynamic continuous hand gesture recognition using FMCW radar sensor. *IEEE Sens J* 18:1–1. <https://doi.org/10.1109/JSEN.2018.2808688>
14. Abraham A, Rohini V, Real time conversion of sign language to speech and prediction of gestures using artificial neural network. *Procedia Comput Sci* 143:587–594. <https://doi.org/10.1016/j.procs.2018.10.435>
15. Rajaganapathy S, Aravind B, Keerthana B, Sivagami M, Conversation of sign language to speech with human gestures. *Procedia Comput Sci* 50. <https://doi.org/10.1016/j.procs.2015.04.004>
16. Poornima NGS, Achuth AY, Anisha Maria Dsilva Chethana S (2015) Review on text and speech conversion techniques based on hand gesture
17. Safeel M, Sukumar T, Shashank S, Arman D, Rudregowda S, Puneeth B, Sign language recognition techniques- a review, 1–9. <https://doi.org/10.1109/INOCOS50539.2020.9298376>
18. Goudar R, Kulloli S, An effective communication solution for the hearing impaired persons: a novel approach using gesture and sentence formation, 168–172. <https://doi.org/10.1109/SmarTechCon.2017.8358363>
19. Desa H (2016) Sign language into voice signal conversion using head and hand gestures
20. Priyakanth R, Sai NM, Abburi R, Hand gesture recognition and voice conversion for speech impaired, *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3734777>
21. Elmahgiubi M, Ennajar M, Drawil N, Elbuni M, Sign language translator and gesture recognition. <https://doi.org/10.1109/GSCIT.2015.7353332>

22. Kausar S, Javed M, Tehsin S, Anjum MA (2016) A novel mathematical modeling and parameterization for sign language classification. *Int J Pattern Recogn Artif Intell.* <https://doi.org/10.1142/S0218001416500099>
23. Wang H, Leu M, Oz C, American sign language recognition using multidimensional hidden Markov models. *J Inf Sci Eng JISE* 22:1109–1123
24. Athira PK, Sruthi CJ, Lijiya A, A signer independent sign language recognition with co-articulation elimination from live videos: an Indian scenario. *J King Saud Univ Comput Inf Sci.* <https://doi.org/10.1016/j.jksuci.2019.05.002>
25. Mahesh Kumar NB (2018) Sathyamanaglam, conversion of sign language into text citation. *Int J Appl Eng Res* ISSN 0973-4562, vol 13, pp 7154–7161
26. Dutta K, Bellary S, Machine learning techniques for indian sign language recognition, 333–336. <https://doi.org/10.1109/CTCEEC.2017.8454988>
27. Shakya S (2021) Multi distance face recognition of eye localization with modified Gaussian derivative filter. *J Innov Image Process* 3:240–254. <https://doi.org/10.36548/jiip.2021.3.006>
28. 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(2):83–95
29. Bashar A (2019) Survey on evolving deep learning neural network architectures. *J Artif Intell* 1(02):73–82
30. Sungheetha A, Rajesh Sharma R (2021) Classification of remote sensing image scenes using double feature extraction hybrid deep learning approach. *J Inf Technol* 3(2):133–149

Decentralized Application (DApp) for Microfinance Using a Blockchain Network



C. Jyothi and M. Supriya

Abstract Blockchain is a distributed ledger for keeping an everlasting and immutable log of information about various transactions. It acts as a decentralized database and is run by processors which are part of a peer-to-peer (P2P) network. A smart contract is an agreement stored as a computer code which would be triggered inevitably when positive settings are encountered by possibly avoiding reliable third-party establishments. With blockchain technology, smart contracts could be decentralized, so that they are fair and trust less and are not controlled by one central institution. By way of consensus mechanism which requires agreement from more than 50% of the peer nodes, it is nearly impossible to hack a blockchain network. Blockchain in microfinance eradicates the need for trusted authorities, increases the speed of KYC process using digital identity creation and increases transparency. The objective of this work is to study various types of block chain networks to implement a DApp for microfinance. Microfinance usually refers to minor amounts of credit for small scale businesses and is intended for excluded customers who are unable to avail conventional sources of capital. This sector faces the challenge of increased interest rates, diminished service availability, sluggish KYC procedure and disproportionate transaction charges. The intent is to enable farmers directly procure a loan from their prospective customers. The customer in turn will get benefitted by getting the produce at a reduced rate from the farmer.

Keywords P2P · DApp · Smart contracts · Blockchain network

C. Jyothi (✉) · M. Supriya
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita
Vishwa Vidyapeetham, Bengaluru, India
e-mail: jyothicherukunnath@gmail.com

M. Supriya
e-mail: m_supriya@blr.amrita.edu

1 Introduction

Blockchain technology is becoming very popular in today's world as it eliminates the need for a trusted central authority. Being a decentralized system, it increases transparency, security, and traceability. Statistics from Statista concludes that, "World-wide spending on blockchain solutions is expected to grow from 1.2 billion in 2018 to an estimated 39.7 billion by 2025". David Chaum an American cryptographer and computer scientist in his 1982 dissertation "Computer Systems Established, Maintained, and Trusted by Mutually Suspicious Groups" proposed a blockchain like protocol. With complete code to implement the protocol, Chaum's dissertation proposed all but one element of the blockchain, proof-of-work which was later detailed in the Bitcoin whitepaper by Satoshi Nakamoto. In 1991, Stuart Haber together with Scott Stornetta presented a work on Time-Stamping a Digital Document which suggested in calculating hash value of document and then timestamping it to save it. Here, the transactions are connected in a data structure which includes hashes of preceding transactions. In 1992, Haber, Stornetta, and Dave Bayer in an attempt to enhance its effectiveness, merged Merkle trees into the earlier proposal, by permitting numerous file certificate additions to a single block. This was later implemented in 2008 by Satoshi Nakamoto and resulted in the blockchain revolution.

Microfinance is the synonym of a financial service offering loans to small businesses and is intended for excluded customers who are unable to avail the benefits of conventional sources of capital meant for businesses of smaller size. Excluded customers here mean, people who are geographically separated or are financially weak. This term was promoted by Muhammad Yunus, who is mentioned to as the 'banker of the poor' and also won the Nobel Peace Prize in 2006. He founded the foremost micro-credit establishment, the Grameen Bank in 1976 in Bangladesh. Although it is very beneficial for the excluded customers, it comes with its own demerits. Some of the challenges faced in this domain are increased interest rates, excessive transaction charges, slow KYC process and decreased availability of services.

Blockchain is a disseminated ledger for maintaining everlasting and immutable log of transactional information where each node in the P2P network has a copy of the ledger. Only the header of the block is visible to everyone whereas private data can only be seen by the owner. Ledger is maintained as append only which can be updated only after receiving consensus from all the participating nodes. The data in the distributed ledger is maintained in the form of Merkle tree. It is a tree in which the leaf node holds the data block's cryptographic hash value, and the labels of the child nodes. Blockchain was designed as a solution to Byzantine General's problem and this solution also known as decentralization. There are various algorithms to arrive at consensus in a blockchain network like proof-of-work, proof-of-authority etc.

Depending on the intended participants, blockchain networks are classified as public, private and consortium. Blockchain networks are of two types permissioned and permissionless. A permissionless/ public network allows everyone to connect to the system to execute contracts. Examples are, Bitcoin, Ethereum etc. Owing to a

larger node participation in public networks, consensus method based on proof-of-work are being made use of to arrange transactional logs as well as block creation. Participant identity is known and authenticated cryptographically in a permissioned network which allows logging of transaction executer's identity. Other than that, this network can have a wide range of access control mechanisms incorporated to give read and add permission on ledger data, issue transactions, control participation etc. Nodes in a permissioned network belongs to different establishments. Enterprises require complex data representations to express situations that can be maintained by means of smart contracts and also the ability to fit in varied systems avoiding the need for a centralized system.

In 1994, Smart contracts were first proposed by Nick Szabo, who introduced the notion of writing contracts in terms of computer code. They are automatically activated when certain conditions are met and are cryptographically secure as it removes the need for trusted third parties. Being an automated system, it is smoother to execute transactions unlike traditional contracts. Blockchain technology enables decentralization of smart contracts making them fair and transparent. Decentralization avoids the need for a central party such as government, banks, brokers etc. Blockchain is a distributed ledger controlled by numerous processors referred to as 'nodes' owned by dissimilar organizations. It is impractical for a hacker to hack as he needs to attack more than half of the nodes to hack the blockchain and smart contracts running on the network. Due to these characteristics, smart contracts on peer nodes can run securely and automatically and remain immutable.

The Ethereum architecture has a set of peer nodes out of which some are mining nodes and others EVM nodes. The mining nodes maintain the ledger and the EVM nodes maintain the smart contract. To invoke a function of the smart contract, message can be sent with private key signature of the invoker for deduction of ether from the account to the miner's account who has solved the cryptographic puzzle which adds the transaction to the ledger. The remix IDE is a web-based IDE which enables a smart contract developer to write Solidity smart contract on a browser, test it using a JavaScript VM to avoid the need for mining and deploy it locally. It enables interaction with the smart contract after deployment. Truffle framework is an IDE for the development and testing of Ethereum based DApp. It has some easy-to-use commands for creating and deploying Dapp and boiler plate codes of the sample applications. These sample applications can be modified based on the requirement and thereby allow reuse of code and save time spent for set up. Ganache provides a local blockchain environment to develop, test and deploy Ethereum based DApps. It comes loaded with ten accounts each with hundred ethers. This helps in testing transaction as they are associated with a cost on the real Ethereum network. The Meta-mask wallet which is available as an extension in the google chrome enables crypto currency movement from one account to another when transactions are executed in the Ethereum network.

Hyperledger Fabric referred to as enterprise-grade open-source permissioned blockchain platform which is modular with a high amount of specificity and has pluggable components. It is presently applied in numerous diverse use cases like

digitizing global trade, secure key, ever ledger etc. Container technology enables hosting of smart contract (Chain code) that comprises the logic of the application.

When user sends transaction proposal over Fabric SDK, it is directed to every endorsing peer. All the peer nodes validate and perform transactions and produce the Read and Write set as output. The response is sent back to the user. The user gathers these responses and direct it to Orderer. The Orderer arranges the transactions in ascending order and produce a block. Block is then broadcasted to committers which on successful verification add a new block in their ledger copy. Fabric security is imposed with the help of digital signatures. Requests sent to fabric, necessitates it to be signed by submitter with suitable enrollment certificates. A user's enrollment certificate is considered legal on the fabric, if it is signed by a trusted certification authority (CA). All typical CA's are supported by fabric. Rather than that, fabric provides a CA server. The two parts of a Ledger are:

World State: Every chain code has its own World state and blockchain. World state database uses key-value pairs to represent the state of the record which is preserved by the peer node. The database contains details regarding business object which could be created, updated, and deleted. Transactions are initially committed in the World state and thereafter logged in to the ledger. Version number of the state gets updated each time the state changes. Initially, the world state is empty when the ledger is just created, and a valid change in world state is logged on the blockchain also. Hence, it could be concluded that the world state could be created whenever needed provided the access to the blockchain is present.

Blockchain: A distributed ledger that records every change that has resulted in the present world state. The data structure is dissimilar as once recorded cannot be deleted and is also tamper-proof. It is an evidence on, in what way the objects reached at the present state by means of transactional history. This is organized as a chronological record of interlinked blocks, in which every block comprises some ordered transactions, which can be a query or update on to the world state.

Each block header contains a hash of block transactions, and a replica of a hash of the preceding block's header. World state is maintained using a database whereas Blockchain is realized as a file. The initial block contained in a blockchain is known as the genesis block. It is the entry point of the ledger that comprises of configuration transaction particulars encompassing preliminary state of the channel and may not carry transactional information. A smart contract describes the transaction logic which decides the lifecycle of a business object present in the world state. For deployment on to the network, it is packaged into a chain code. Chaincode decides the packaging of smart contracts for deploying to the network. Blockchain networks are used in various applications such as cross-border payments, individual identity safekeeping, anti-money laundering tracking system, voting mechanism, publicity insights, novel content conception, safe sharing of medical information, music royalties tracking, cryptocurrency exchange, real-time IoT operating systems, real estate processing platform, logistics and supply chain monitoring.

The Blockchain's enables microfinance establishments to eliminate intermediaries and trust authorities. It aids people in trusting each other to give and take financial assistance. Small scale industries and agriculturalists needing finances can take it by

using a digital identity procured using blockchain technology. This in turn increases the speed of KYC process. Smart contracts ensures that processes are streamlined, faster and more economical. This also improves the transparency and assist people and banks to view processes.

2 Literature Survey

A two-phase model is proposed by Lalitha et al. [1] for KYC verification and availing loans. In phase one, KYC verification of the client is performed. A detailed discussion on the roles of various stake holders such as social collateral group, loan officer, Credit report agency etc. and how they record the necessary information on to the blockchain has been done. In the second phase discussion on various steps to be taken for loan recovery is done. All though this work has suggested a framework to enable KYC verification, the implementation has not been performed. Wu et al. [2] surveyed 995 Ethereum DApps and looked into 29,846,075 operations over them. A detailed investigation on the increase in number of Dapps and the acceptance of Dapps is performed. Few suggestions has been given for DApp users on appropriate DApp selection, on enhancing the competence of DApps for developers, and to boost the sustenance of DApps for vendors.

A blockchain reference architecture proposed by Viswanathan et al. [3] demonstrates different perspectives with stakeholders about the overall blockchain solution. Since different stakeholders would need varying complexity levels to be injected, this would be used to prepare a progressive complexity on the same view. BSRA provides a set of architectural frameworks to build solutions based on Blockchain technologies and onboarding members to the network. A survey conducted by Kim et al. compares the different state-of-the-art DApps and foresee the required features of future Dapps [4]. It discusses the deliberations when picking a blockchain implementation and development of next generation blockchain systems. It also discusses issues such as synchronization issue, double spending issue, proof of work consensus etc. Since the Ethereum blockchain network comprises of single network, when it is united with another application, the information in the ledger may be visible to all nodes. This results in significant information in the power trading industry becoming visible to everyone. Thus, the work conclude that it is not conceivable to construct a network in the Ethereum only for power trading. Only participating organization associates of the channel have accessibility to the distributed ledger in Hyperledger Fabric. This network comprises of channels that keep separate jobs. There will be a smart contract besides a distributed ledger for every channel. The data in the distributed ledger of a channel is invisible to members in other channels.

A survey conducted by Sajan et al. [5] analyses various blockchain networks based on Byzantine fault tolerance mechanism such as Hyperledger fabric, Ripple and stellar. It also surveys various consensus mechanisms. Another study conducted by Shree et al. [6] suggests an automatic billing system to avoid long queues by enabling transparency through blockchain technologies. However the work was still

in progress and results has not been discussed. A study conducted by Ambili et al. [7] has conducted a survey on various consensus mechanism and specifically conducted a comparative study on federated consensus mechanism and proof of validation mechanism. An implementation conducted by Srivatsa et al. [8] has developed a product authentication system on Ethereum network. However this is a command line application and does not provide a client side application and has not focussed on confidentiality and availability. A proposed system by Priyanka et al. [9] has implemented a secure way to store birth certificates with the aid of Inter Planetary File System (IPFS) and blockchain networks. The application is deployed inside a container using docker-compose in order to enable a Multi-Container Docker application. A proposed system by Dhanush et al. [10] has implemented a buyer and seller's protocol on ethereum platform by way of smart contracts on blockchain networks. A discussion on the need for Blockchain networks instead of conventional storage technologies has also been conducted. The working of the proposed system has been discussed in detail.

A study conducted by Thakkar et al. [11] identifies the influence of configuration parameters such as information content of the block, endorsement rule, channels, reserve distribution, state database selection on throughput and latency of transaction. Despite that, the work also recognizes performance holdups and trouble spot. The phase two, attempts to improve Hyperledger Fabric v1.0 grounded on these findings. Several improvements like aggressive caching were proposed and projects that it gives three times enhanced performance. The study suggests verification of endorsement policy in a parallelized manner can give up to seven times improvement in performance. In addition, improved and gauged the consequences of usual bulk reads and writes in CouchDB while in state validation and commit stage which resulted in two and a half times improvement. By merging the three optimizations, the total throughput was enhanced by sixteen times.

Table 1 highlights the above findings of the literature survey and presents it precisely for a comparative study.

This paragraph presents the additional survey for a deeper understanding. A study conducted by David Lee Chaum presented the notion of blockchain [12]. It suggests a way for organizations who do not trust each other to come to create and preserve a highly-secured computer system that all of them can trust by agreeing on a practical proposal. Cryptographic techniques aids in protecting information while storing and communicating. A mechanism, named vault is to be protected. Once examined and wrapped, any effort to see the vault will result in destruction of the data. A verdict by a set of trustees can let a physically destroyed vault to be re-established securely. A study by Haber et al. [13] proposed two solutions to time stamp digital documents, one using unidirectional hash functions, the results of which are treated with the real data as well as digital signatures. As per first method, the hashes of documents send to TSS will be chained, and certificates showing linkage of documents are sent to all customers above and below in the Merkle tree of that document. In the second approach, some of the members of the customers should time-stamp hashes of the document. Both these approaches differ in the method by which timestamp is made tamper proof.

Table 1 Literature survey findings

| References No | Domain | Nature of work | Type of blockchain |
|---------------|-------------------------------|-----------------------------|--|
| [1] | Finance | Architecture proposal | – |
| [2] | General | Theoretical study | – |
| [3] | General | Architectural framework | Hyperledger Fabric |
| [4] | Smart grid | Architecture proposal | Hyperledger Fabric |
| [5] | General | Theoretical study | Hyperledger fabric, Ripple and stellar |
| [6] | Billing system | Architecture proposal | – |
| [7] | General | Theoretical study | – |
| [8] | Product authentication system | Command Line Implementation | Ethereum Network |
| [9] | Storage of birth certificates | Implementation | – |
| [10] | Buyer and seller's protocol | Implementation | Ethereum platform |
| [11] | General | Theoretical study | – |

A study conducted by Satoshi Nakamoto [14] suggests that a peer-to-peer form of e-cash will help to send money online avoiding the need for a third-party organization with the aid of digital signatures and peer nodes. A discussion on proof-of-work consensus mechanism has been done. A study conducted by Smys et al. [15] has incorporated a blockchain network to accomplish a secure smart vehicle system. Although many security issues are taken care the work has not addressed how to protect the system from benign entities which exists internally. The proposed work has suggested incorporating secure data transmission as an enhancement for the system. A smart farming system proposed by Siva Ganeshan et al. [16] suggests the usage of blockchain networks for improving the data security. A work by Christy et al. [17] suggests the usage of blockchain network along with an IOT platform for real time analysis of patient data and secure automation of medical records by insurance agencies and hospitals.

Based on the literature survey, it could be concluded that there are not many implementation-based studies on blockchain based Dapps especially on permissioned networks. Hyperledger fabric based DAPPS will be a better choice as it will avoid malicious participants and transaction cost on Ethereum networks and will thereby eliminate the need for a crypto wallet for the farmers. The proposed system was built on Hyperledger fabric based blockchain after a thorough analysis of both Ethereum and Hyperledger fabric based smart contracts. Section 3 discusses the details of the system design and architecture.

3 System Design and Implementation

Visual Studio has an extension named IBM Blockchain Platform which provides a local fabric blockchain network. The extension assists programmers in creation, testing and debugging of smart contracts, join fabric environments, and develop applications which perform transactions on to the blockchain network. Once tested, these applications can be deployed to IBM cloud. This platform enables creation of a new smart contract project by giving the name of the asset to be tracked. A skeletal smart contract will be generated in the language of preference. This can be further modified based on the functionalities required.

3.1 System Architecture

The architecture diagram of the full stack Microfinance DApp is shown in Fig. 1. This is a MVC based architecture comprising of model which is the Order fetched from the blockchain. The View is the Embedded JavaScript (EJS) files rendered on the client's browser. The controller is the request handler function running on the express .js server.

The application flow is as follows. The client sends the request to server. The server takes the request parameters and with the aid of a request handler function running on it, connects to the blockchain network with the Gateway connection information and the client's wallet stored on it and fetches or adds data to it. The response is prepared based on this data and sent back to the client as http response. EJS template engine enables rendering of dynamic data on to the client's browser.

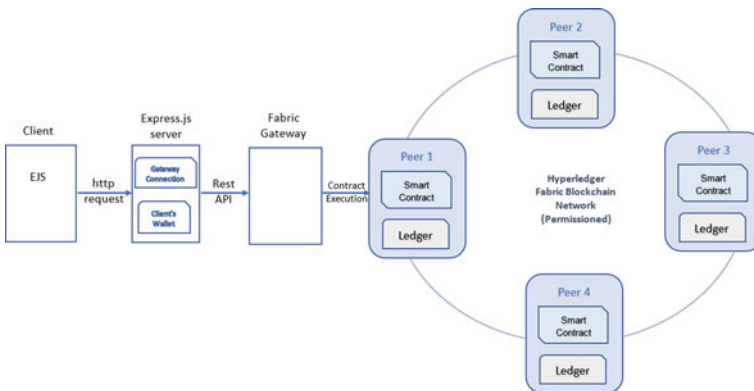


Fig. 1 Architecture diagram

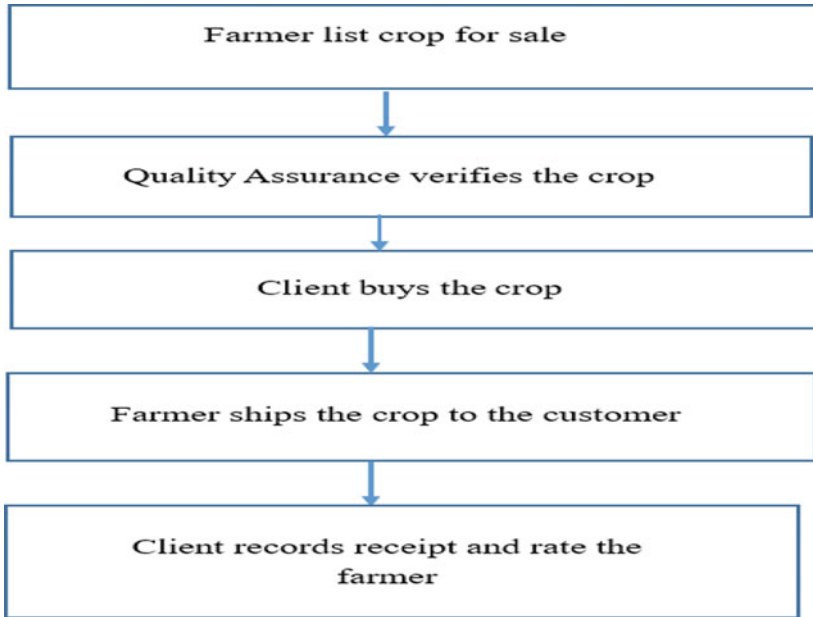


Fig. 2 Flow chart of daap

3.2 Flow Chart

The flow of the full stack Microfinance DApp is shown in Fig. 2.

3.3 Smart Contract Functionalities

The smart contract which contains the business logic of the application was implemented by extending the Contract class available in fabric-contract API. The various functionalities provided by smart contracts are as follows.

- **List Crop-** The farmer identifies himself with his wallet and list the crop for which he expects finance from the client to grow.
- **Verify Crop-** The Quality Assurance team can be a government authorized body such as the local agricultural officer can visit the site and verify the crop details elaborated by the farmer
- **Buy Crop-**The Client can buy the crop which is verified by the quality assurance team and assigned a verified date. The agricultural officer can arrange the finance to the farmer with the aid of accounts linked to the digital identity of farmer and customer

- Ship Crop- Once the crop is ready the farmer can issue a ship request which can be taken up by a shipping company and send to client location.
- Receipt of Shipment -The receipt of shipment can be recorded by the client by giving a received date.
- Rate the farmer- The client can rate the farmer based on the quality of the product received. This allows more transparency and avoid the need for third parties like banks and middlemen.
- Query Order History- There are two functions available to query. One is query order history based on item to buy and the other one is to query order history based on seller.

The Front-End- Hyperledger fabric doesn't provide support for REST API, hence a server based on express .js is introduced to receive http request and to handle them using REST API. The server in turn connects to the fabric gateway with the help of the connection information and the user's wallet. Fabric's high-level APIs packaged as fabric-network has classes such as Gateway, Wallet etc. enables this. With the aid of EJS as the template engine, the dynamic data is displayed on to the client's browser using the data fetched from blockchain.

The results are discussed in Sect. 4.

4 Results and Discussion

The study was conducted in two parts. The first part consisted of setting up sample applications in both Ethereum and Hyperledger fabric to evaluate their applicability. The second part consists of implementing Microfinance DApp on Hyperledger Fabric. Smart contract methods were implemented to record transactions on to the blockchain network. The Hyperledger fabric chain code is tested using transaction view available in the local set up. A screen shot of the sell transaction executed through transaction view of IBM blockchain platform extension for VS Code is shown in Fig. 3.

There are three main type of users the farmer, verifier, and buyer. The farmer logs in to either create a new sell order or to ship an order which is ready. Once he performs these actions, he can see the order history report showing that his transaction has been added to the blockchain. He has an option to log out once the necessary action has been taken. Once logged in, the verifier gives the seller number of the order he wants to verify. Then he verifies the order giving the necessary details. Upon submitting this, he gets an order history report which reflects the verified date added. Then he gets an option to log out or verify more orders. The buyer logs in and query the item he wants to buy. Then he gets the order history report of the required item. He selects a verified order and place a buy request. The buy date gets recorded on the blockchain. This will show the order history report with the buy date updated. Once he receives the order, he can confirm receipt by logging the received date. He also has an option to rate the order based on his experience as a reference information

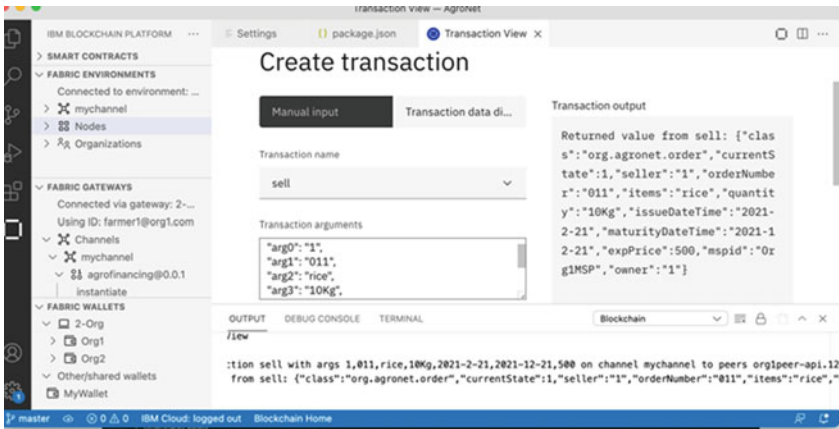


Fig. 3 Sell transaction executed through transaction view

for future buyers. This also gives the buyer an option to log out once the action is completed. The completed order history report as can be seen from buyer's login is as shown in Fig. 4

The various modules of the system are.

Farmer Module- Display pages to select the action to be performed, create order, query order history, ship Order etc.

Verifier Module- Display pages to query the order history of seller to be verified, verification screen for the verifier etc.

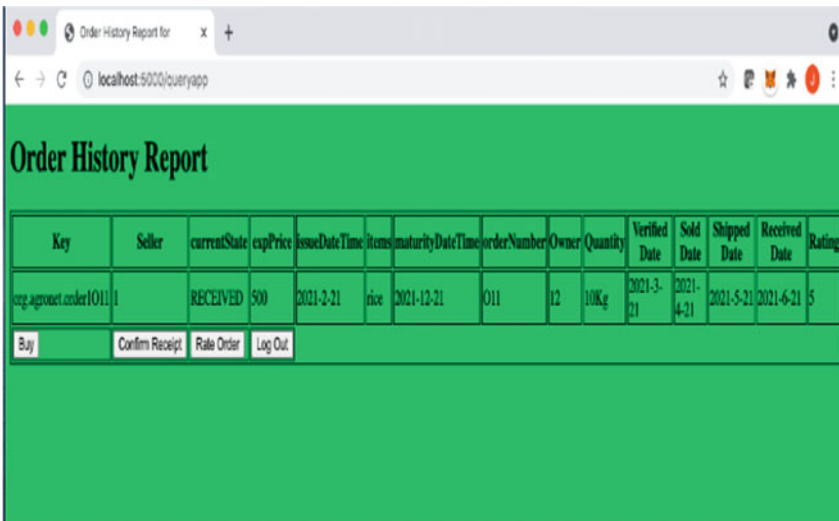


Fig. 4 Order history report

Buyer Module- Display pages to query the item to be bought, buy, record receipt and rate order.

Based on the literature survey Table 1 it was concluded that there are not many implementation-based studies on blockchain based Dapps especially on permissioned networks. This work has addressed this by implementing a Microfinance Dapp with a user interface on Hyperledgerfabric based block chain network.

5 Conclusion and Future Scope

For a comparative study of the two types of block chain network, the pet store sample application in Ethereum and commercial paper sample application in hyper ledger fabric were set up. The smart contract was written in both solidity using remix IDE and Javascript using VS code extension. From the testing of smart contracts it was concluded that in a public blockchain, anyone with a crypto wallet can perform transactions and can result in malicious activity by hackers. Hence keeping in mind, the necessity for avoiding malicious activity such as price inflation, fake orders which cannot be verified etc. and to eliminate crypto wallet requirement by farmers for issuing transactions, Hyperledger fabric network was arrived at as the most appropriate blockchain in this scenario. A user interface for the Dapp was developed and tested using express .js.

In the future, the system can be extended to add more users with a database support for user information such as seller id. Last order number can be retrieved from user table each time a new Order is added to make the order number unique and will avoid manual entry. The user type can be stored in a session to give more personalized features for the DApp. Form validations can be done on server side to ensure that the data gets added to the blockchain accurately. More stake holders can be added to the application such as Shipping company, the payment agency etc. The system can be deployed on IBM Blockchain available for IBM Cloud users. This can be loaded with additional data and can be studied for performance tuning if required.

References

1. Lalitha N, Soujanya D (2019) Financial sector innovations: empowering microfinance through the application of KYC Blockchain technology. 2019 international conference on digitization (ICD), Sharjah, United Arab Emirates, pp 237–243, doi: <https://doi.org/10.1109/ICD47981.2019.9105874>
2. Wu K, Ma Y, Huang G, Liu X (2019) A first look at blockchain-based decentralized applications. *Softw Pract Exp*
3. Viswanathan R, Dasgupta D, Govind Swamy SR (2019) Blockchain solution reference architecture (BSRA). *IBM J Res Dev* 63(2/3), 1:1–1:12. <https://doi.org/10.1147/JRD.2019.2913629>
4. Kim Y, Kim K, Kim J (2020) Power Trading Blockchain using Hyperledger Fabric. *Int Conf Inf Netw (ICOIN)* 2020:821–824. <https://doi.org/10.1109/ICOIN48656.2020.9016428>

5. TIFAC-CORE, Sajana P, Sindhu M, Sethumadhavan M (2018) On blockchain applications : hyperledger fabric and ethereum. *Int J Pure Appl Math* 118(18):2965–2970, ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
6. Shree J, Kanimozhi NR, Dhanush GA, Haridas A, Sravani A, Kumar P (2020) To design smart and secure purchasing system integrated with ERP using Block chain technology. 2020 IEEE 5th international conference on computing communication and automation (ICCCA), pp 146–150. <https://doi.org/10.1109/ICCCA49541.2020.9250767>
7. Ambili KN, Sindhu M, Sethumadhavan M (2017) On federated and proof of validation based consensus algorithms. In: Blockchain published under licence by IOP Publishing Ltd, IOP conference series: materials science and engineering, vol 225, International conference on materials, alloys and experimental mechanics (ICMAEM-2017) 3–4 July 2017, Narsimha Reddy Engineering College, India
8. Srivatsa D, Aakash N, Sahishnu S, Kumar P (2020) A product authentication scheme for supply chain system using smart contract and facial recognition, at AICTE sponsored international E-conference on data analytics, intelligent systems, and information security (ICDIIS' 20), December 11–12, 2020, IOP-Institute of Physics-Journal of Physics Conference Series with volnNo 1767. <https://iopscience.iop.org/issue/1742-6596/1767/1>
9. Kumar P, Shah M (2020) To build scalable and portable blockchain application using docker. In: Soft computing: theories and applications SoCTA 2019, December 27–29, International journal of computer networks & communications (IJCNC), Singapore
10. Dhanush GA, Srivatsa D, Sivaprakash N, Selvaraj S (2019) A Buyer and Seller's protocol via utilization of smart contracts using blockchain technology. https://doi.org/10.1007/978-981-15-0108-1_43
11. Thakkar P, Nathan S, Viswanathan B (2018) Performance benchmarking and optimizing hyperledger fabric blockchain platform. 2018 IEEE 26th international symposium on modeling, analysis, and simulation of computer and telecommunication systems (MASCOTS), Milwaukee, WI, USA, pp 264–276. <https://doi.org/10.1109/MASCOTS.2018.00034>
12. Chaum, David, Computer systems established, maintained and trusted by Mutually suspicious groups, Memorandum UCB/ERL M79/10, UC Berkeley, February 1979
13. Haber S, Stornetta WS (1991) How to time-stamp a digital document. *J Cryptol* 3:99–111. <https://doi.org/10.1007/BF00196791>
14. Nakamoto S (2009) Bitcoin: a peer-to-peer electronic cash system. Cryptography mailing list at <https://metzdowd.com>
15. Smys S, Wang H (2021) Security enhancement in smart vehicle using blockchain-based architectural framework. *J Artif Intell* 3(2):90–100
16. Sivaganesan D (2021) Performance estimation of sustainable smart farming with Blockchain technology. *IRO J Sustain Wirel Syst* 3(2):97–106
17. Christy A, Praveena A, Suji Helen L, Vaithyasubramanian S (2021) A framework for monitoring Patient's vital signs with internet of things and blockchain technology. In *Computer networks, Big Data and IoT*, pp 697–708. Springer, Singapore

Experimental Study on LoRaWAN Technology Applied to Vehicular-to-Infrastructure (V2I) Communication



Samuel Alexander A. Pasia, Vince Matthew A. Rivera,
Jereme Adriane D. G. Sy, Bianca Clarisse Y. Tan, Gerald P. Arada,
and Elmer R. Magsino

Abstract This paper presents an experimental study on the suitability of Long Range Wide Area Network (LoRaWAN) in the communication and monitoring of vehicles in Metropolitan Manila, Philippines. The system has two components—the sensor nodes and the gateway. The sensor nodes send packets that contain the location of the vehicle and the received signal strength indicator (RSSI) values to the LoRaWAN gateway. The characterization of the coverage range of the gateway was conducted along a highway in an urban area. After characterization of the coverage range of the gateway was conducted, it was found that the maximum distance is approximately 700 m within the RSSI acceptable range of -30 to -120 dBm. To test the communication and possible interferences between the vehicles and the base station, two sensor nodes, one for each vehicle, and one gateway at the base station were set up. The experiments were carried out in different scenarios during sunny and rainy weather conditions. The results show that the signal from the two sensor nodes do not interference with each other as there is no significant effect on the RSSI values measured from both vehicles. The communication between the vehicles and the gateway went smoothly as the LoRaWAN gateway was placed on top of a building to maintain line-of-sight (LOS) conditions.

Keywords LoRa · LoRaWAN · LoRaWAN gateway · LPWAN · vehicular communication · V2I

S. A. A. Pasia · V. M. A. Rivera · J. A. D. G. Sy · B. C. Y. Tan · G. P. Arada (✉) · E. R. Magsino
Department of Electronics and Computer Engineering, Gokongwei College of Engineering, De La
Salle University, Manila, Philippines
e-mail: gerald.arada@dlsu.edu.ph

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
G. Ranganathan et al. (eds.), *Pervasive Computing and Social Networking*, Lecture Notes
in Networks and Systems 475, https://doi.org/10.1007/978-981-19-2840-6_9

109

1 Introduction

LoRa (short for **Long Range**) being a LPWAN has found suitability in different Internet-of-Things(IoT)-based applications due to its long-range capabilities and low energy consumption [1–3]. LoRa devices and the LoRaWAN network have been providing opportunities for smart technologies, such as smart building [4], smart waste management [5], smart agriculture [6] and smart cities [7], to thrive more. As IoT now largely supports Intelligent Transportation System (ITS), the future of smart mobility is becoming bigger and brighter. With the fast roll out of advanced telecommunications technology, like LTE and 5G, low latency and reliable communication between IoT devices and intelligent transportation systems are now highly achievable. LTE and 5G indeed complement AI and IoT [8, 9].

IoT and LoRa technology in vehicular communication are now gaining momentum in improving intelligent transportation systems particularly, in vehicle-to-everything (V2X) communication [10, 11], which encompasses vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications [12]. V2X communication help alleviate traffic and road accidents, and efficiently monitors the mobility of vehicles. In [13], the significance of optimized transmission of data in static and dynamic environment was presented and determined public transport characteristics through empirical mobility traces. While in [14], an experimental study investigated the use of LoRa technology in V2V and V2I communications by conducting field tests to measure signal-to-noise ratio, reception ratio and signal strength. V2X communication supports self-driving vehicles and smart parking systems. Reference [15] studied how to optimize the speed regulation in self-driving vehicles by utilizing metaheuristics algorithms, while [16] introduced a smart parking system that prioritizes the security of the owner and the drivers.

A study in [17] tested an RSSI LoRa system in three locations in Malaysia. An Arduino UNO board with a LoRa module including an antenna acting as the main device to communicate with the base station was used to track a moving vehicle in rural, suburban, and urban settings. The study showed that the RSSI values are noticeably lower in urban settings due the presence of more obstacles and interference in urban areas. Continuous tracking was accomplished through a handover session involving the connectivity between the node and the base stations when transitioning from one to another. Furthermore, it was found during testing that buildings significantly affect the RSSI values, specifically those that are 38m or greater in height.

This experimental study aims to test the suitability of LoRaWAN as wireless communication between the vehicles and the base station given the dynamic nature of the environment and various external factors that cause multipath fading and other interferences. Previous studies have shown the feasibility of using LoRaWAN in vehicular communication but because of the uniqueness of the geographical location and weather conditions [18] in every region or country, they yield different results. At present, LoRaWAN in V2I communication is not yet implemented in the Philippines. Hence, the results of this study may contribute to its actual implementation in the transportation industry in the future.

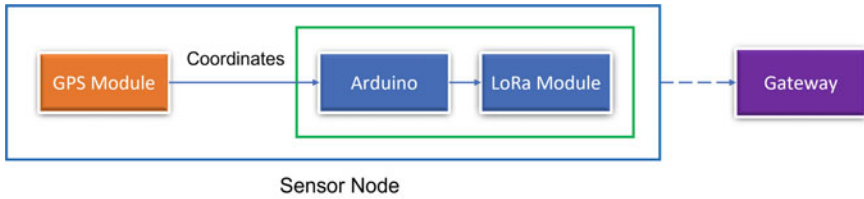


Fig. 1 The system block diagram

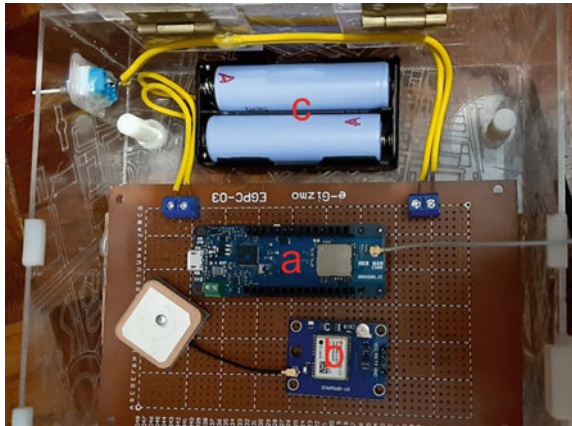


Fig. 2 The sensor node or LoRa node with a MKR1300, b NEO6M GPS module and c 18650 batteries

2 System Design and Methods

2.1 System Design

As shown in Fig. 1, the two major components of the system are: the sensor nodes and the gateway. To monitor the location of the vehicle, a sensor node was developed. The sensors attached to the vehicle collect the data and are then processed in the Arduino board with built-in LoRa module. The data will be sent wirelessly to the gateway via LoRa.

The Sensor Node The sensor node developed is consist of Arduino MKR1300, Neo-6M GPS module, and batteries as shown in Fig. 2. Arduino MKR1300 is a board that has LoRa connectivity, which allows the users to enable LoRa communication. Neo-6M GPS is the module used for determining the current GPS coordinates of the sensor node. The sensor node is powered by two 18650 1700 mAh cells and is placed inside the vehicle.

The Gateway A RAK7246 LoRaWAN developer gateway was utilized that can accommodate simultaneous receptions over 8 LoRa channels. It has already a Rasp-

berry Pi Zero W module inside which makes it ideal for prototyping and development evaluation. It covers the LoRa high-frequency band space including RU864, IN865, EU868, US915, AU915, KR920, AS923. This study will operate at the frequency plan of 868MHz as the gateway also utilizes an 868 MHz antenna, thus the software is programmed to operate at EU868.

2.2 Methods

Although the main purpose of this experimental study is to test the communication by sending data packets from the sensor nodes inside vehicles to the gateway or the base station, some minor experiments are also conducted to ensure that all the system components are functioning properly before proceeding to the major experiments.

Experiment on the Accuracy of Vehicle’s Location In measuring the accuracy of the GPS, the node was placed inside the vehicle and was driven around Macadamia Greenwoods Executive Village, Pasig City (Fig. 3). The road in the area is concrete, traffic condition is light, and the noise level is around 55 dBA since it is a residential area. While driving around the village, the authors recorded both the coordinates from the node and from a third-party app and then computed the displacement from a reference point. The third-party application used was Maps on the iPhone X. This is because the GPS points obtained from the iPhone are accurate within less than 10 ms for 98% of the time [19]. The car was only driven for short distances to determine if the GPS module in the sensor node can perform well.

Experiment on Gateway’s Coverage Range The experiment was carried out along the Southeast Metro Manila Expressway (C6) as shown in Fig. 4. The purpose of



Fig. 3 The experiment site to test accuracy of the vehicle’s location taken from Google maps

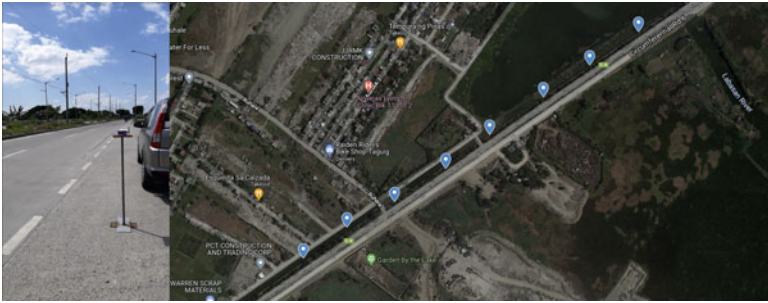


Fig. 4 The experiment site to test the range of coverage of the gateway

this test is to obtain RSSI values at varying distances to determine the range of coverage of LoRaWAN while maintaining the desirable RSSI, which is from -30 to -120 dBm [20]. The gateway is positioned at an emergency bay of the highway as the vehicle drives away from the gateway. A marker is placed at every 100m distance. The gateway was stationed at the 0m marker and positioned 4 feet above the ground. At every station (i.e., every 100ms), the sensor node sends packets containing GPS coordinates and RSSI values to the gateway. Only one packet per station was recorded to avoid causing congestion or traffic accidents on the public highway. The experiment continues until the gateway fails to receive the data.

Experiment on Communication and Possible Interferences

This is the actual experiment to test the communication and possible interferences between the vehicles and the base station or gateway. All scenarios below are done during sunny and rainy weather conditions. The gateway and the antenna are situated 12m above ground and placed on top of a building to be able to satisfy line-of-sight (LOS) conditions (Fig. 5). The whip antenna operating at 868 MHz is oriented vertically and transmitting omnidirectionally. The modulation type is a “chirp spread spectrum” that is unique to LoRa communication. The data rate is at 12.5 kbps at a spreading factor of 12. The vehicle speed is at 20 km per hour, on the average.

Scenario 1 The first scenario simulates the situation where two vehicles or cars are parked beside each other at the base station(Fig. 6a).

Scenario 2 The second scenario (Fig. 6b) simulates the event where one car is parked at the base station and another car is moving in the opposite direction. The RSSI value of the stationary car is recorded as the moving car is approaching or moving away to determine any possible interference. The RSSI of the moving car is also recorded with respect to the distance traveled away from the base station.



Fig. 5 The gateway on top of a building in order to have LOS conditions

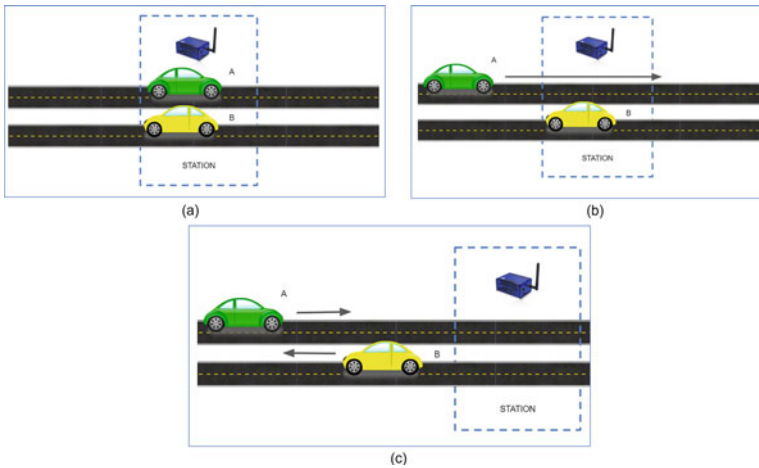


Fig. 6 The three scenarios: **a** two cars are stationary **b** one car is moving and the other car is stationary **c** two cars are moving in opposite directions

Scenario 3 The third scenario (Fig. 6c) simulates the event where cars are moving in the opposite direction. The RSSI of both cars are recorded as well as the RSSI of the point where both cars meet. The RSSI values of the cars are again recorded with respect to the distance traveled away or going to the base station. The researchers observe the effect on the RSSI after the two cars meet.

3 Experiment Results and Discussion

In this section, the pertinent results from all the experiments are hereby presented.

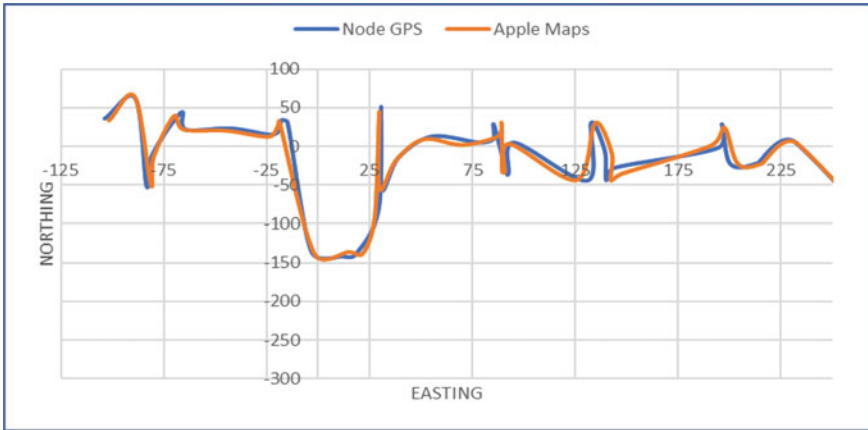


Fig. 7 Pathways recorded by the sensor node’s GPS and Maps App

3.1 Experiment on the Accuracy of Vehicle’s Location Result

Figure 7 shows the pathway taken by the vehicle along the experiment site. To better visualize the data obtained, the latitude and longitude coordinates have been converted to easting and northing coordinates, respectively. The center of the plane represents the reference points. The points in the graph were obtained from different times of the experiment. The graph easily shows that the obtained coordinates from the sensor node are very similar to that of the coordinates obtained from the third party application. The maximum difference between the points is 10.51 m while the minimum difference is equal to 0.39 m. Differences that are greater than 10m happened for 2.7% of the time, while differences that are less than 5 m happen around 77.77% of the time. Here, it was found that the node has a 95.67% similarity with the performance with that of the Maps app which further proves its accuracy.

3.2 Experiment on Gateway’s Coverage Range Result

Due to the different external factors encountered while testing, such as weather and road conditions, the obtained RSSI values were not constant. The weather conditions during the different trials varied from hot to light rain weather. As for the road conditions, traffic varied from light-to-moderate and moderate-to-heavy traffic. Additionally, the test site has a weak mobile carrier signal which resulted to a slow internet connection. Due to this, the delay between packets went from around 1 to 3 min for distances between 0 to 100 ms as compared to a 6 s delay with a fast internet speed. A total of 10 trials were conducted for this experiment. In each trial, a total of 8 RSSI values were recorded as there are 8 different stations. Thus, a total of 80 data

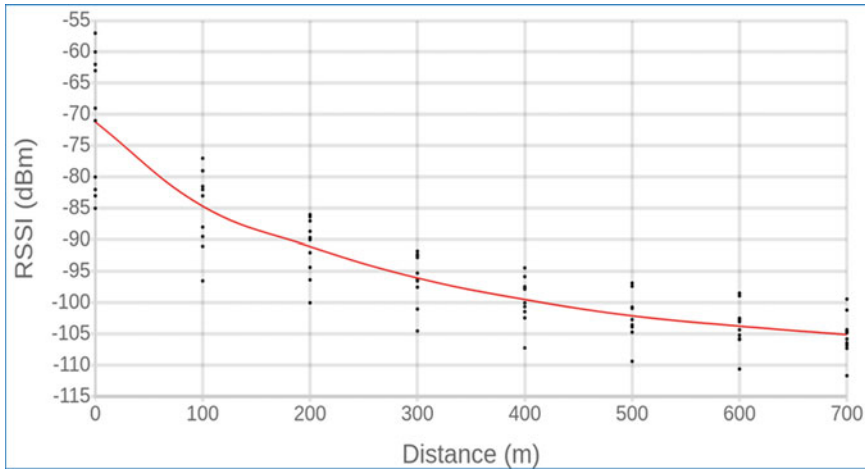


Fig. 8 The Kalman filtered plot of RSSI versus distance during the coverage range experiment

points were obtained over all. The RSSI value versus distance was then plotted in Fig. 8. Due to the adverse effects of the weather and multipath fading brought about by the obstructions surrounding the test area, some data inaccuracies were obtained. In order to reduce these inaccuracies, Kalman filtering was applied in the measured RSSI values to remove these noisy measurements. Beyond 700 m the gateway failed to receive the packets, since the sensitivity of the gateway is less than the received signal strength. The maximum distance obtained was 700 m while maintaining the desirable RSSI values of -30 to -120 dBm.

3.3 Experiment on Communication and Possible Interferences Result

Below are the experiment results on three scenarios to test the communication between the vehicles and the base station or gateway, and to test for possible interferences between vehicles or cars. The goal of this experiment is to ensure that there are no problems in receiving and sending of data packets, thereby maintaining the desired range of RSSI values.

3.3.1 Scenario 1 Result

The first scenario tested was when two cars are parked beside each other. The sensor node, located inside the car, sends one packet to the gateway containing the GPS

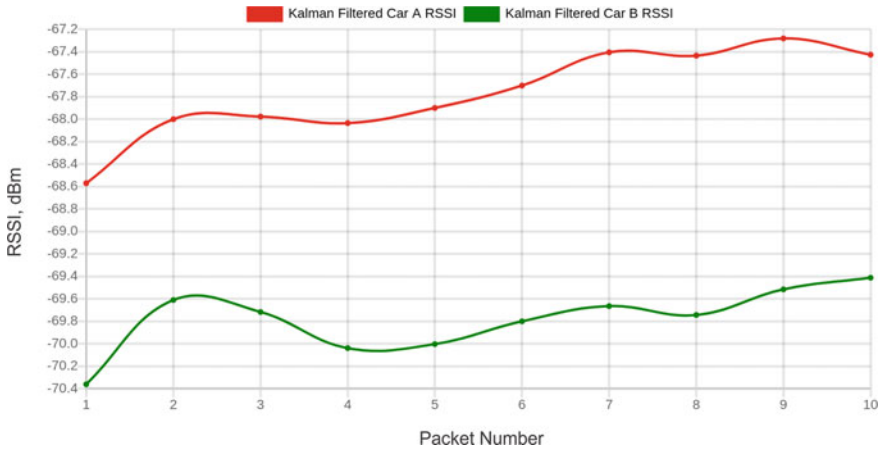


Fig. 9 Kalman filtered RSSI for 10 packets when the two cars are stationary during sunny weather

coordinates and RSSI values. A total of 10 packets per sensor node were recorded per trial. This test was done for a total of 20 trials.

For the 20 trials conducted for this experiment, the weather conditions during 6 trials were rainy, while the 14 other trials were sunny. The average of the RSSI values obtained during sunny conditions and rainy weather conditions are shown in Figs. 9 and 10, respectively. Car A obtained RSSI values from -68.57 to -67.28 dBm, while Car B obtained RSSI values from -70.36 to -69.41 dBm during sunny weather. Car A obtained RSSI values from -82.62 to -81.17 dBm, while Car B Obtained RSSI values from -73.83 to -73.34 dBm during rainy weather. It can be immediately observed that the RSSI values obtained during rainy weather conditions degraded as compared to that of RSSI values obtained during sunny weather condition. Note that the data gathered during sunny and rainy weather conditions are done in different days, hence the environment scenarios are different which affected the drastic change of RSSI values of Car A from sunny weather to rainy weather. There is no significant effect on the RSSI values when 2 cars with sensor nodes are parked next to each other, as all the data packets were received successfully. However, the obtained RSSI values may have fluctuated due to the weather conditions, multipath due to vehicles passing by, and the slow internet speed.

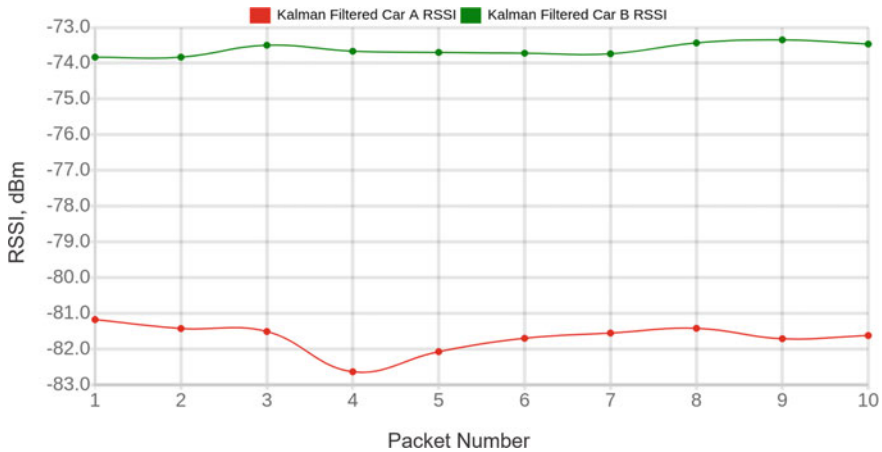


Fig. 10 Kalman filtered RSSI for 10 packets when the two cars are stationary during rainy weather

3.3.2 Scenario 2 Result

The second scenario tested was when one car is parked while the other is in motion. Each trial recorded 6 packets and 12 out of the 20 trials were taken during rainy weather condition, while the other 8 trials were sunny. During sunny weather, Car A and Car B obtained RSSI values from -73.82 to -71.75 dBm and from -80.83 to -75.87 dBm, respectively. While the RSSI values of Car A and Car B during rainy weather were from -86.16 to -80.83 dBm and from -84.95 to -75.75 dBm, respectively. The plots are shown in Fig. 11 for sunny weather and Fig. 12 for rainy weather. It can also be observed from the results of Car B that the RSSI value decreases as the packet number increases; this is mainly because Car B is moving away from the gateway. However, the result of Car A does not have any significant variation because Car A remains stationary during this experiment. It can be clearly seen that this scenario does not affect the RSSI values nor does it put the RSSI values out of the desirable range.

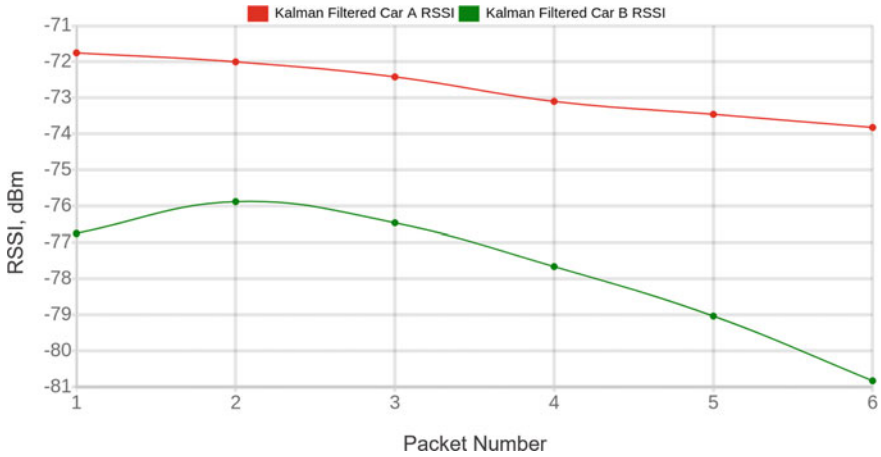


Fig. 11 Kalman filtered RSSI for 6 packets when one car is moving during sunny weather

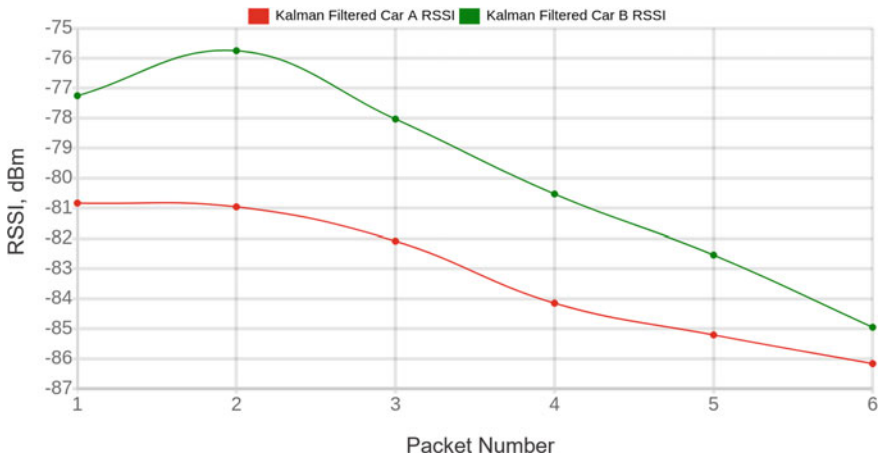


Fig. 12 Kalman filtered RSSI for 6 packets when one car is moving during sunny weather

3.3.3 Scenario 3 Result

The third scenario shown in Fig. 13 tests the two cars possible interference when they are moving in the opposite directions.

There were 8 packets obtained for each trial. 10 out of the 20 trials were during sunny weather, while the rest were during rainy weather. As shown in Fig. 14 and Fig. 15, the range of RSSI values for both cars during sunny weather are from -82.43 to -78.92 dBm, while during rainy weather the RSSI values obtained for both cars are from -99.80 and -88.39 dBm. As in scenario 2, the changes of RSSI values were due to the increasing distance from the base station or gateway.



Fig. 13 The scenario where two cars are travelling in the opposite directions

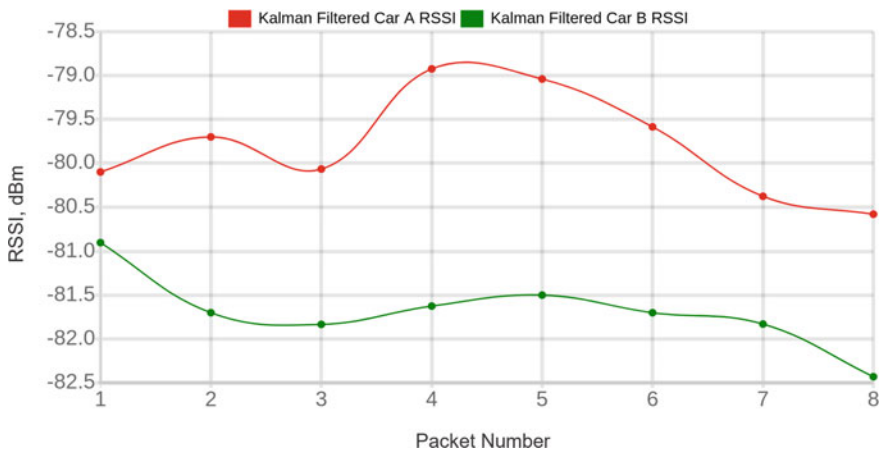


Fig. 14 Kalman filtered RSSI for 8 packets when two cars are moving in the opposite directions during sunny weather

Overall, there were no significant changes in the RSSI values obtained from both cars, which means that no interference was caused to the sensor nodes inside the vehicles. Some fluctuations or variations in RSSI values during the entire experiment were caused by varying weather conditions, slow internet speed, and multipath due to nearby buildings and vehicles passing by the test site.

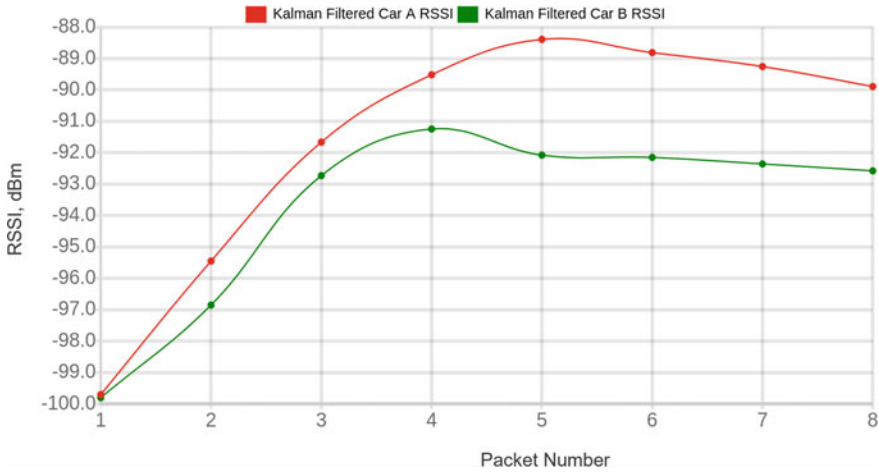


Fig. 15 Kalman filtered RSSI for 8 packets when two cars are moving in the opposite directions during rainy weather

4 Conclusion

An experimental study of a small scale set-up of LoRa network with the capability to track the mobility of vehicles was found to be suitable for vehicular communication in Metropolitan Manila, Philippines. Through several experiments and measurements, the study was able to characterize the capability of LoRa nodes and the gateway having a maximum detection range of 700ms, while the accuracy of tracking was measured to be 95.67%. The study was able to prove the suitability of LoRaWAN to V2I communication demonstrated by the successful experiments in three different scenarios, being able to maintain the desirable RSSI range of -30 to -120 dBm. The RSSI values were affected by weather conditions but not enough to put the RSSI values out of the desirable range. To implement this on a larger scale, the range of coverage can be further increased by adding multiple gateways along the route. This evaluation can also be extended to address actual problems in traffic, road accidents and vehicle monitoring in Metropolitan Manila. In the future, an experimental study on V2V communication, i.e., the communication between sensor nodes or vehicles, will be done.

References

1. Osmeni T, Ali M (2021) LoRa IoT WSN for E-agriculture. In: Miraz MH, Southall G, Ali M, Ware A, Soomro S (eds) emerging technologies in computing. iCETiC 2021. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 395. Springer, Cham. https://doi.org/10.1007/978-3-030-90016-8_6

2. Drăgulinescu AMC, Manea AF, Fratu, Drăgulinescu A (2020) LoRa-based medical iot system architecture and testbed. *Wirel Personal Commun*, 1–23
3. Fuentes AF, Tamura E, LoRa-based iot data monitoring and collecting platform. In: Nesmachnow S, Hernández Callejo L (eds) smart cities. ICSC-CITIES 2019. Communications in Computer and Information Science, vol 1152. Springer, Cham. https://doi.org/10.1007/978-3-030-38889-8_7
4. Xu W, Zhang J, Kim JY, Huang W, Kanhere SS, Jha SK, Hu W (2019) The design, implementation, and deployment of a smart lighting system for smart buildings. *IEEE Int Things J* 9(4):7266–7281
5. Sheng TJ, Islam MS, Misran N, Baharuddin MH, Arshad H, Islam MdR, Chowdury MEH, Rmili H, Islam MT (2020) An internet of things based smart waste management system using LoRa and Tensorflow deep learning model. *IEEE Access* 8:148793–148811
6. Caruso A, Chessa S, Escolar S, Barba J, López JC (2020) Collection of data with drones in precision agriculture: analytical model and LoRa case study. *IEEE Int Things J* 8(22):16692–16704
7. Premsankar G, Ghaddar B, Slabicki M, Francesco MD (2020) Optimal configuration of LoRa networks in smart cities. *IEEE Trans Indus Inf* 16(12):7243–7254
8. Smys S, Wang H, Basar A (2021) 5G network simulation in smart cities using neural network algorithm. *J Artif Intell* 3(1):43–52
9. Bashar A (2020) Artificial intelligence based LTE MIMO antenna for 5th generation mobile networks. *J Arti Intell* 2(3):155–162
10. Haque KF, Abdelgawad A, Yanambaka VP, Yelamarthi K (2020) LoRa architecture for V2X communication: an experimental evaluation with vehicles on the move. *Sensors* 20(23)
11. Haque KF, Abdelgawad A, Yanambaka VP, Yelamarthi K (2020) A LoRa based reliable and low power vehicle to everything (V2X) communication architecture. In: 2020 IEEE international symposium on smart electronic systems (iSES), pp 177–182. IEEE
12. Han B, Peng S, Wu C, Wang X, Wang B (2020) LoRa-based physical layer key generation for secure V2V/V2I communications. *Sensors* 20(3)
13. Magsino ER, Arada GP, Ramos CML (2020) Investigating data dissemination in urban cities by employing empirical mobility traces. In: 2020 IEEE 12th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM), pp 1–5. IEEE
14. Torres APA, Da Silva CB, Filho HT (2021) An experimental study on the use of LoRa technology in vehicle communication. *IEEE Access* 9:26633–26640
15. Satesh A (2020) Metaheuristics optimizations for speed regulation in self driving vehicles. *J Inf Technol Digital World* 2(1):43–52
16. Vivekanadam B (2021) Smart parking with fair selection and imposing higher privacy constraints in parking owner and driver information. *IRO J Sustain Wirel Syst* 1(1):11–22
17. Tan ZA, Rahman MTA, Rahman A, Hamid AFA, Amin NAM, Munir HA, Zabidi MMM (2019) Analysis on LoRa RSSI in Urban, suburban, and rural area for handover signal strength-based algorithm. *IOP Conference series: materials science and engineering*, vol 705
18. Elijah O, Rahim SKA, Sittakul V, Al-Samman AM, Cheffena M, Din JB, Tharek AR (2021) Effect of weather condition on LoRa IoT communication technology in a tropical region: Malaysia. *IEEE Access* 9:72835–72843
19. Merry K, Bettinger P (2019) Smartphone GPS accuracy study in an urban environment. *PLOS ONE* 14(7)
20. Vazquez-Rodas A, Astudillo-Salinas A, Sanchez C, Arpi B, Minchala LI (2020) Experimental evaluation of RSSI-based positioning system with low-cost LoRa devices. *Ad Hoc Netw* 105

Optimal Densely Connected Networks with Pyramid Spatial Matching Scheme for Visual Place Recognition



P. Sasikumar and S. Sathiamoorthy

Abstract Visual place recognition (VPR) is the procedure of identifying a formerly visited place by the use of visual details under distinct appearance conditions and viewpoint changes. With the recent advances in imaging technologies, camera hardware, and deep learning (DL) models, the VPR becomes a hot research topic among computer vision and robotics communities. The VPR is found to be a challenging process, particularly in uncontrolled outdoor environments and over long duration owing to the environmental factors, light variations, and geometric aspects. With this motivation, this paper presents optimal densely connected network with spatial pyramid matching scheme for VPR, named ODNPSM-VPR technique. The ODNPSM-VPR technique mainly aims to achieve maximum detection performance under varying viewpoints and conditions. The VPR can be considered as an image retrieval process by the inclusion of two stages namely retrieval of closest candidate images and re-ranking. The ODNPSM-VPR technique employs DenseNet-121 model as a feature extractor and the hyperparameters of the DenseNet-121 model are tuned by the grasshopper optimization algorithm (GOA). Besides, image filtering database (IFD) with Manhattan distance is used for the retrieval of top candidate images and pyramid spatial matching scheme is employed for comparing the query image with the candidate images to recognize the place. In order to demonstrate the performance of the proposed ODNPSM-VPR technique, a series of simulations were carried out. The experimental results reported the promising performance of the ODNPSM-VPR technique interms of different measures.

Keywords Visual place recognition · Spatial matching · Similarity measurement · Deep learning · Hyper parameter tuning

P. Sasikumar (✉) · S. Sathiamoorthy
Department of Computer and Information Science, Annamalai University, Annamalai Nagar,
Chidambaram, Tamil Nadu, India
e-mail: mailto:ausasikumar@gmail.com

S. Sathiamoorthy
e-mail: ks_sathia@yahoo.com

1 Introduction

Visual place recognition (VPR) is aimed at assisting robots or vision-based navigation scheme which determines whether is located in a formerly visited place [1]. It is the most challenging and essential problem in the computer vision (CV) and robotics field. In the last decade, it has witnessed an increase in the usage of VPR for different applications. For instance, in a visual Simultaneous Localization and Mapping (SLAM) scheme [2], VPR refers to loop closure detection (LCD), which is an essential element. Despite the fact, VPR has been widely investigated in robotics and computer vision research also has attracted growing interest, but still, there are several open issues to handle. In general, the challenges for VPR are twofold [3]. Firstly, false VPR renders interference to a localization approach that reduces the performance and also leads to the failure of catastrophic localization for a navigation scheme. Thus, VPR algorithm should be capable of achieving a higher detection accuracy (100% accuracy is required in few instances). Next, almost all the current VPR approaches are based on appearance. But, the appearance of a similar place might dramatically change together with distinct illumination seasons, viewpoints, conditions, distance, background, and/or occlusion clutter. Therefore, it is difficult to properly identify the same place if it endures appearance change.

At the same time, it is common that VPR method suffers perceptual aliasing problems [4], viz. about distinct places. This is due to (i) usually there are similar objects namely buildings and trees in the environment that implies images from distinct places might have same appearance, results in difficulty for detection, and (ii) the visual sensor (e.g., camera) obtains partial data of the environment, that limit the discrimination of visual information. The modern trend in VPR study is stimulated by the considerable achievement of Convolution Neural Network (CNN) in CV task.

Recently, there has been a prosperous development in visual content recognition using an effective image representation-extractor—CNN [5] that provides advanced performance on several category-level recognition tasks includes scene recognition, image classification, and object classification [6]. The fundamental idea of CNN dates back to 1980s, and the two main reasons why CNN is very effective in CV systems are the advancement in graphical processing unit (GPU) based data volume and computational power, correspondingly. The current study shows that features extracted by CNN could be generalized well [7] and transferable to other visual-recognition processes. The VPR is found to be a challenging process, particularly in uncontrolled outdoor environments and over long duration owing to the environmental factors, light variations, and geometric aspects.

This paper presents an optimal densely connected network with spatial pyramid matching scheme for VPR, named ODNPSM-VPR technique. The VPR can be considered as an image retrieval process by the inclusion of two stages namely retrieval of closest candidate images and re-ranking. The ODNPSM-VPR technique employs DenseNet-121 model as a feature extractor and the hyperparameters of the DenseNet-121 model are tuned by the grasshopper optimization algorithm (GOA).

Besides, image filtering database (IFD) with Manhattan distance is used for the retrieval of top candidate images and pyramid spatial matching scheme is employed for comparing the query image with the candidate images to recognize the place. The performance validation of the ODNPSM-VPR technique is performed using standard datasets and the outcomes are investigated in terms of different measures.

2 Literature Review

The authors in [8] introduced a VPR framework that attains better accuracy than other techniques. It is depending on an image retrieval configuration, initially, it retrieves closest candidate to a query from a database and next re-ranked the list of candidates. The latter is realized by the overview of geometric authentication process which employs the activation of pretrained CNN. Zaffar et al. [9] proposed a training-free, and compute-efficient method-based HOG descriptor to achieve an advanced efficiency for each compute unit in VPR. The motivation for this method (viz. CoHOG) depends on the convolution scanning and region-based extracted features utilized by CNN.

The authors in [10] introduced a VPR method which follows two-phase format used to image retrieval pipeline. The scheme encrypts images of place by applying the activation of distinct layers of an off-the-shelf, pretrained, visual geometry groups (VGG)16-CNN framework. Initially, a number of top candidate images, query image of place is retrieved from a formerly stored database. Next, proposed a comprehensive assessment of query image against this candidate by encoding spatial and semantic data by using CNN feature. Chen et al. [11] presented a deep distance learning architecture for VPR. By comprehensive review of the various limitations of the distance relation in the VPR problems, the multiple-constraint loss function is presented for optimizing the distance constraint relationship in the Euclidean space. This architecture could assist any type of CNN namely VGGNet, AlexNet and other user-determined network extracts distinctive features.

Tomitã et al. [12] proposed a handcrafted VPR approach such as ConvSequential-SLAM that accomplishes advanced place matching performance under challenging conditions. The study used a sequence of block-normalization and data to manage appearance changes while utilizing regional-convolution matching to accomplish viewpoint-invariance. Tsintotas et al. [13] proposed a lower complexity loop closure detection pipeline where the traversed trajectory (map) is characterized by sequence-based location. All this group of images, refer to the place, is produced online via a point tracking repeatability check applied on the perceived visual sensory data. By using an image-to-image search, the proper location is selected, and while querying the database, the appropriate candidate place is carefully chosen. Schubert et al. [14] presented a graph-based architecture for systematically exploiting distinct kinds of information and structure. The graphical model is utilized for formulating nonlinear

least square problems that are enhanced by conventional tools. Beyond standardization and sequences, we proposed the use of intra-set similarity within the database and/or the query image set as further source of data.

3 The Proposed Model

In this study, the VPR is considered as an image retrieval process which commonly takes place in two stages namely IFD process and re-ranking process. Initially, in the IFD, the reference image dataset of places is fed into the optimal DenseNet-121 model. The features from the DenseNet-121 model are saved in two individual databases namely IFD and spatial matching database (SMD). During the testing process, the query image (QI) is fed into the optimal DenseNet-121 model and the recognition process takes place in two levels. At the first level, the extracted features are compared with the features stored in the IFD using the Manhattan distance metric to rank N candidate images. Next, the QI features are compared with the features saved in SMD for every candidate image, and the optimum matching candidate is considered as the recognized place. Figure 1 depicts the overall process of ODNPSM-VPR technique.

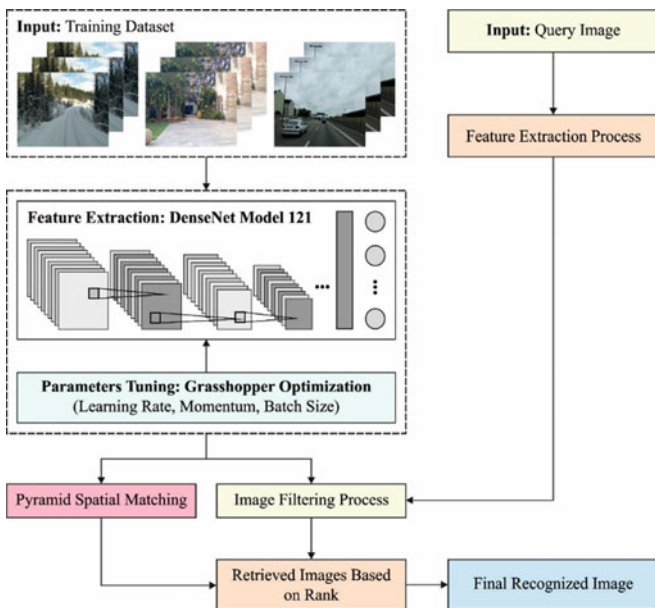


Fig. 1 Overall process of ODNPSM-VPR technique

3.1 Feature Extraction Using Optimal DenseNet Model

During the feature extraction process, the images are fed into the DenseNet model to generate feature vectors. The DenseNet model is a DL algorithm proposed on the ResNet. Recently, DenseNet has accomplished effective outcomes in the area of image classification. The fundamental concept of DenseNet and ResNet is the same; but, DenseNet established dense connections among the last and prior layers [15]. This feature makes DenseNet accomplish good results when compared to ResNet with less parameter and computation cost and alleviates vanishing gradient problems. The DenseNet is consists of Transition layer and Dense_Block. Dense_Block adapts a radical dense connection model; i.e., each layer is interconnected to one another. Especially, all the layers except the output from each preceding layer as input. In Dense_Block, all the layers have a similar size and all the layers are concatenated to each preceding layer. For network with L layer, Dense_Block comprises an overall of $L(L + 1)/2$ connection [16].

$$x_L = H_L([x_1, x_2, \dots, x_{L-2}, x_{L-1}]), \quad (1)$$

whereas L characterizes the layer count. H_L denotes non-linear transformation, that is an integration of Conv, Batch Normalization (BN), ReLU, and Pooling functions. In the study, the DenseNet-B framework is used, as well as the bottleneck layer is utilized for reducing the calculation amount; viz., the architecture $BN + \text{ReLU} + 1 \times 1 \text{ Conv} + BN + \text{ReLU} + 3 \times 3 \text{ Conv}$ is adapted. Set the channel number of input Dense_Block as k_0 , afterward the input channel number of L layer is $k_0 + k(L - 1)$. All the layers in Dense_Block output k feature map afterward convolution, viz., and the amount of convolutional kernels. Now, the last convolution layer is k , and k is named as growth rate.

As the input size of the model afterward passes with Dense_Block remain unchanged, the channel dimension would be continuously increasing. Thus, reduction dimension is needed to minimize computation difficulty. The Transition layer consists of 1×1 convolutional layer and 2×2 MaxPooling or AvgPooling layer, and its architecture is $BN + \text{ReLU} + 1 \times 1 \text{ Conv} + 2 \times 2 \text{ AvgPooling}$ layer. It interconnects 2 adjacent Dense_Blocks and minimizes the dimensionality of the output of Dense_Block. In this study, DenseNet-121 model is applied. The output vector of all the layers are treated as image descriptor \tilde{X} . The descriptors for all layers of DenseNet-121 model are attained. Prior to the usage of \tilde{X} , it undergoes normalization as a unit vector X using Eq. (2):

$$\left(\frac{\tilde{x}_1}{\sqrt{\sum_{i=1}^d \tilde{x}_i^2}}, \dots, \frac{\tilde{x}_d}{\sqrt{\sum_{i=1}^d \tilde{x}_i^2}} \right) \rightarrow (x_1, \dots, x_d), \quad (2)$$

where $(\tilde{x}_1, \dots, \tilde{x}_d)$ refers the descriptors \tilde{X} with d-dimensional vector and (x_1, \dots, x_d) signifies the normalization descriptor X with similar dimensional as \tilde{X} . This normalization process helps to be robust against varying viewpoints and conditions.

In order to optimally select the hyperparameters of the DenseNet-121 model such as learning rate, batch size, and momentum, GOA is employed and thereby enhances the recognition performance. The GOA is an evolutionary technique established by simulating t of the swarm of grasshoppers while searching for food. The development of grown-up grasshopper involves 3 stages namely egg, nymph, and adult. The swarming nature of the grasshopper can be mathematically formulated in Eq. (3) for resolving optimization problems.

$$Y_i^d = cx \left\{ \begin{array}{l} \sum_{j=i}^n \quad cx(ul_d - ll_d/2)sf(|Y_j^d - Y_i^d|)^{Y_j - Y_i} / D_{ij} \\ j \neq i + \widehat{T}_d \end{array} \right\} \quad (3)$$

where Y_j, Y_i implies the place of j^{th} and i^{th} grasshoppers. The j^{th} and i^{th} places of grasshoppers from D^{th} dimensional is demonstrated as Y_j^d and Y_i^d correspondingly. The amount of grasshoppers, human interactions, and distance amongst j^{th} and i^{th} grasshopper was illustrated as , sf , and D_{ij} correspondingly. \widehat{T}_d signifies the value of target from the D^{th} dimensional, while ul_d and ll_d refers the upper and lower limits from D^{th} dimensional. The adaptive parameter cz has been utilized for reducing the comfort zones. For balancing exploration and exploitation of entire grasshopper swarm toward the optimum global solutions, a primary cx value outside the bracket from the formula was employed. In addition, the comfort zone, repulsion zone, and attraction amongst the grasshopper are reduced by creating utilize of secondary cx value inside the brackets [17]. The co-efficient cx reduces the comfort zone proportional to the amount of iterations is computed as Eq. (4).

$$cz = cz_{\max} - t(t_{\max}) \quad (4)$$

where cz_{\max} implies the maximal value, cz_{\min} represents the minimal value, t refers the existing iteration, and r_{\max} stands for the maximal amount of iterations. The theoretical method of comfort zone and the attraction as well as repulsion forces amongst the grasshoppers.

3.2 Image Filtering Using Manhattan Distance

At the image filtering stage, the image features in the database which are highly similar to the QI are retrieved. For a given QI, the feature vectors are extracted and compared with each feature vector in the IFD. For every feature vector and Manhattan

distance measure, the nearest N candidate images are included in a histogram of places. Once every query vector is considered, the resultant histogram can be employed for extracting a list of N top-ranked candidate images. The Manhattan distance is the amount of absolute differences among two vectors. In 2D space, the Manhattan distance can be determined as follows:

$$d = |p_1 - q_1| + |p_2 - q_2| \quad (5)$$

For n -dimension space, the Manhattan distance for two data points p_i and q_i can be represented using Eq. (5):

$$d_n = \sum_{i=1}^n |p_i - q_i| \quad (6)$$

Pyramid Spatial Matching Scheme.

The output of the previous process is an ordered list of N images which resembles the QI in a vector distance sense. In this stage, a pyramid spatial matching scheme is designed for capturing the semantic information and spatial relationship among the features [18].

The pyramid match kernels work with orderless image representations. It permits to accurate match of 2 sets of features from higher-dimension appearance space but discards each spatial data. The study advocate an “orthogonal” method: implement pyramid match from the 2D image space and utilize conventional clustering technique in feature space. Especially, quantize each feature vector as to M discrete type, and makes the simplifying assumption that feature of the similar types could be matched with each other. All the channel m provide two sets of 2D vectors, X_m and Y_m , represents the coordinate of feature of type m found in the corresponding image.

$$K^L(X, Y) = \sum_{m=1}^M \kappa^L(X_m, Y_m). \quad (7)$$

This method is a benefit of keeping continuity with common “visual vocabulary” model—indeed, it minimizes to bag of features while $L = 0$. Since the pyramid matching kernel is a weighted sum of histogram intersection, and $\text{cmin}(a, b) = \min(ca, cb)$ to positive number, perform K^L as histogram connection of “long” vector generated by concatenating the accurately weighted histogram of each channel at each resolution. For M channels and L levels, the resultant vector has dimensionality $\sum_{l=0}^L 4^l = M \frac{1}{3} (4^{L+1} - 1)$. For maximal computation efficacy, we normalized each histogram by the overall weight of each feature from the image, in effects force the overall amount of features from each image would be similar.

4 Performance Validation

The performance validation of the ODNPSM-VPR technique can be examined using three benchmark datasets [19] namely Garden Point, Berlin A100, and Synthesized Nordland. The Garden point dataset comprises videos captured by a handheld mobile phone through the Gardens Point Campus at QUT University. It includes variations in viewpoints and conditions. Next, the Berlin A100 dataset includes day time series of frames over the identical urban route, with every traversal saved from various cars and users of Mapillary. The Synthesized Nordland is a huge dataset, comprising a collection of 1415 frames/sequence. The hyperparameter values of the DenseNet121 model re-learning rate: 0.0001, momentum: 0.8, and batch size: 64. Figures 2, 3 and 4 illustrates the sample images of three datasets.

Table 1 provides a comparative AUC analysis of the ODNPSM-VPR technique with recent methods on distinct datasets. Figure 5 examines the performance of the ODNPSM-VPR technique with existing techniques on the Berlin A100 Points dataset interms of AUC. The results portrayed that the SeqSLAM and AlexNet models have depicted poor results with the minimal AUC of 14.30 and 42%. Besides, the Region-BoW and Region-VLAD techniques have resulted in certainly improved AUC of 78.40% and 73.80% respectively. Though the NetVLAD technique has accomplished reasonable AUC of 92.30%, the ODNPSM-VPR technique has surpassed the existing techniques with a higher AUC of 94.60%.

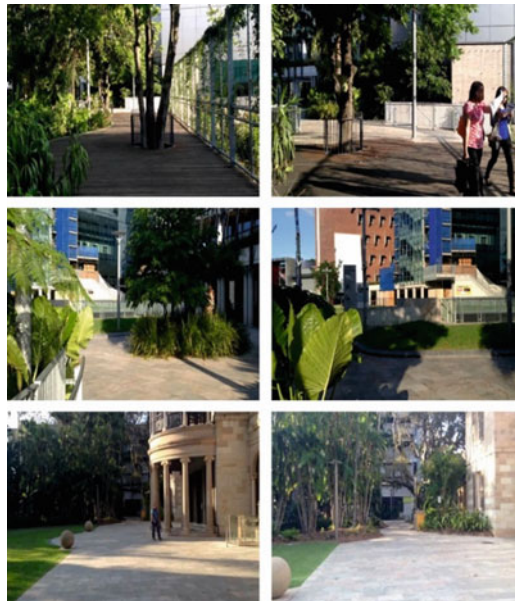


Fig. 2 Sample images Garden point



Fig. 3 Sample images Berlin A100

Figure 6 offers a comparative study of the ODNPSM-VPR technique with existing techniques on the Synthesized Nordland dataset interms of AUC. The figure displayed that the SEQSLAM model has shown worse performance with a lower AUC of 15.90%. Besides, the Region-BoW, Region-VLAD, AlexNet, and NetVLAD techniques have attained somewhat improved AUC of 44, 54.70, 54.8, and 43.70. However, the ODNPSM-VPR technique has outperformed the other methods with a maximum AUC of 63%.

Figure 7 inspects the performance of the ODNPSM-VPR technique with existing techniques on the Garden Points dataset interms of AUC. The figure reported that the AlexNet model has shown worse performance with the lower AUC of 36.60%. At the same time, the Region-BoW, Region-VLAD, and SEQSLAM techniques have obtained slightly increased AUC of 68.90%, 72.60%, and 75.10% respectively. Though the NetVLAD technique has reached to moderately enhanced AUC of 88.20%, the ODNPSM-VPR technique has outperformed the other methods with a maximum AUC of 91.20%.

Figure 8 shows the precision-recall curves obtained by the ODNPSM-VPR technique with recent methods on the test Garden Point dataset. The figure



Fig. 4 Sample images synthesized Nordland

Table 1 Result analysis of odnpsm-vpr technique with existing approaches

| AUC (%) | | | |
|-------------|---------------|-------------|----------------------|
| Methods | Garden Points | Berlin A100 | Synthesized Nordland |
| Region-BoW | 68.90 | 78.40 | 44.00 |
| Region-VLAD | 72.60 | 73.80 | 54.70 |
| SEQSLAM | 75.10 | 14.30 | 15.90 |
| AlexNet | 36.60 | 42.00 | 54.8 |
| NetVLAD | 88.20 | 92.30 | 43.70 |
| ODNPSM-VPR | 91.20 | 94.60 | 63.00 |

depicted that the Region-VLAD and Region-BoW techniques have attained ineffective performance. Followed by, the NetVLAD, AlexNet, and SEQSLAM techniques have obtained somewhat improved and closer performances. However, the ODNPSM-VPR technique has showcased effective outcomes over the other methods.

Figure 9 illustrates the precision-recall curves obtained by the ODNPSM-VPR approach with recent algorithms on the test Berlin A100 dataset. The figure depicted

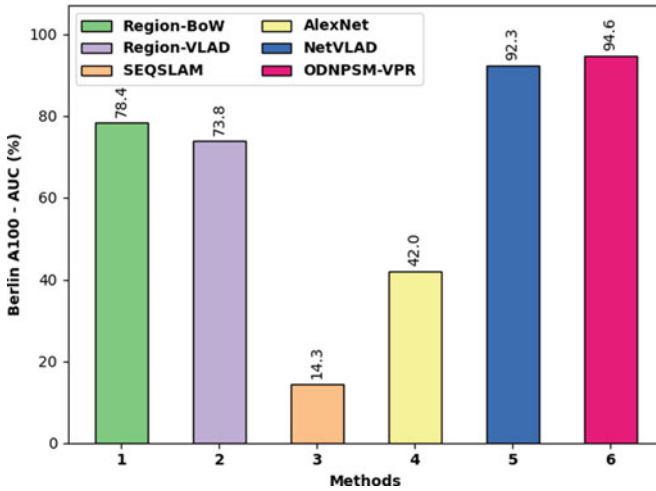


Fig. 5 AUC analysis of ODNPSM-VPR technique under Berlin A100 points dataset

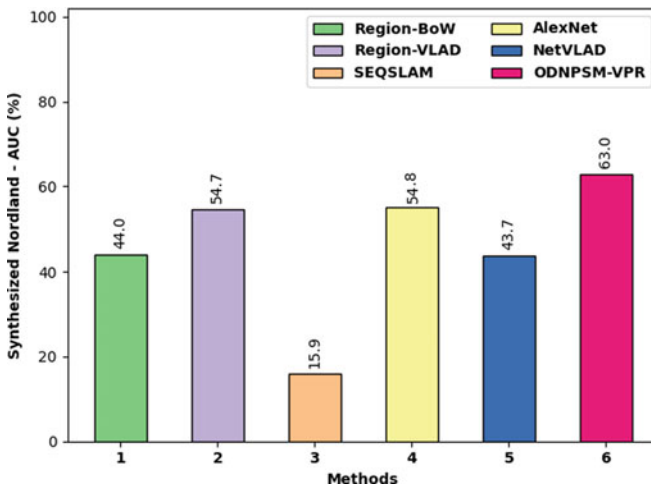


Fig. 6 AUC analysis of ODNPSM-VPR technique under synthesized Nordland dataset

that the Region-VLAD and Region-BoW methodologies have reached ineffective performance. Followed by, the NetVLAD, AlexNet, and SEQSLAM systems have reached somewhat improved and closer performances. At last, the ODNPSM-VPR algorithm has showcased effective outcomes over the other methods.

After examining the above mentioned tables and figures, it is apparent that the ODNPSM-VPR technique has been found to be an effective tool for VPR.

Figure 10 depicts the precision-recall curves obtained by the ODNPSM-VPR algorithm with recent techniques on the test Synthesized Nordland dataset. The figure

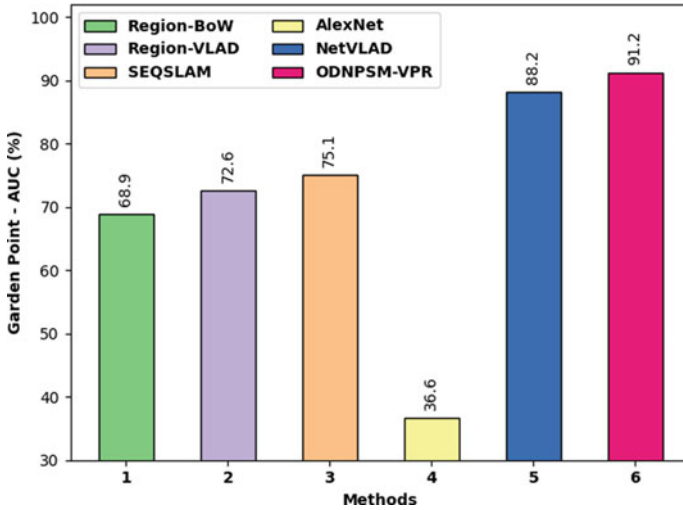


Fig. 7 AUC analysis of ODNPSM-VPR technique under Garden points dataset

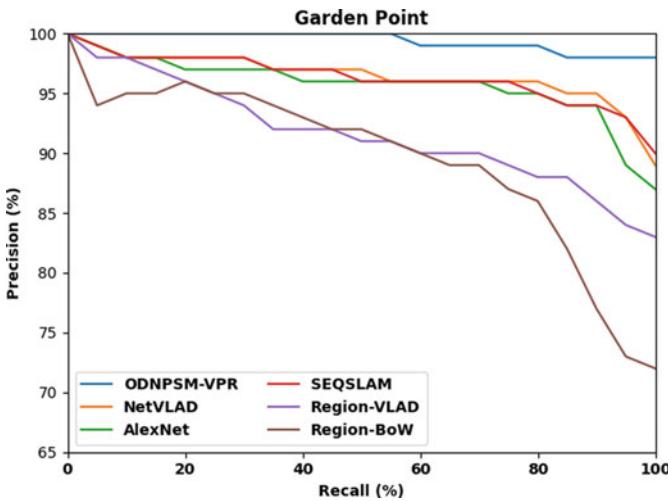


Fig. 8 Precision-recall curves of ODNPSM-VPR technique under Garden point dataset

revealed that the Region-VLAD and Region-BoW systems have attained ineffective performance. Followed by, the NetVLAD, AlexNet, and SEQSLAM approaches have obtained somewhat improved and closer performances. However, the ODNPSM-VPR methodology has exhibited effective outcomes over the other approaches.

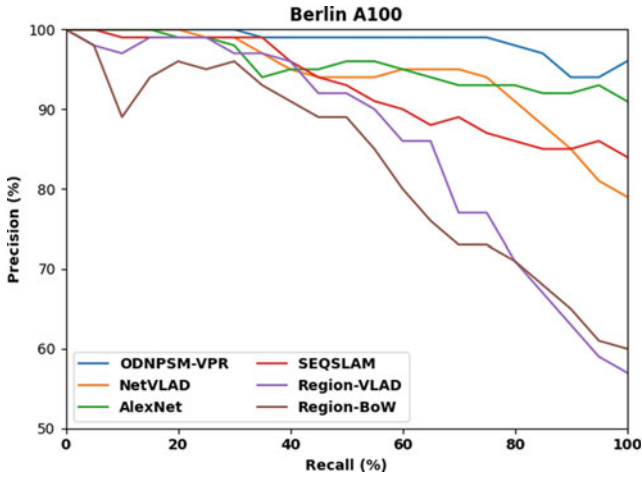


Fig. 9 Precision-recall curves of ODNPSM-VPR technique under Berlin A100 dataset

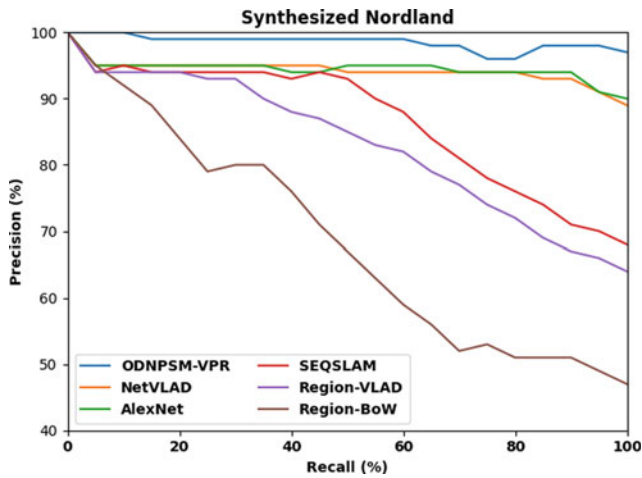


Fig. 10 Precision-recall curves of ODNPSM-VPR technique under synthesized Nordland dataset

5 Conclusion

In this study, a new ODNPSM-VPR technique has been developed for the recognition of visual places. The proposed ODNPSM-VPR technique is robust of simultaneous viewpoint and condition variations. The ODNPSM-VPR technique involves three major elements namely feature extractor, IFD, and SMD. With the inclusion of pyramid spatial matching and hyperparameter tuning of the DenseNet

model, the recognition performance of the ODNPSM-VPR technique is considerably improved. The performance validation of the ODNPSM-VPR technique is performed using benchmark datasets and the outcomes are examined with respect to various measures. The experimental results reported the better performance of the ODNPSM-VPR technique compared to recent approaches. In future, the performance of the ODNPSM-VPR technique can be tested on large scale datasets.

References

1. Zhang X, Wang L, Su Y (2021) Visual place recognition: a survey from deep learning perspective. *Pattern Recogn* 113:107760
2. Manoharan S (2019) Study on Hermitian graph wavelets in feature detection. *J Soft Comput Paradigm (JSCP)* 1(01):24–32
3. Bhalaji N (2021) Cluster formation using fuzzy logic in wireless sensor networks. *IRO J Sustain Wirel Syst* 3(1):31–39
4. Chancán M, Hernandez-Nunez L, Narendra A, Barron AB, Milford M (2020) A hybrid compact neural architecture for visual place recognition. *IEEE Robot Autom Lett* 5(2):993–1000
5. Das R, Kumari K, Thepade S, Manjhi PK (2021) Improved classification of content-based image features using hybrid classification decision. In: *Computer networks and inventive communication technologies*. Springer, Singapore, pp 451–459
6. Akey Sungheetha RSR (2021) Classification of remote sensing image scenes using double feature extraction hybrid deep learning approach. *J Inf Technol* 3(02):133–149
7. Shakya S (2021) Unmanned aerial vehicle with thermal imaging for automating water status in vineyard. *J Electr Eng Autom* 3(2):79–91
8. Camara LG, Přeučil L (2020) Visual place recognition by spatial matching of high-level cnn features. *Robot Auton Syst* 133:103625
9. Zaffar M, Ehsan S, Milford M, McDonald-Maier K (2020) Cohog: a light-weight, compute-efficient, and training-free visual place recognition technique for changing environments. *IEEE Robot Autom Lett* 5(2):1835–1842
10. Camara LG, Přeučil L (2019) Spatio-semantic convnet-based visual place recognition. In: *2019 European conference on mobile robots (ECMR)*. IEEE, pp 1–8
11. Chen L, Jin S, Xia Z (2021) Towards a robust visual place recognition in large-scale vSLAM scenarios based on a deep distance learning. *Sensors* 21(1):310
12. Tomitá MA, Zaffar M, Milford MJ, McDonald-Maier KD, Ehsan S (2021) Convsequential-slam: a sequence-based, training-less visual place recognition technique for changing environments. *IEEE Access* 9:118673–118683
13. Tsintotas KA, Bampis L, Gasteratos A, FIET, (2021) Tracking-DOSeqSLAM: a dynamic sequence-based visual place recognition paradigm. *IET Comput Vision* 15(4):258–273
14. Schubert S, Neubert P, Protzel P (2021) Graph-based non-linear least squares optimization for visual place recognition in changing environments. *IEEE Robot Autom Lett* 6(2):811–818
15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 4700–4708
16. Wang C, Zhao Z, Wang F, Li Q (2021) A novel malware detection and family classification scheme for IoT based on DEAM and DenseNet. *Secur Commun Netw*
17. Mirjalili SZ, Mirjalili S, Saremi S, Faris H, Aljarah I (2018) Grasshopper optimization algorithm for multi-objective optimization problems. *Appl Intell* 48(4):805–820

18. Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06), vol 2. IEEE, , pp 2169–2178
19. <http://imr.ciirc.cvut.cz/Datasets/Ssm-vpr>

Virtual Reality as a Teaching Resource in Higher Education: Professors' Assessment



Álvaro Antón-Sancho, Diego Vergara-Rodríguez, David G. Calatayud, and Pablo Fernández-Arias

Abstract In this paper, a quantitative research is carried out on the perception of a group of Spanish professors about the didactic use of virtual reality in higher education classrooms. Thirty-six Spanish university professors with no previous experience in the use of virtual reality participated in the study. These professors attended a talk-workshop on VR and expressed their assessment of the didactic use of this technology before and after the workshop through a survey of our own creation. The answers were statistically analyzed to compare the mean answers before and after the workshop and to identify gender and age differences in the valuations. The results allow us to conclude that the training received positively increased the participants' perception of all aspects of VR studied and that males and older professors increased their assessment of the technical aspects of virtual reality to a greater extent. In addition, the longest-serving professors positively increase their perception of the ease of use of virtual reality after the workshop.

Keywords Virtual reality · University professor · Survey · Didactic resource · Assessment · Quantitative study

Á. Antón-Sancho · D. Vergara-Rodríguez (✉) · D. G. Calatayud · P. Fernández-Arias
Catholic University of Ávila, C/Canteros s/n 05005, Ávila, Spain
e-mail: diego.vergara@ucavila.es

Á. Antón-Sancho
e-mail: alvaro.anton@ucavila.es

D. G. Calatayud
e-mail: david.glez@ucavila.es

P. Fernández-Arias
e-mail: pablo.fernandezarias@ucavila.es

1 Introduction

Virtual reality (VR) is the set of computerized technologies (hardware and software) that generates a sensory interaction experience in a virtual environment [1]. Immersive virtual reality (IVR) generates a perfectly immersive and fully inclusive experience in the virtual environment represented, while non immersive virtual reality (NIVR) gives rise to a non-complete immersive experience [2]. Immersion refers to the degree to which the user's experience with the simulated environment allows them to connect with it through his/her senses [3]. The degree of immersion of the VR experience depends on how intensely it engages the user's senses and the number of senses brought into play. VR technology is part of a wide range of computational image processing resources that may eventually have didactic applicability [4]. It enjoys a multitude of applications in the most diverse areas of knowledge (Fig. 1), such as Health Sciences [5, 6], Architecture or Engineering [7, 8], Marketing [9, 10], Security [11, 12] or Entertainment [13, 14]. In recent years it has been frequently

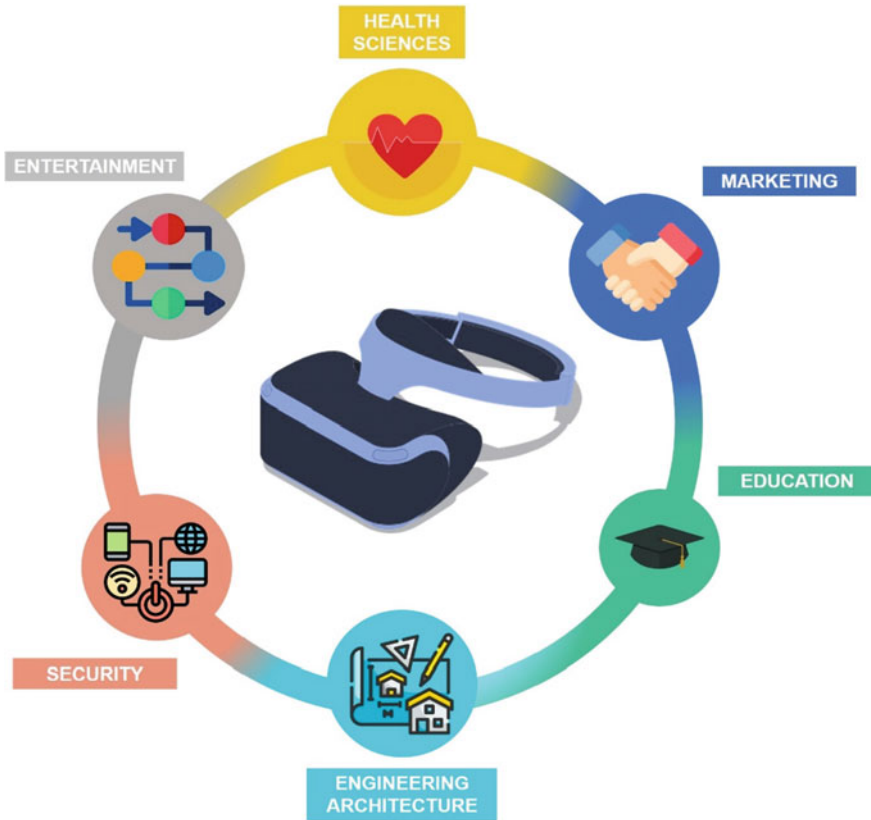


Fig. 1 Virtual reality applications

used in higher education classrooms as a didactic resource [15–18]. As a result, there are numerous publications analyzing the educational effects of VR in higher education and the assessment of professors and students [19–21]. These works usually highlight the positive effects of VR on students' training and identify certain factors that influence the opinion on the use of VR of the affected agents (e.g., the area of knowledge, as shown in [21]). Other works study VR technologies in terms of some of the technical and usability characteristics discussed in this study, but in social contexts other than higher education, such as commerce [22].

This paper presents the results of a quantitative research study based on a VR lecture-workshop experience with Spanish university professors with no previous experience in the use of VR technologies. Perceptions about different aspects of the didactic use of VR in higher education classrooms were collected before and after the workshop in which the teachers participated. The main objective of the paper is to compare the answers given by the participating professors in order to conclude whether the training that took place had a statistically significant influence on their perception of VR. We also analyzed the existence of significant differences by gender or age in the influence of the training on the participants' opinions.

2 Methodology

2.1 Participants

Thirty-six university professors from two different Spanish universities participated in the study (Fig. 2). These participants were selected by a non-probabilistic convenience sampling process. All of them initially stated that they had no experience in the use of VR, so their initial specific knowledge about this technology is that of the general public. The participants attended a lecture-workshop about VR technology and its didactic use in higher education classrooms, so it can be assumed that there was no conceptual disparity among the participants about VR technology. All professors voluntarily and anonymously agreed to answer a questionnaire before the workshop (pre-test) to assess VR and its didactic employability, which they repeated after the workshop (post-test).



Fig. 2 Methodology flowchart

2.2 Objectives

The main objective of this study is to analyze the perception of a set of Spanish university professors with no previous VR experience on the didactic use of VR in higher education classrooms and to study the influence that training acquired during a VR workshop has on this perception. The specific objectives are: (i) to describe the teachers' opinion about the didactic aspects of VR; (ii) to check whether their ratings increase positively after the VR workshop; and (iii) to analyze the influence of gender and age on this perception.

2.3 Variables and Hypotheses

Gender and age of participants act as independent variables. Both are considered dichotomous variables, with the values expressed in Table 1. Age has been divided into two brackets, with a border at 50 years old, in order to analyze whether there are significant differences between young or middle-aged teachers (under 50 years old) and those who are older (over 50 years old) in terms of their perception of VR technologies.

Eight dependent variables are considered in the study. Seven of them are discrete quantitative, with values from 1 to 10, where 1 represents the lowest possible rating and 5 represents the highest (Table 2). These quantitative variables are used to evaluate different aspects of VR, which have been grouped into three families (Fig. 3): (i) technical aspects; (ii) usability; and (iii) didactic effectiveness. The eighth variable

Table 1 Independent variables

| Variable | Possible values |
|----------|-----------------|
| Gender | Female |
| | Male |
| Age | ≤50 years old |
| | >50 years old |

Table 2 Dependent variables

| Family | Variable |
|------------------------|--------------------------------|
| Technical aspects | Interactivity |
| | Immersion |
| Usability | Ease of use |
| | Cost-effectiveness |
| Didactic effectiveness | Motivation |
| | Improved classroom performance |
| | Attention grabbing features |



Fig. 3 Scheme of the variables used and the research objective

is qualitative and asks to identify the dominant characteristics of VR as a teaching resource, as judged by the participant. Within the technical aspects, interactivity refers to the set of relationships established between the model simulated by the VR technology and the user. In particular, it indicates the degree of possibility for the user to influence the shape of the simulated environment or to establish communication channels with it [3]. Within the usability variables, cost effectiveness refers to the assessment of the economic cost of implementing VR technologies in university centers [23]. Here it is a matter of assessing whether the economic cost of implementing VR technologies by the actors involved is affordable, in the sense that the investment can be used in frequent learning experiences and that the learning outcomes justify the investment. Finally, within the didactic effectiveness variables, the attention grabbing features refer to the ability of VR technologies to capture the attention of students, through the sensory experience that it arouses [24].

Throughout the work, we tried to corroborate the following hypotheses: (H1) a specific training session increases the perception of university professors on the didactic use of VR in higher education classrooms; and (H2) the gender and age of university professors condition the way in which the VR workshop influences their evaluations.

2.4 Instrument

A self-designed survey consisting of twelve questions was used. The first four questions were used to determine the sociological profile of the professors and their knowledge of VR (i) gender, (ii) age, (iii) whether they had any previous experience

in the use of VR, and (iv) whether they knew how to distinguish between IVR and NIVR). The following seven asked to rate each of the dependent variables set out in Table 2 on a Likert scale from 1 to 10 where 1 means the lowest rating and 10 means the highest rating. The last question asks to assess the eighth dependent variable by choosing one or more characteristics of VR as those considered most defining of this technology as a teaching resource. The options proposed were as follows: (i) innovative; (ii) flexible; (iii) dynamic; (iv) economical; (v) enriching; (vi) fun; (vii) easy of use; (viii) attractive; (ix) educational; and (x) interactive.

2.5 Methods

This work consists of a quantitative descriptive study based on the results of a survey that was given to the participants as a pre-test and a post-test and that was completed by them before and after a VR workshop, respectively. The mean value was obtained for each of the variables measured in the survey and these statistics were compared before and after the workshop using Student's t-test for comparison of means. Next, the multifactor ANOVA (MANOVA) test was used to study whether the gender and age of the participants influenced their valuations of the different aspects of VR. All tests were performed at the 0.05 significance level. In summary, the study has been developed according to the following phases: (i) design of the training session for teachers on RV and of the evaluation instrument (questionnaire); (ii) performance of the pre-test and development of the training action; (iii) performance of the post-test and (iv) statistical analysis of the results obtained.

3 Results

As for the sample of participants, 50% are males and 50% are females, so the gender distribution is perfectly homogeneous. By age, 66.67% are 50 years old or younger and 33.33% are over 50 years old, so the age distribution is not homogeneous.

The Cronbach's alpha parameter of the Likert scale of the survey is 0.8306, so the quantitative scale of the instrument has an optimal internal consistency. The overall valuations given by the participants to VR are very high in all the families of quantitative variables, with the exception of the cost effectiveness variable, where the valuation is very low in the pre-test and intermediate in the post-test. Thus, the participating professors have an excellent concept of VR, but identify the economic impact as a real disadvantage. On the other hand, the t-test statistics (Table 3) identify that there is a positive and statistically significant increase in the mean value of all the quantitative aspects of VR that have been measured. This increase is especially notable in the variables of ease of use and cost-effectiveness, so that the workshop has been particularly influential with respect to these variables. The t-test was performed in a situation of homoscedasticity, corroborated by Levene's test.

Table 3 Means and t-test statistics of the pre-test and the post-test

| Variable | Mean pre-test | Mean post-test | T | p-value |
|--------------------------------|---------------|----------------|--------|---------|
| Interactivity | 9.67 | 10.00 | 2.6458 | 0.0121* |
| Immersion | 9.83 | 10.00 | 2.6458 | 0.0121* |
| Ease of use | 7.33 | 10.00 | 9.8687 | 0.0000* |
| Cost-effectiveness | 2.67 | 6.67 | 6.2996 | 0.0000* |
| Motivation | 9.67 | 10.00 | 4.1833 | 0.0000* |
| Improved classroom performance | 9.33 | 10.00 | 4.1833 | 0.0000* |
| Attention grabbing features | 9.67 | 10.00 | 4.1833 | 0.0000* |

* $p < 0.05$

The MANOVA test identifies the existence of gender gaps in the interactivity and immersion variables, i.e., in the family of technical aspects of VR (Table 4). From the mean values it can be deduced that the workshop has not had any influence on the perception of female professors with respect to the technical aspects (all of them gave the maximum valuation to these aspects), but it has significantly influenced the valuation of males, in the sense of generating a positive increase (in the post-test, all males gave the maximum valuation to the technical aspects of RV).

The age of the participants was significantly discriminative for the variables of technical aspects (interactivity and immersion) and also for the ease of use variables, within the family of usability variables (Table 5). In the variables of technical aspects, the workshop only had a positive influence on the perception of those over 50 years old. As for the usability variable, the workshop had a positive influence on the perception of the participants in the two age groups considered, but in a notably more intense way in those over 50 years old (there is a difference of four points

Table 4 Means and MANOVA statistics of the pre-test and the post-test when differentiating by gender

| Variable | Females | | Males | | MANOVA | p-value |
|--------------------------------|----------|-----------|----------|-----------|--------|---------|
| | Pre-test | Post-test | Pre-test | Post-test | | |
| Interactivity | 10.00 | 10.00 | 9.33 | 10.00 | 8.5000 | 0.0048* |
| Immersion | 10.00 | 10.00 | 9.67 | 10.00 | 8.5000 | 0.0048* |
| Ease of use | 7.67 | 10.00 | 7.00 | 10.00 | 1.5455 | 0.2181 |
| Cost-effectiveness | 3.33 | 7.33 | 2.00 | 6.00 | 0.0000 | 1.0000 |
| Motivation | 9.67 | 10.00 | 9.67 | 10.00 | 0.0000 | 1.0000 |
| Improved classroom performance | 9.33 | 10.00 | 9.33 | 10.00 | 0.0000 | 1.0000 |
| Attention grabbing features | 9.67 | 10.00 | 9.67 | 10.00 | 0.0000 | 1.0000 |

* $p < 0.05$

Table 5 Means and MANOVA statistics of both pre- and post-test when differentiating by age

| Variable | ≤50 years old | | >50 years old | | MANOVA | p-value |
|--------------------------------|---------------|-----------|---------------|-----------|---------|---------|
| | Pre-test | Post-test | Pre-test | Post-test | | |
| Interactivity | 10.00 | 10.00 | 9.00 | 10.00 | 22.6670 | 0.0000* |
| Immersion | 10.00 | 10.00 | 9.50 | 10.00 | 22.6670 | 0.0000* |
| Ease of use | 8.00 | 10.00 | 6.00 | 10.00 | 18.1330 | 0.0000* |
| Const-effectiveness | 2.50 | 6.75 | 3.00 | 6.50 | 0.3027 | 0.5840 |
| Motivation | 9.75 | 10.00 | 9.50 | 10.00 | 2.2667 | 0.1368 |
| Improved classroom performance | 9.50 | 10.00 | 9.00 | 10.00 | 2.2667 | 0.1368 |
| Attention grabbing features | 9.75 | 10.00 | 9.50 | 10.00 | 2.2667 | 0.1368 |

* p < 0.05

between the pre-test and post-test means in the perceptions of these participants, while the difference is only two points in younger participants).

As for the qualitative dependent variable, on the dominant characteristics of VR, there were notable changes in the post-test compared to the pre-test (Fig. 4). The innovative, dynamic, fun and attractive options are chosen more frequently after the workshop than before, while the educational and interactive options do not change their frequency in the post-test compared to the pre-test. Particularly interesting is the case of the options flexible, enriching and easy to use, which were not chosen by any professor before the workshop, but in the post-test reach frequencies of 50, 100 and 83%, respectively.

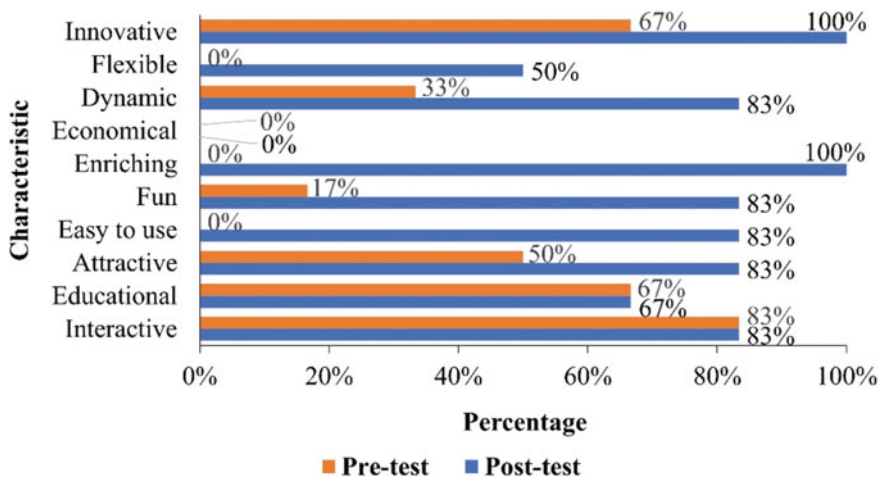


Fig. 4 Frequencies of VR characteristics chosen by participants

On the other hand, the characteristic on cost-effectiveness is not considered as dominant in the VR by any participant, neither in the pre-test nor in the post-test. The economic aspect, therefore, although its evaluation improves positively with respect to the pre-test, is not considered a defining characteristic of VR in any case in the post-test.

4 Discussion and Conclusions

The research objectives have been met throughout the work. Indeed, the opinions expressed by the professors have been analyzed and the answers of the pre-test and post-test have been compared. Likewise, the existence of both gender or age gaps in the mean values provided by the participants was studied. As a result of this analysis, it has been shown that the mean values of the participants suffer a positive increase in the post-test with respect to the pre-test in all the variables studied (interactivity, immersion, ease of use, cost-effectiveness, motivation, improved classroom performance and attention grabbing features), which implies that the workshop actually has a positive influence on the mean perception that professors have of VR as a teaching resource (Table 3). In [25] it is noted that some of the reasons for these high ratings may be that VR technologies allow for a more practical presentation of didactic content than traditional presentations and thus facilitate more meaningful learning. Despite this, it should be noted that the design of VR applications directly affects the level of meaningful learning [26]. This positive increase is especially notable in the variable that measures the economic cost-effectiveness of VR.

It is clear that university professors with no experience in the use of VR have the idea that the cost of VR technologies is a real impediment to their implementation in classrooms. However, that idea changes moderately after the training received. Nevertheless, even after the workshop, professors still see the economic cost as a disadvantage of VR (Fig. 1). All this allows to conclude that hypothesis H1 is verified. Likewise, this conclusion supports the hypothesis previously established in other papers [15], where it is suggested that training professors in VR would increase their appreciation of VR. Furthermore, these results presented here are in line with work exploring the perception of VR use in various higher education settings, such as military training [27], where it is shown that effectiveness in terms of academic outcomes softens the perception about the economic impact that users perceive VR technologies to have. In view of the above results, it can be concluded that the implementation of specific faculty training plans on VR by universities would help to increase its use in the classroom by professors. However, analogous studies with longer workshop experiences and larger samples of participants would be necessary to universalize these results.

In addition, gender and age gaps have been identified in some families of variables studied (Tables 4 and 5), which means that the workshop has not influenced males and females or younger and older participants in the same way in terms of their perception of VR with respect to these variables. Specifically, the aspects in which

these differences are statistically significant are the technical aspects (and ease of use in the case of age), with males and older professors positively increasing their perceptions more significantly. This allows us to conclude that hypothesis H2 has been partially verified, because the gender and age gaps occur only in some of the variables studied. These results are novel in the literature, because many previous works, like [19, 20], do not analyze gender or age differences in the perception of the agents involved and others, such as [21], had not identified gender gaps, although partially by age, and, in any case, did not compare opinions before and after a training experience on VR. Universalization of these results would require, again, expanding the sample and the training experiences. However, these results suggest the need for the training actions on VR that universities implement to have a corrective perspective of the observed gender and age differences.

Acknowledgements The authors wish to acknowledge the financial support provided by the Spanish Association “Amigos de la Universidad Católica de Avila” (Award 2021).

References

1. Steuer J (1992) Defining virtual reality: dimensions determining telepresence. *J Commun* 42:73–93
2. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I (2020) A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput Educ* 147:103778
3. Roussou M, Oliver M, Slater M (2006) The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning. *Virtual Reality* 10:227–240
4. Sathesh A, Adam EEB (2021) Hybrid parallel image processing algorithm for binary images with image thinning technique. *J Artif Intell* 3(3):243–258
5. Longo UG, De Salvatore S, Candela V, Zollo G, Calabrese G, Fioravanti S, Giannone L, Marchetti A, De Marinis MG, Denaro V (2021) Augmented reality, virtual reality and artificial intelligence in orthopedic surgery: a systematic review. *Appl Sci* 11:3253
6. Smys S (2019) Virtual reality gaming technology for mental stimulation and therapy. *J Inf Technol* 1(01):19–26
7. Noghabaei M, Heydarian A, Balali V, Han K (2020) Trend analysis on adoption of virtual and augmented reality in the architecture, engineering, and construction industry. *Data* 5:26
8. Silva J, Gaitán M, Varela N, Pineda Lezama OB (2019) Engineering teaching: simulation, industry 4.0 and big data. In: International conference on computational vision and bio inspired computing. Springer, Cham, pp 226–232
9. Zhao H, Zhao QH, Ślusarczyk B (2019) Sustainability and digitalization of corporate management based on augmented/virtual reality tools usage: china and other world IT companies’ experience. *Sustainability* 11:4717
10. Ho JCF, Zhang X (2020) Strategies for marketing really new products to the mass market: a text mining-based case study of virtual reality games. *J Open Innov Technol Mark Compl* 6:1
11. Lv Z, Chen D, Lou R, Song H (2021) Industrial security solution for virtual reality. *IEEE Internet Things J* 8(8):6273–6281
12. Ott R, Gutierrez M, Thalmann D, Vexo F (2006) Advanced virtual reality technologies for surveillance and security applications. In: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications (VRCA ‘06). pp 163–170

13. Allen RC, Singer MJ, McDonald DP, Cotton JE (2000). Age differences in a virtual reality entertainment environment: a field study. *Proc Hum Fact Ergon Soc Annual Meet* 44:542–545
14. Kodama R, Koge M, Taguchi S, Kajimoto H (2017) COMS-VR: mobile virtual reality entertainment system using electric car and head-mounted display. In: 2017 IEEE Symposium on 3D User Interfaces (3DUI). pp 130–133
15. Vergara D, Rubio MP, Lorenzo M (2014) Interactive virtual platform for simulating a concrete compression test. *Key Eng Mater* 572:582–585
16. Vergara D, Rubio MP, Lorenzo M (2018) A virtual resource for enhancing the spatial comprehension of crystal lattices. *Educ Sci* 8:153
17. Soliman M, Pesyridis A, Dalaymani-Zad D, Gronfula M, Kourmpetis M (2021) The application of virtual reality in engineering education. *Appl Sci* 11:2879
18. Vergara D, Fernández-Arias P, Extremera J, Dávila LP, Rubio MP (2022) Educational trends post COVID-19 in engineering: virtual laboratories. *Mater Today Proc* 49:155–160
19. Fairén M, Farrés M, Moyés J, Insa E (2017) Virtual reality to teach anatomy. *Eurographics* 51–58
20. Fabris CP, Rathner JA, Fong AY, Sevigny CP (2019) Virtual reality in higher education. *Int J Innov Sci and Math Educ* 27:69–80
21. Vergara D, Antón-Sancho Á, Extremera J, Fernández-Arias P (2021) Assessment of virtual reality as a didactic resource in higher education. *Sustainability* 13:12730
22. Kumar TS (2021) Study of retail applications with virtual and augmented reality technologies. *J Innov Image Proc (JIIP)* 3:144–156
23. Vergara D, Extremera J, Rubio MP, Dávila LP (2020) The proliferation of virtual laboratories in educational fields. *ADCAIJ Adv Distrib Comput Artif Intell J* 9:85–97
24. Vergara D, Rubio MP (2012) Active methodologies through interdisciplinary teaching links: industrial radiography and technical drawing. *J Mater Educ* 34:175–186
25. Pagano K, Haddad A, Crosby T (2017) Virtual reality-making good on the promise of immersive learning: the effectiveness of in-person training, with the logistical and cost-effective benefits of computer-based systems. *IEEE Consum Electr Magaz* 6:45–47
26. Vergara D, Extremera J, Rubio MP, Dávila LP (2019) Meaningful learning through virtual reality learning environments: a case study in materials engineering. *Appl Sci* 9:4625
27. Bhagat KK, Liou WK, Chang CY (2016) A cost-effective interactive 3D virtual reality system applied to military live firing training. *Virt Real* 20:127–140

Automated Intelligent Hematology Classification System Using Image Processing and Neural Networks



B. G. Taralekar, Prithviraj Chauhan, Shrinath Palwankar, Celsy Phillips, and Sarang Patil

Abstract In this paper, a method has been proposed which uses an Image Processing and Deep learning-based approach to classify microscopic blood smear images based on 7 classes of blood diseases namely, Leukemia, Anemia, Lymphoma (CLL, FL, MCL), Myeloma and Malaria, from the healthy blood images. Image preprocessing techniques based on Feature Extraction and Ni-black Thresholding were used on image dataset to obtain features for identification and classification of Leukemia and Anemia. Thereafter, a neural network based on VGG16 was implemented to train the model for classification of all the diseases which included pretrained weights from ImageNet. For validation of the model, the scores of Precision, Recall, and F-score were taken into account to calculate the accuracy of the model. Through this methodology, the model was able to achieve an accuracy of 98.6% with minimum loss of 0.47. The proposed system will help hematologists to identify blood diseases more accurately and faster with this automatic analysis system.

Keywords Deep learning · Feature extraction · Hematology · Image processing · Neural network

B. G. Taralekar · P. Chauhan (✉) · S. Palwankar · C. Phillips · S. Patil
Department of Electronics and Telecommunication, Vishwakarma Institute of Technology, Pune,
Maharashtra, India
e-mail: prithviraj.chauhan18@vit.edu

B. G. Taralekar
e-mail: bharat.taralekar@vit.edu

S. Palwankar
e-mail: shrinath.palwankar19@vit.edu

C. Phillips
e-mail: celsy.phillips18@vit.edu

S. Patil
e-mail: sarang.patil18@vit.edu

1 Introduction

The analysis of microscopic images plays a huge role in the fields of medicine and computer science. Several research issues are involved in the analysis of microscopic images which are the first steps in detecting and diagnosing diseases like malaria, leukemia, and anemia. There have been many researchers conducted on identification of diseases which involves algorithms based on image processing, computer vision, machine learning and deep learning. Following are the diseases that will be identified and classified in the proposed work.

“Leukemia”, which refers to the cancer of the blood and bone marrow (where blood cells are produced) is one such disease where there is a rapid growth of abnormal WBCs. The diagnosis of leukemia is based on the fact that the white blood cell count is increased with immature last cells (lymphoid or myeloid) and the neutrophils and platelets decrease. In “Anemia”, this disease is diagnosed based on the shape of the RBCs. People with anemia have red blood cells that have an unusual shape and might look larger or smaller than normal. In the case of “Lymphoma”, which is a cancer affecting lymph nodes, comes in 3 types i.e., CLL (chronic lymphocytic leukemia), FL (follicular lymphoma) and MCL (mantle cell lymphoma). In this disease, the lymphocytes, which are a type of WBCs that are present in the lymph, become abnormally cancerous and grow uncontrollably. As these cells travel through the lymphatic system, they settle in the lymph nodes and cause them to swell. And in “Myeloma”, which is a plasma cell related disease, there is an abnormal growth of the plasma cells in the bone marrow and based on that this disease is diagnosed. Lastly, there is “Malaria” which is a disease caused by a bite from a malaria-carrying mosquito. The malaria parasite enters the bloodstream, multiplies, infects, and destroys the RBCs and based on the number of RBCs present, this disease is detected.

In order to detect these diseases, hematologists have to thoroughly examine blood smears under microscope for proper classification of diseased cells and identify the disease which can be quite time consuming and inaccurate. This approach will help the hematologists by reducing the time for diagnosing the diseases with existing technologies using chemical and electro voltaic methods, this will also reduce the expense of machines and human resource as this is an integrated system for 7 types of diseases. Therefore, this research aims to develop an image processing and deep learning-based system to detect these diseases in an efficient and more accurate way. The model will classify about 7 diseases and the classification model is validated based on the accuracy of the model, precision, recall and f1 score of the confusion matrix obtained from the model.

2 Literature Work

In the past, there have been many works related to classification and detection of hematologic diseases.

Tiancheng Xia et al. detected CBCs (complete blood cells) by using deep neural networks through YOLOv3 algorithm and applied the detector to the blood samples of COVID 19 patients where blood cell clots are its typical symptom. Their precision and recall values of test results were low which led to a lesser model performance accuracy [1]. Asem H. and Ashraf Y. used data mining techniques through decision tree, rule induction and Naive Bayes on a test blood dataset which detected blood diseases of adult and children hematology and tumor. The accuracy of data mining techniques was quite less. Maximum accuracy to detect tumor was achieved by Naïve Bayes which was only 56% and for adult and children hematology, rule induction technique achieved accuracy in the range of 57–67%. Hence, the use of data mining algorithms was ruled out for the proposed system [2]. Angelo Galiano et al. developed a home health assistance web-based communication system to manage medical services measuring data from patients at home which would be automatically transferred to the hospital [3]. Also, in [4], another health care monitoring system was developed which was an IoT based biotelemetry system to analyze ECG signals of patients. Both of the above works have built a hardware-based system to diagnose medical symptoms which are relatively expensive and requires high maintenance. Same is the case with [5] in which Rao et al. have developed a hardware-based image processing system which works on image segmentation and localization of the iris of the eye. Lei Zhao et al. did a study based on the blood quality control materials for analyzing hematologic diseases. The equipment stated in this work is a hematology analyzer which is used to monitor the blood quality and analyze its components. This system makes use of many chemicals to store blood and even temperature conditions are to be considered carefully for storage [6]. The system proposed in this paper does not require to store huge amounts of blood for detection, only a microscopic image of a drop of blood smear is required to perform detection and classification which makes it more efficient. Guclu Ongun et al. developed an automatic differential blood count system using active contour models and for classification used K-nearest neighbor, learning vector quantization, multilayer perceptron and support vector machine [7]. Wei Yu et al. developed an automatic cell recognition system using deep neural networks to classify different types of leukocytes. The overall accuracy achieved by this system was about 88.5% which is relatively lesser than the accuracy of the proposed system [8]. Pooja and Nagaraj used ensemble learning methods like stacking, bagging, voting, Adaboost and Bayesian boosting and applied classifiers such as Decision tree, artificial neural network, Naive Bayes and K-nearest neighbor to detect anemia from red blood cells [9]. Preeti and Virani devised a methodology for leukemia detection using image processing techniques and image segmentation based on K means clustering, Marker controlled watershed algorithm and HSV color-based segmentation then used SVM for further classification of leukemia types [10]. Both of the above works have implemented ensemble algorithms and classifiers to detect

the diseases but were able to obtain only 80–90% accuracy. Deep learning algorithms are known to produce greater accuracy and classification performance as compared to other machine learning algorithms. Subhash et al. proposed an acute lymphoblastic leukemia detection system that used OpenCV and skimage for image processing to extract features from blood images and used classifiers such as CNN, FNN, SVM and KNN [11]. Prashanth built an automatic microscopic blood smear RBC classification using Principal Component Analysis (PCA) and SVM classifier [12–18] were based on detection and classification of myeloma cells in microscopic images which was done using various deep learning algorithms like CNN, ANN, Mask R-CNN, along with classification algorithms based on SVM, Random Forest, etc. Hend Mohamed et al. developed an approach to automatically detect WBC cancer diseases by dividing the disease categories based on similar symptoms, extracting features from them and applying Random Forest classifier to classify the diseases [19–22] discusses approaches to detect lymphoma cells using a computer vision method, an immunophenotyping approach and a segmentation-based approach called DenseX-Net respectively. The problem with the DenseX-Net approach is that it faced problems with detecting lymphoma images where it behaved badly on boundary delineation wherever the boundary of lymphoma cell was blurred. Also, it misdiagnosed certain images since its maximum accuracy was 72.84%. Hence, for the proposed system, image processing methods were implemented to detect the boundaries of infected cells thoroughly and extract required features and the infected cells were labeled properly in the images for better classification. [23–29] discusses various approaches for classification of types of blood cells. Deep learning approaches such as Inception Recurrent Residual Convolutional Neural Network (IRRCNN), Faster R-CNN were used, some other classification algorithms based on Otsu and Naïve Bayes were used and some image processing techniques such as masking, and morphological operations were implemented. [30–32] explains approaches to count different types of cells from microscopic images based on computer vision and machine learning algorithms.

Most of the previous works are based on deep neural networks or a combination of more than one image processing and machine learning based algorithms to detect only one particular type of blood disease. The performance accuracy of most of the literature works was found to be less than the accuracy of the proposed system. Deep learning algorithms are found to achieve better accuracies as compared to other approaches for classification. Some of the works mentioned above have used deep learning methods such as YOLO, Faster and Mask R-CNN and IRRCNN. Even [33, 34] have developed image detection and classification systems based on CNN and have proved that neural networks give the best accuracy. More training images should be added to enhance the classification performance. But the above approaches have worked only on classification of one or few diseases and the sample size of training images taken were less.

Hence, the proposed system has been built using a deep learning approach to classify and detect 7 types of blood diseases with a large dataset which contains ample number of images for each disease class. The model has achieved a great performance accuracy of 98.6%. Moreover, the implementation of preprocessing techniques on image datasets has contributed to the enhancement of the classification performance of the model which led to achieve more accuracy than the approaches mentioned in the literature work. Further sections will explain the dataset and methodology used for the system along with the results and performance parameters that were used to validate the accuracy of the model.

3 Methodology

In the proposed methodology, an approach based on Image processing and Deep learning was implemented to detect and classify 7 classes of diseases i.e., Leukemia, Anemia, CLL, FL, MCL, Myeloma and Malaria from the healthy blood smear images. The approach is divided into 2 sections namely, *Image Preprocessing* and *Classification*. The first section will explain the application of image processing operations on datasets of Leukemia and Anemia based on Feature Extraction and Ni-black Thresholding. For the remaining 5 diseases, ready-made preprocessed datasets were available on the net so the preprocessing for the same was not required. In the second section, a deep learning model was built to classify the diseases which is based on a neural network. VGG16 was used as the neural network for the model along with pretrained weights of ImageNet. These sections will be explained in depth in the following Proposed Work segment of the methodology. Following is the flowchart of the proposed methodology (Fig. 1).

A. Dataset

Images dataset of anemia and Leukemia was taken from google images. Anemia dataset consists of 20 images and Leukemia dataset consists of 40 images. Images taken from google was not pre-processed, for Anemia detection the Ni-Black preprocessing is used to get better results. For Leukemia the feature extraction method is used, after applying preprocessing on both the dataset it is finally used in model. Pixel size of both the dataset is resolved to $(224 \times 224,3)$ because the model accepts the $(224 \times 224,3)$ resolution [35, 36].

Image datasets for each of the disease classes were taken from Kaggle which were already preprocessed. Three datasets of Lymphoma, Myeloma, and Malaria were used.

Lymphoma: The Lymphoma dataset consists of three classes of diseases (CLL, FL, MCL), totaling 5400 images [37]

Myeloma: For Myeloma, a total of 298 images were taken. The Bone marrow photos were collected in a microscope of a pixel size of 2040×1536 pixels and then preprocessed [38–40].

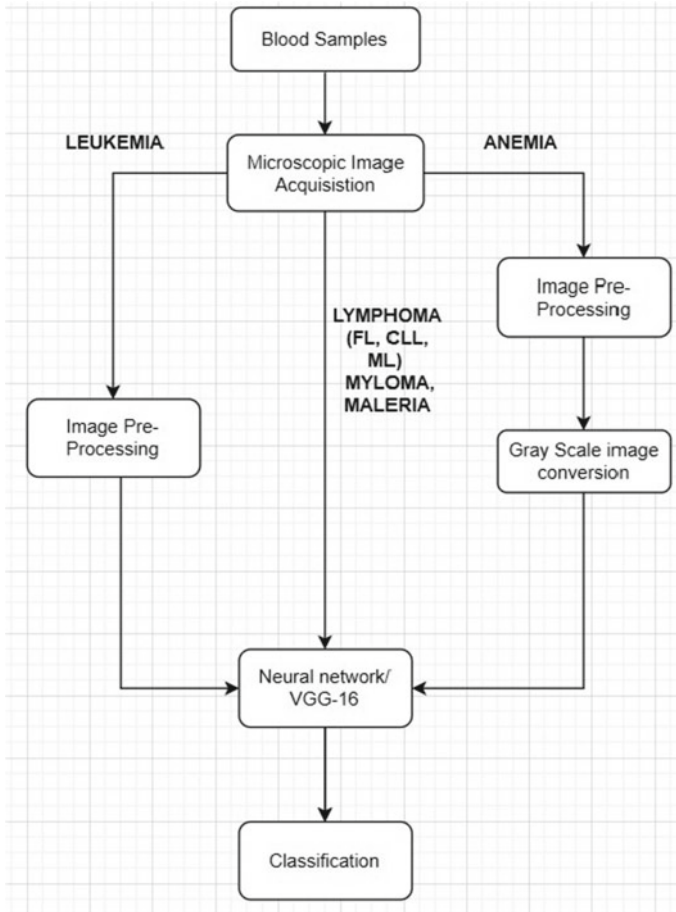


Fig. 1 Flowchart of proposed methodology

Malaria: The Malaria dataset consists of two classes those are infected and non-infected, combining the total number of images are 27,558 [41].

B. Proposed Work

1. Image Preprocessing

In this section, image processing algorithms were used on two different datasets of Leukemia and Anemia. The two main algorithms used are:

- 1.1 *Feature Extraction*: This algorithm was used to detect Leukemia from blood sample images of Leukemic patients. The images were first converted to grayscale from its BGR form. Then a feature extraction method was applied on the grayscale images to extract features which is helpful for distinguishing the resulting values from the standard ones. This feature extraction was done by segmenting blood

cells and cell nuclei into binary equivalent images. The features were based on its color, geometrical features like symmetry and concavity, texture features like entropy and homogeneity and statistical features like skewness and mean gradient matrix.

- 1.2 *Ni-black Thresholding*: This algorithm was used to detect Anemia from microscopic blood images of Anemic patients. The niblack thresholding method aims to achieve better results specially for microscopic images, it separates the object of our interest from the rest of the image for further processing. [42] compared different binarization methods like Niblack, Bernsen, Sauvola and Otsu for counting of WBCs and RBCs. In their experiments over different samples with different conditions showed that Niblack is the most reliable method. It maintains disjoint components which is necessary for avoiding over or under segmentation. The images were first converted to grayscale from its BGR form since it is a crucial step for applying Niblack Thresholding. The gray level of pixels belonging to RBCs are entirely different from the gray levels of pixels belonging to the other blood contents. The niblack thresholding method is applied on the images by segmenting the foreground region to isolate the normal and abnormal RBCs from the background region. It provides us with the binary image from the grayscale one, effectively separating the normal and abnormal cells from the rest of the contents of blood. It is done by calculating the local mean and standard deviation of the pixel value within a confined size of a window in the grayscale image. The thresholding value for each window is computed based on the mean m , standard deviation σ of the pixels in that window.

$$T = m + k \times \alpha$$

k is -0.2 suggested by Ni-black. In this way, the images are binarized by showing sickle cells darker than the rest of the contents.

2. Classification

In this section, a deep neural network-based approach was used to perform the classification of the diseases. A neural network model called VGG16 was used. VGG16 is Very Deep Convolutional Networks for Large-Scale Image Recognition with 93% using ImageNet weight, in 2014 it was considered as the best architecture for image classification. It is a convolutional neural network (CNN) that is 16 layers deep. A pre-trained version of the network which was trained on more than a million images is loaded and this pretrained weight was taken from the ImageNet database. This weight can classify images up to 1000 object categories. Hence, this network was chosen for its outstanding classification property. Following is the architecture

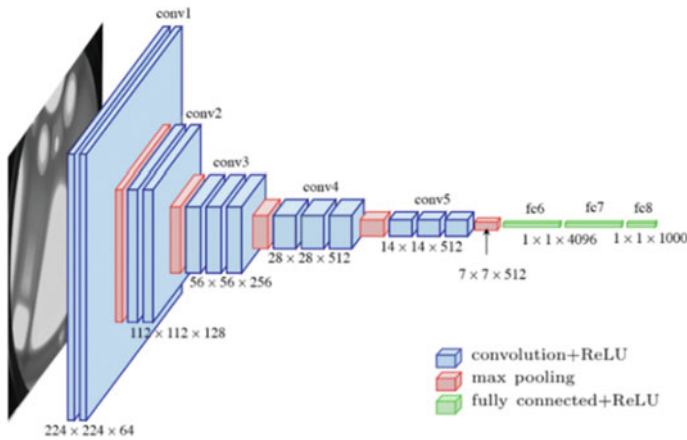


Fig. 2 VGG16 architecture (Source: <https://medium.com/mlearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484>)

of VGG16 which shows a schematic arrangement of all the layers i.e., convolutional, pooling, and dense (fully connected) layers (Fig. 2).

The model consists of 6 layers of which there are 3 dense layers, 2 dropout layers and a flatten layer. Firstly, a flatten layer was added, then a dense layer was added with an activation function of ‘LeakyReLU’ and its learning rate as 0.3. This is followed by a dropout layer of rate 0.5 and again a dense layer. Then another dropout layer of rate 0.3 followed by a dense layer added on 8 classes (7 disease classes and 1 healthy image class) with an activation function of ‘SoftMax’. Then, a loss function of ‘categorical_crossentropy’ was used with an optimizer function of ‘adam’. For training of the model, 100 epochs were set for a batch size of 10. Tensor board was used to extract the graph results of loss and accuracy over epoch of the model. This research tried other loss such as ‘root mean square’ but it gave 31.3% so this was dropped and ‘adam’ was chosen.

4 Results

In this paper, a novel problem has been solved which involves including more than two diseases into a CNN based classifier. A high accuracy has been achieved which is specifically 98.6% and were able to predict diseases correctly. For acquiring a dataset of two diseases, image pre-processing techniques was implemented and for the rest of diseases which directly fed the data into the model. Following results are divided into 2 sections, (4.1) Image Preprocessing, and (4.2) Classification.

4.1 Image Preprocessing

This section provides results for the preprocessing of disease classes of Leukemia and Anemia.

Preprocessing of Leukemia: (Fig. 3)

The image above shows the original microscopic image of blood smear containing leukemia affected cells. The image below shows the output after preprocessing the image and extracting features. The leukemia affected cells are highlighted as shown in the figure below.

Preprocessing of Anemia: (Fig. 4)

The image above shows the original microscopic image of blood smear containing anemia affected cells. The image below shows the output after preprocessing the image using Ni-black Thresholding. The anemic cells are threshold from the rest of the blood contents as shown in the figure below.

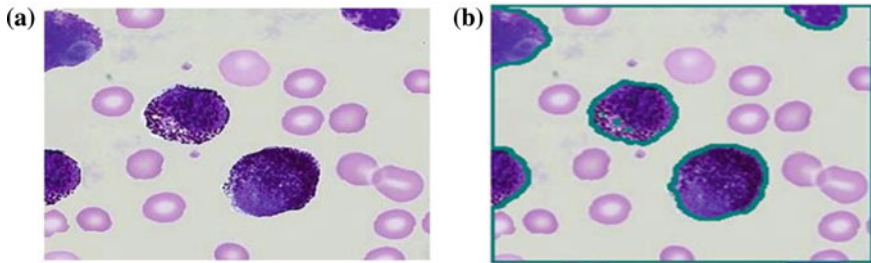


Fig. 3 a Original image b LEUKEMIA detected image

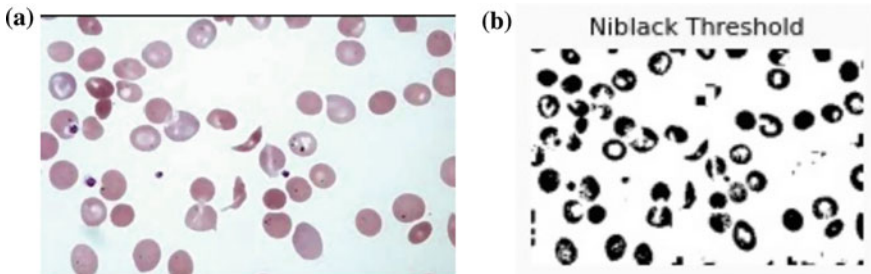


Fig. 4 a Original blood image b Ni-black thresholding


```
ID: 5, Label: Lukemia
/usr/local/lib/python3.7/dist-packages/tensorflo
warnings.warn("`model.predict_classes()` is de
```

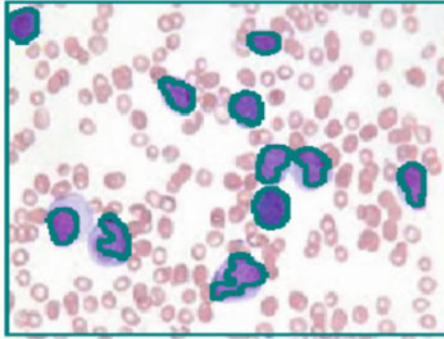


Fig. 5 Prediction of leukemi

4.2 Classification

This section shows the results obtained from the classification model built using Neural networks. The section is further divided into 2 parts—(a) Disease Identification, and (b) Performance Parameters for the Classification Model.

(a) Disease Identification

Following results show the prediction outputs of all the disease classes performed by the model after applying the neural network by training all the 7 disease classes of image datasets.

Leukemia detection: (Fig. 5)

Anemia detection: (Fig. 6)

MCL Lymphoma detection: (Fig. 7)

CLL Lymphoma detection: (Fig. 8)

FL Lymphoma detection: (Fig. 9)

Myeloma detection: (Fig. 10)

Malaria detection: (Fig. 11)

4.3 Performance Parameters for the Classification Model:

Following parameters which include the classification matrix and graphs of Epoch versus Loss and Epoch versus Accuracy are shown and explained.

1. Classification Matrix

The image below shows the classification matrix of the model. It is observed that most of the precision, recall and f1 scores for each disease class is 1.00 which means

```
ID: 1, Label: Anemia  
/usr/local/lib/python3.7/dist-packages/ter  
warnings.warn(`model.predict_classes()`
```

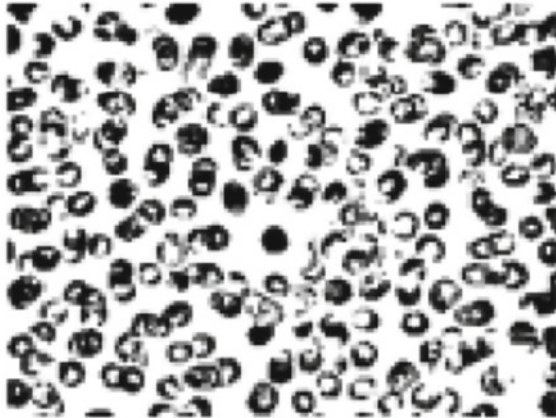


Fig. 6 Prediction of anemia

```
ID: 6, Label: MCL  
/usr/local/lib/python3.7/dist-packages/ter  
warnings.warn(`model.predict_classes()`
```

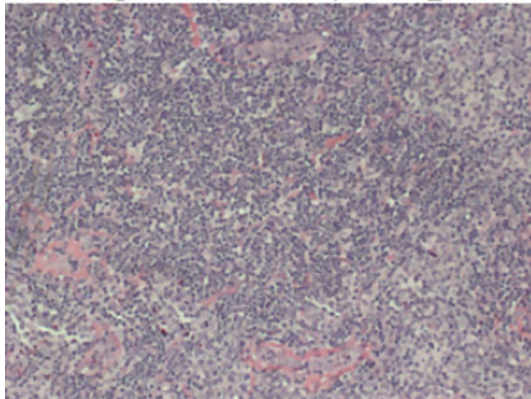


Fig. 7 Prediction of MCL lymphoma

that the model has excellent performance measures based on the number of positive predictions. Precision as 1.00 proves that the model was able to make 100% correct positive predictions. Recall as 1.00 proves that the model was able to classify 100% positive cases from the whole dataset of that particular disease class (Table 1).

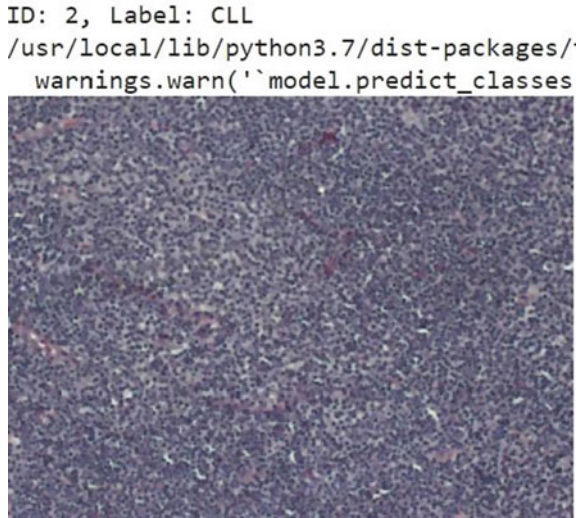


Fig. 8 Prediction of CCL lymphoma

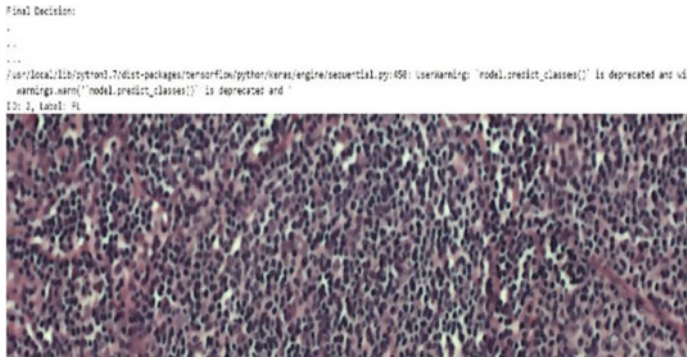


Fig. 9 Prediction of FL lymphoma

For example, Anemia has a precision and recall of 1.00 which means all the Anemia images were correctly classified and detected from the rest of the image data so there were 0 wrong predictions for Anemia detection whereas in the case of Healthy image classification, the precision score is 0.59 which implies there were 59% correct predictions of healthy images from the data and the remaining 41% were wrong predictions.

And finally, an F1 score takes both precision and recall into account to ultimately measure the accuracy for each disease classification. Following is the formula to calculate F1 score:

```
ID: 8, Label: myeloma  
/usr/local/lib/python3.7/dist-packages/tens  
warnings.warn(`model.predict_classes()` i
```

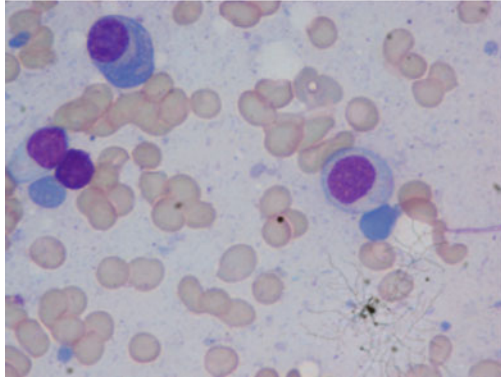


Fig. 10 Prediction of myeloma

```
ID: 7, Label: Malaria  
/usr/local/lib/python3.7/dist-packages/tens  
warnings.warn(`model.predict_classes()` i
```

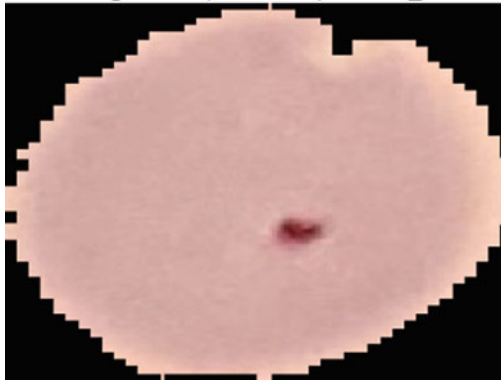


Fig. 11 Prediction of malaria

$$F_1 = 2 \times \left(\frac{precision \times recall}{precision + recall} \right)$$

As for the overall performance of the model, an accuracy of 98.6% was achieved with a loss of 0.47, which is pretty good for a multi class classification.

- 1.1 Epoch versus Loss
- 1.2 Epoch versus Accuracy

Table 1 Classification matrix of model

| | Precision | recall | F1 score | Accuracy (%) | Support |
|--------------|-----------|--------|----------|--------------|---------|
| Anemia | 1.00 | 1.00 | 1.00 | 100 | 20 |
| CL | 0.98 | 1.00 | 0.99 | 99 | 113 |
| FL | 1.00 | 1.00 | 1.00 | 100 | 139 |
| Healthy | 0.59 | 0.46 | 0.52 | 52 | 50 |
| Lukemia | 1.00 | 1.00 | 1.00 | 100 | 40 |
| MCL | 1.00 | 0.98 | 0.99 | 99 | 112 |
| Malaria | 0.56 | 0.68 | 0.61 | 61 | 50 |
| Myeloma | 1.00 | 1.00 | 1.00 | 100 | 227 |
| Micro avg | 0.96 | 0.94 | 0.94 | 94 | 811 |
| Micro avg | 0.79 | 0.79 | 0.79 | 79 | 811 |
| Weighted avg | 0.94 | 0.94 | 0.94 | 94 | 811 |
| Samples avg | 0.94 | 0.94 | 0.94 | 94 | 811 |

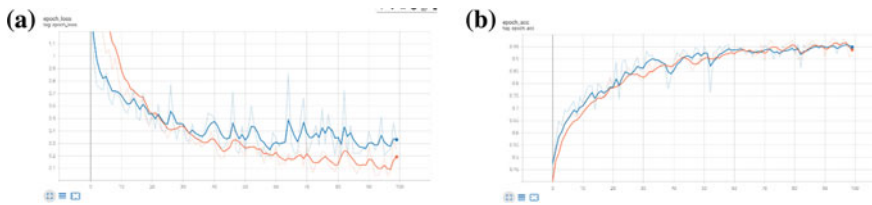


Fig. 12 a Epoch versus Loss graph b Epoch versus Accuracy graph

In the above graphs X-axis indicates the epoch and Y-axis indicates loss in Fig. 12a and accuracy in Fig. 12b. From this it can be seen that at the epoch range between 90 and 100, an optimized value of accuracy at 98.6% and loss value at 0.47 is obtained.

5 Conclusion

In this paper, an automated system has been developed to detect various hematologic diseases such as leukemia, anemia, lymphoma, myeloma, and malaria by using deep learning and image processing techniques. An accuracy of 98.6% with a minimum loss of 0.47 was achieved by the model and was able to classify 7 classes of these diseases from the healthy type. It can be concluded that an automated system saves time and is cheaper than manual testing methods. It can help medical practitioners to a great extent. This research can be further extended to other diseases and a real time hardware model can be developed which could be used in the public health sector as an effective tool.

Acknowledgements We would like to extend our heartfelt gratitude towards Prof. Shripad Bhatlawande, Head of Department, Electronics and Telecommunication Engineering for providing us with this golden opportunity. We would also like to thank Prof. Medha Wyawahare for her constant support and guidance in the right direction towards this paper. This paper would have been an uphill task without her continuous directional and unwavering support.

References

1. T. Xia, Y. Q. Fu, N. Jin, P. Chazot, P. Angelov and R. Jiang, "AI-enabled Microscopic Blood Analysis for Microfluidic COVID-19 Hematology," 2020 5th International Conference on Computational Intelligence and Applications (ICCIA), 2020, pp. 98–102, doi: <https://doi.org/10.1109/ICCIA49625.2020.00026>.
2. M. A. Al-Ameri, B. Ciyilan and S. D. S. Almassri, "Blood Diseases Detection Using Data Mining Techniques," 2021 6th International Conference on Computer Science and Engineering (UBMK), 2021, pp. 73–77, doi: <https://doi.org/10.1109/UBMK52708.2021.9558942>.
3. A. Galiano et al., "Improvements in haematology for home health assistance and monitoring by a web based communication system," 2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2016, pp. 1-5, doi: <https://doi.org/10.1109/MeMeA.2016.7533762>.
4. Balasubramaniam, Vivekanadam. "IoT based Biotelemetry for Smart Health Care Monitoring System." *Journal of Information Technology and Digital World 2*, no. 3 (2020): 183-190.
5. Rao, Smaran S., Gajanan Maske, and Antara Roy Choudhury. "Iris Image Segmentation and Localization using Dynamic Reconfigurable Processor." *Journal of Innovative Image Processing 2*, no. 3 (2020): 147–155.
6. L. Zhao, X. Tang and X. Yan, "The Study in Whole Blood Quality Control Materials of Hematology Analyze," 2007 IEEE/ICME International Conference on Complex Medical Engineering, 2007, pp. 357-360, doi: <https://doi.org/10.1109/ICME.2007.4381756>.
7. G. Ongun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac and S. Beksac, "An automated differential blood count system," 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001, pp. 2583–2586 vol.3, doi: <https://doi.org/10.1109/IEMBS.2001.1017309>.
8. W. Yu et al., "Automatic classification of leukocytes using deep neural network," 2017 IEEE 12th International Conference on ASIC (ASICON), 2017, pp. 1041–1044, doi: <https://doi.org/10.1109/ASICON.2017.8252657>.
9. P. T. Dalvi and N. Vernekar, "Anemia detection using ensemble learning techniques and statistical models," 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2016, pp. 1747–1751, doi: <https://doi.org/10.1109/RTEICT.2016.7808133>.
10. P. Jagadev and H. G. Virani, "Detection of leukemia and its types using image processing and machine learning," 2017 International Conference on Trends in Electronics and Informatics (ICEI), 2017, pp. 522–526, doi: <https://doi.org/10.1109/ICOEI.2017.8300983>.
11. S. Rajpurohit, S. Patil, N. Choudhary, S. Gavasane and P. Kosamkar, "Identification of Acute Lymphoblastic Leukemia in Microscopic Blood Image Using Image Processing and Machine Learning Algorithms," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 2359–2363, doi: <https://doi.org/10.1109/ICA CCI.2018.8554576>.
12. P. Kannadaguli, "Microscopic Blood Smear RBC Classification using PCA and SVM based Machine Learning," 2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT), 2020, pp. 82–86, doi: <https://doi.org/10.1109/MPCIT51588.2020.9350389>.

13. M. T. Vyshnav, V. Sowmya, E. A. Gopalakrishnan, S. Variyar V.V., V. K. Menon and K. Soman, "Deep Learning Based Approach for Multiple Myeloma Detection," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1–7, doi: <https://doi.org/10.1109/ICCCNT49239.2020.9225651>.
14. S. Tehsin, S. Zameer and S. Saif, "Myeloma Cell Detection in Bone Marrow Aspiration Using Microscopic Images," 2019 11th International Conference on Knowledge and Smart Technology (KST), 2019, pp. 57–61, doi: <https://doi.org/10.1109/KST.2019.8687511>.
15. R. Guilal, A. F. Bendahmane, N. Settouti, A. Benazzouz and M. A. Chikh, "Clinical and paraclinical factors selection for multiple myeloma diagnosis," 2019 International Conference on Advanced Electrical Engineering (ICAEE), 2019, pp. 1-6, doi: <https://doi.org/10.1109/ICAEE47123.2019.9014837>.
16. S. Domanskyi, A. Hakansson, G. Paternostro and C. Piermarocchi, "Modeling disease progression in Multiple Myeloma with Hopfield networks and single-cell RNA-seq," 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2019, pp. 2129–2136, doi: <https://doi.org/10.1109/BIBM47256.2019.8983325>.
17. Sanju and A. Kumar, "Classification of Multiple Myeloma Cancer Cells Using Convolutional Neural Networks and Transfer Learning," 2021 Asian Conference on Innovation in Technology (ASIANCON), 2021, pp. 1–6, doi: <https://doi.org/10.1109/ASIANCON51346.2021.9544557>.
18. S. P. Kamma, G. S. S. Chilukuri, G. S. Ram Tholeti, R. K. Nayak and T. Maradani, "Multiple Myeloma Prediction from Bone-Marrow Blood Cell images using Machine Learning," 2021 Emerging Trends in Industry 4.0 (ETI 4.0), 2021, pp. 1–6, doi: <https://doi.org/10.1109/ETI4.051663.2021.9619385>.
19. H H. Mohamed et al., "Automated detection of white blood cells cancer diseases," 2018 First International Workshop on Deep and Representation Learning (IWDRL), 2018, pp. 48-54, doi: <https://doi.org/10.1109/IWDRL.2018.8358214>.
20. N. V. Orlov et al., "Automatic Classification of Lymphoma Images With Transform-Based Global Features," in IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 4, pp. 1003-1013, July 2010, doi: <https://doi.org/10.1109/TITB.2010.2050695>.
21. D. J. Foran, D. Comaniciu, P. Meer and L. A. Goodell, "Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy," in IEEE Transactions on Information Technology in Biomedicine, vol. 4, no. 4, pp. 265-273, Dec. 2000, doi: <https://doi.org/10.1109/4233.897058>.
22. H. Li et al., "DenseX-Net: An End-to-End Model for Lymphoma Segmentation in Whole-Body PET/CT Images," in IEEE Access, vol. 8, pp. 8004-8018, 2020, doi: <https://doi.org/10.1109/ACCESS.2019.2963254>.
23. M. Z. Alom, C. Yakopcic, T. M. Taha and V. K. Asari, "Microscopic Blood Cell Classification Using Inception Recurrent Residual Convolutional Neural Networks," NAECON 2018 - IEEE National Aerospace and Electronics Conference, 2018, pp. 222–227, doi: <https://doi.org/10.1109/NAECON.2018.8556737>.
24. I. T. Young, "The Classification of White Blood Cells," in IEEE Transactions on Biomedical Engineering, vol. BME-19, no. 4, pp. 291–298, July 1972, doi: <https://doi.org/10.1109/TBME.1972.324072>.
25. J. M. Sharif, M. F. Miswan, M. A. Ngadi, M. S. H. Salam and M. M. bin Abdul Jamil, "Red blood cell segmentation using masking and watershed algorithm: A preliminary study," 2012 International Conference on Biomedical Engineering (ICoBE), 2012, pp. 258–262, doi: <https://doi.org/10.1109/ICoBE.2012.6179016>.
26. A. Gautam, P. Singh, B. Raman and H. Bhadauria, "Automatic classification of leukocytes using morphological features and Naïve Bayes classifier," 2016 IEEE Region 10 Conference (TENCON), 2016, pp. 1023–1027, doi: <https://doi.org/10.1109/TENCON.2016.7848161>.
27. R. R. Tobias et al., "Faster R-CNN Model With Momentum Optimizer for RBC and WBC Variants Classification," 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech), 2020, pp. 235–239, doi: <https://doi.org/10.1109/LifeTech48969.2020.1570619208>.
28. J. Rawat, H. S. Bhadauria, A. Singh and J. Virmani, "Review of leukocyte classification techniques for microscopic blood images," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 1948–1954.

29. I. Cseke, "A fast segmentation scheme for white blood cell images," Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol. III. Conference C: Image, Speech and Signal Analysis., 1992, pp. 530–533, doi: <https://doi.org/10.1109/ICPR.1992.202041>.
30. Habibzadeh M., Krzyzak A., Fevens T. (2013) White Blood Cell Differential Counts Using Convolutional Neural Networks for Low Resolution Images. In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L.A., Zurada J.M. (eds) Artificial Intelligence and Soft Computing. ICAISC 2013. Lecture Notes in Computer Science, vol 7895. Springer, Berlin, Heidelberg.
31. N. Christy Evangeline and M. Annalatha, "Computer Aided System for Human Blood Cell Identification, Classification and Counting," 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), 2018, pp. 206–212, doi: <https://doi.org/10.1109/ICBSII.2018.8524636>.
32. B. Venkatalakshmi and K. Thilagavathi, "Automatic red blood cell counting using hough transform," 2013 IEEE Conference on Information & Communication Technologies, 2013, pp. 267–271, doi: <https://doi.org/10.1109/CICT.2013.6558103>.
33. Sungheetha, Akey, and Rajesh Sharma. "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network." Journal of Trends in Computer Science and Smart technology (TCSST) 3,no. 02 (2021): 81-94.
34. Deepankan, B. N., and Ritu Agarwal. "A Two-Phase Image Classification Approach with Very Less Data." In International Conference On Computational Vision and Bio Inspired Computing, pp. 384–394. Springer, Cham, 2019.
35. <https://drive.google.com/drive/folders/1WQ5xND2XuQyuDOQTX4isOJUMfit9fJC?usp=sharing>
36. <https://drive.google.com/drive/folders/1wgzvWNOsrdrmiWB3BkQlwa4THgLDwSkk?usp=sharing>
37. andrewmvd larxel: Orlov, Nikita & Chen, Wayne & Eckley, David & Macura, Tomasz & Shamir, Lior & Jaffe, Elaine & Goldberg, Ilya. (2010). Automatic Classification of Lymphoma Images With Transform-Based Global Features. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society. 14. 1003-13. <https://doi.org/10.1109/TITB.2010.2050695>.
38. Anubha Gupta, Rahul Duggal, Shiv Gehlot, Ritu Gupta, Anvit Mangal, Lalit Kumar, Nisarg Thakkar, and Devprakash Satpathy, "GCTI-SN: Geometry-Inspired Chemical and Tissue Invariant Stain Normalization of Microscopic Medical Images," Medical Image Analysis, vol. 65, Oct 2020. DOI: <https://doi.org/10.1016/j.media.2020.101788>. (2020 IF: 11.148)
39. Shiv Gehlot, Anubha Gupta and Ritu Gupta, "EDNFC-Net: Convolutional Neural Network with Nested Feature Concatenation for Nuclei-Instance Segmentation," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 1389–1393.
40. Anubha Gupta, Pramit Mallick, Ojaswa Sharma, Ritu Gupta, and Rahul Duggal, "PCSeg: Color model driven probabilistic multiphase level set based tool for plasma cell segmentation in multiple myeloma," PLoS ONE 13(12): e0207908, Dec 2018. DOI: <https://doi.org/10.1371/journal.pone.0207908>
41. <https://ceb.nlm.nih.gov/repositories/malaria-datasets>. Photo by Egor Камелев on Unsplash <https://unsplash.com/@ekamelev>
42. Habibzadeh, Mehdi & Krzyzak, A. & Fevens, Thomas & Sadr, Ali. (2011). Counting of RBCs and WBCs in noisy normal blood smear microscopic images. Proc SPIE. 7963. <https://doi.org/10.1117/12.878748>.

Nonlinear Integrity Algorithm for Blockchain Based Supply Chain Databases



Mani Deep Karumanchi, J. I. Sheeba, and S. Pradeep Devaneyan

Abstract With an ever-growing global supply chain, supply chain management(SCM) is becoming more complex and competitive because of continuous development in global supply chain. In recent years, the blockchain has grown with the promise of becoming a “ trust machine “. However, this technological innovation can foster trust and transparency in many other sectors. In the industrial sector, for example, it could have interesting applications in production ecosystems and in particular in improving supply chain performance, offering the opportunity to increase speed and quality with potential cost savings ranging from 12 to 50%. the computation power required to setup and maintain a blockchain is expensive because of Computing hash for each block added to the blockchain and validating blocks. In this paper we proposed a nonlinear Integrity algorithm which works efficiently in terms of run-time and hash-bit variation with dynamic hash size and data size when compared it with traditional hashing algorithms like whirlpool, SHA, MD5.

Keywords Integrity algorithm · Blockchain · Information security · Supply chain management(SCM)

1 Introduction

SCM includes the control of all the supply chain activities which includes procuring raw materials, manufacturing goods, Shipping manufactured goods to end-user with the help of different SC participants using the logistics process. The main aim of SCM

M. D. Karumanchi (✉)

Department of Computer Science and Engineering, Bapatla Engineering College, Bapatla, India
e-mail: manideep.karumanchi@becbapatla.ac.in

J. I. Sheeba

Department of Computer Science and Engineering, Puducherry Technological University,
Puducherry, India
e-mail: sheeba@ptuniv.edu.in

S. P. Devaneyan

Department of Mechanical Engineering, Sri Venkateshwaraa College of Engineering and
Technology, Puducherry, India

is to offer high quality, and high-value products to its customers while maintaining a competitive edge [1]. The best customer service and products that a company offers to its customers will increase both the business growth and the profitability. Figure 1. shows the participants of Supply Chain.

Blockchain is a distributed, and growing list of records or blocks proposed by Satoshi Nakamoto[2] in2008 for Bitcoin which was used by many other cryptocur- rency networks later. Blockchain is a possible revolution in the financial world. But the other area where Blockchainshows great assurance is SCM. Blockchain can considerably enhance supply chains by allowing quick and reduced costs for deliver- ing the products, improving product traceability, enabling co-ordination between actors, and assist access to funds. A blockchain is a digital decentralized ledger for recording transactions among different parties in a tamperproof, andverifiablemanner [3] across a peer-to-peer network [4, 5]. Figure 2. As shown in Fig. 2. the transactions



Fig. 1 Participants of supply chain

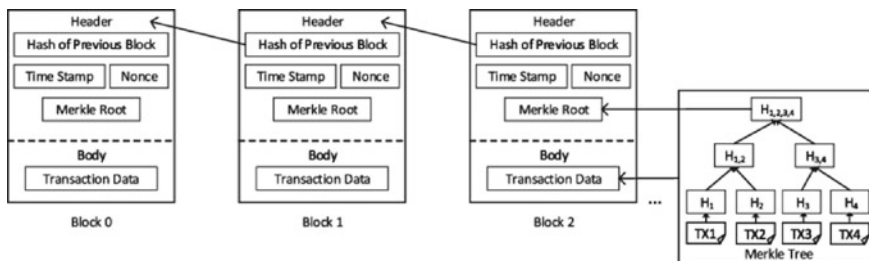


Fig. 2 The structure of a blockchain



Fig. 3 Supply chain members interacting with blockchain

included in the block will be hashed. Generally hashing processes the given data using mathematical functions and produces fixed-length output. Because of fixed-length output, it is difficult to decrypt the hash, and also guessing the input length is not possible. Shows the typical structure of Blockchain[6]. In cryptocurrency networks which are designed to substitute current currencies, the major role of blockchain is to allow an limitless number of transactions between multiple incognito parties with one another secretly in a secure way without having a central authority.

In supply chains, blockchain is used to allow a finite number of familiar persons to protect their business transactions against malicious users while offering better performance. Figure 3. Shows Interaction of supply chain members under a blockchain network. Recent blockchain applications which are successful for SCM require new permissioned blockchains, new guidelines for constituting transactions on a block, and new regulations to control the system all are in various phases of being developed. Supply Chains with Blockchain based solutions allows secured information sharing between supply chain partners, product quality monitoring, transparency and visibility throughout the supply chain[7, 8]. There are many examples of supply Chains which are successfully transformed using blockchain but there is still facing difficulty regarding security, privacy, usability, [9, 10] and cost [11].

2 Literature Survey

Azzi et al. [7] supply chain integrated with blockchain transformed to the authentic, transparent, reliable, and secure system. In this paper authors also explain various benefits that are brought by the blockchain into the SCM.

Dietrich et al. [12] discussed about various blockchain projects in SCM. They identified the main objective of blockchain applications in Supply chain is to improve transparency. they also pointed that majority of these projects are implemented on Supply Chains with simpler requirement like Food supply chains. They also pointed that the biggest limitation for blockchain technology in complex supply chains is scalability.

Agrawal et al. [13] implemented and presented a traceability based framework using blockchain. In this paper they demonstrated how blockchain ensure data safety, traceability trust in supply chain management using a smart contract.

Karumanchi et al. [14] demonstrated Cloud based Supply Chain Management challenges and blockchain improves security, and transparency and trust between trade partners.

Kshetri et al. [15] In this paper authors considered multiple case studies to show how blockchain affects supply chain management. they also address how blockchain can be used to achieve key SCM objectives like transparency and accountability.

Wang et al. [16] identified the main reason behind adopting blockchain in supply chain is trust. authors also identified the value of blockchain in SCM lies in extended visibility and traceability, supply chain digitalization and reduction in the use of intermediaries, improved data security and smart contracts.

Stefan et al. [17] analyzed multiple case studies to develop a descriptive prototype for the interaction of participants in a blockchain integrated with supply chain. Authors analyzed status quo to show impact of blockchain on logistics business.

Lee et al. [18] identified blockchain has potential to improve supply chain transparency and traceability in consumer electronic industry. they also identified blockchain technology is very beneficial in consumer electronics industry.

Agrawal et al. [19] discussed about how drug recall is a major issue in pharmaceutical industry and how blockchain effectively reduces this problem by monitoring the drug in the pharmaceutical supply chain with enhanced transparency and security throughout the process.

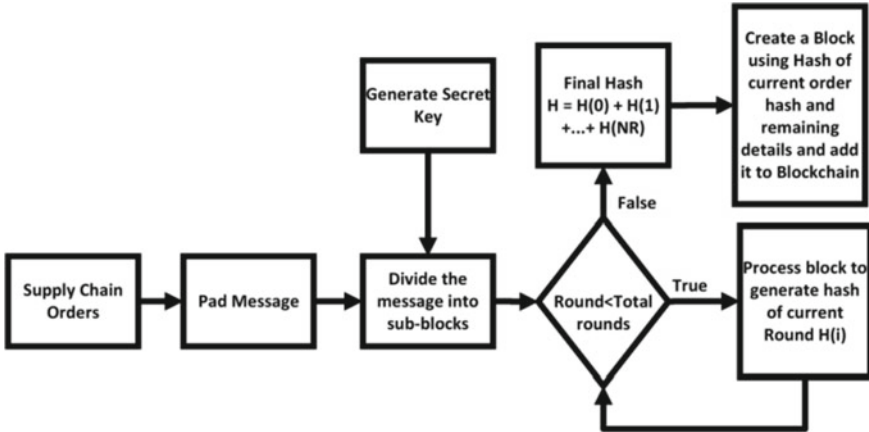


Fig. 4 Proposed framework

Smys et al. [20] proposed a secure smart vehicle system using blockchain framework. In this work they implemented a technique to identify when ecu fails or compromised and also to inform authorities before vehical causes any damage.

3 Proposed Technique

In this paper we design a conceptual model using Blockchain. Since Blockchain is being used for product tracking in supply chain only members of the supply chain can access the data or enter the data and also once entered into the Blockchain data can never be modified. The proposed framework is shown in Fig. 4.

The Hashing algorithm implemented to calculate the hash of supply chain orders is shown below.

```

1: Input: SID (Supply Chain ID)and D(Order Data), S(Hash Size), NR(Number
of Rounds).

2: Input Message: M={SID||D}

3: Initialize the round secret key SK in each iteration.

4: Partition M into blocks of size 8.

5: while(len(M)>0)
  Do
    If(len(M)<8)
      Pad message with sequence of ...0000001;
    else
      Step 6 Block Processing;
  Done
6: BlockProcessing
Divide the block into 32 bit size sub-blocks;
Sb[]=BlockPartition[S/32];
For i=0 to len(Sb)
  Do
    While(round<NR)
      Do
        Perform Step 7 Subblock Processing(Sb[i])
      Done
    Done
  Done
7: Sub block processing
For each byte in Sb[i]
  Do
    [Q R] = QRDecompose(SK)
    h1 = SKT . [Q, Rank(SK), (CLB(Poly(SK))) ]
    h2 = ( [ Q.trace SK,CLB(Poly(SK))] ) /
           (Σ SK[i])/max{solve(Poly(SK))}
    h3=Σ Solve(Poly(Q, R))
    H[i] = h1 ⊕ h2 ⊕ h3
  Done
8: Final Hash =Concat(H0||H1||H2||.....||Hn||);

```

In the above algorithm the input is Message which is the combination of SCM_ID and the Order data. The output is the final hash of the message. While generating the hash the algorithm generates a secret key used in the hashing process. In the next step the message will be divided into blocks of size 8 bytes to process quickly. in

the next step the blocks are sub divided into 32 bit blocks for non-linear transformation and then generate the hash for each sub-block using the above mathematical transformations. Finally combined the all sub-blocks hash to generate final hash.

4 Results

Experiments are simulated using Java, MySQL and AWS RDS. results are simulated using different datasets for the parameters bit-change variation and hash-runtime computation. Bit-change variation is defined as the number of hash bits change for a small change of input data. Sample supply chain dataset is shown in the Table 1.

Dataset URL: <https://github.com/mjosaarinen/weesrc/blob/master/corpus/cantenbury>

Table 2, shows the hash generated by the different traditional integrity algorithms for different datasets.

Table 3, shows the hash value of the proposed hash algorithm on the different sample training datasets. In this result, a large dynamic key(4096) hash size is generated on different data sets.

Figure 5. shows the runtime of the proposed algorithm is significantly less compared to the traditional hashing algorithms. So, the proposed algorithm reduce burden on the nodes which calculates the hash to propose a block in blockchain.

When we simulated the algorithm with different supply chain orders the algorithm shows huge variation in output hash for a small change in input data as shown in Fig. 6.

Table 1 Sample supply chain orders data

| Order company | Order no | Item no | Ordered quantity | Shipped quantity | Price | Line no | Invoice date | Document number | Reference |
|---------------|----------|---------|------------------|------------------|-------|---------|--------------|-----------------|-----------|
| 1295 | 118 | FULK215 | -710 | -710 | 2000 | 1 | 2016-12-30 | 54 | 85,311 |
| 1295 | 156 | FULK109 | -214.67 | -214.67 | 2001 | 2 | 2017-01-30 | 56 | 85,225 |
| 1295 | 1084 | FMCV20 | -200 | -200 | 2002 | 1 | 2018-02-28 | 157 | 85,481 |
| 1295 | 1084 | FMCV21 | -200 | -200 | 2003 | 2 | 2018-02-28 | 157 | 85,483 |
| 1295 | 2251 | FULK214 | 400 | 400 | 2004 | 1 | 2018-02-14 | 1573 | 85,485 |
| 1295 | 2462 | FULK237 | -28 | -28 | 2005 | 1 | 2018-03-31 | 1566 | 85,486 |
| 1295 | 2462 | FULK238 | -28 | -28 | 2006 | 2 | 2018-03-31 | 1566 | 85,488 |
| 1295 | 2462 | FULK162 | -60 | -60 | 2007 | 3 | 2018-03-31 | 1566 | 85,489 |
| 1295 | 2480 | FULK239 | -90 | -90 | 2008 | 1 | 2018-03-31 | 1567 | 85,490 |
| 1295 | 2516 | FULK141 | 90 | 90 | 2009 | 1 | 2017-03-31 | 1693 | 85,491 |
| 1295 | 2601 | FMCV19 | -65 | -65 | 2010 | 2 | 2018-08-31 | 1978 | 85,225 |
| 1295 | 2601 | FULK267 | -21.33 | -21.33 | 2011 | 3 | 2018-08-31 | 1978 | 85,311 |
| 1295 | 2601 | FULK245 | -19 | -19 | 2012 | 4 | 2018-08-31 | 1978 | 85,311 |
| 1295 | 2601 | FMCV26 | -15 | -15 | 2013 | 5 | 2018-08-31 | 1978 | 85,225 |
| 1295 | 2601 | FMCV28 | -10 | -10 | 2434 | 6 | 2018-08-31 | 1978 | 85,481 |

Table 2 Traditional hash algorithms results on different datasets

| | | | |
|------------|---|-----------|---|
| — | md5 | ALGORITHM | — |
| File Name: | alice29.txt | | |
| | 74c3b556c76ea0cfae111c4db64d08255 | | |
| File Name: | asyoulik.txt | | |
| | 2183e4e23c67c1dccc6cb84e13d8863bf | | |
| File Name: | cp.html | | |
| | d4b4e81b46ae7a3c3bc2b733bbd6d8cc8 | | |
| — | sha256 | ALGORITHM | — |
| File Name: | alice29.txt | | |
| | 7467306e0feed4971260f3c87421154a05be571d944e9cb021a5713700c38f0 | | |
| File Name: | asyoulik.txt | | |
| | ea3526fc53859f34ecd255712f9ecf0b2c903451d4755b2edaa2e2599cb0fc | | |
| File Name: | cp.html | | |
| | e0cd21cef5b6c4069461e949be100080c3ce887de6f1dd8626c480528efaaf61 | | |
| — | sha512 | ALGORITHM | — |
| File Name: | alice29.txt | | |
| | d93d674d66b227d7b3f4e1b7c35b102c40800e728b6ff68c7821109e7db7ad72f0b76a67bc9bd53b0202ac8daa0b22145f004dbdc6b59a48a6c8c72061bf1989f | | |
| File Name: | asyoulik.txt | | |
| | 38a78dc3ac6748f60195eadca9b39b637ae9df93feac45f08b610d1eccc4e9b8e89b22066f735c56397acd8f3ace1921295d77ba80607727b24f3880e1839c894 | | |
| File Name: | cp.html | | |
| | f26100cd21b26974b8eca2a9f24c753c5d413387abe055d62fc2707e244ab92e12538263c2bced8819253f10545786f91bd76b41482e2cde1c42e9f9887e6219 | | |
| — | whirlpool | ALGORITHM | — |
| File Name: | alice29.txt | | |
| | ca4847e73da8651f4544cb78af21d94d9a091a3d43755f4445828649b8a83f6907d3d650cfac4b53f0e8b982d6025dda670f439aa08e3ffea2fb79cd6f1c956 | | |
| File Name: | asyoulik.txt | | |
| | 2cf5df73be349eeff0a2560639e750f201621ce9434148bdea29e63cb7768c9c092bfb675f7ab87a9c28162ae74beeacc1e8a4a5e9cd2de6d4b5b8ce04c4523b6 | | |
| File Name: | cp.html | | |
| | 09330fe3c7799fd3f0709951607cce41641299043510809f8acf94f21e54e0111227a2c5f3ccba1f5ddb5d3c497804efc29f631b9927deef50b679a474dd7fd | | |

Table 3. Proposed Hash algorithm on different datasets

File Name:alice29.txt

```

e1805db332bcd629fb4ed37d4707b701553bc0e84bda8b1809ccad945389593ea340c7fd29a4a6d714e2dcbcb5da8ef8bc0ca582cd1136846a2e238a798da1829
add538abea5937aaeb0aec7103800e4bea019e7e088a68740c7261c404558a06332accdef5d70ca62a6e1b8b532c965f6ac2fe2700156ba8312cfiae5a37c8365b65
875bd131e7a428fee8d4604d70de80bbe507fa39e1858b4a9773eb64222eda96645cb234894bca63a8b57ee42c415ecd1812ee0cd8466f4e4ca7a02f09f3539a579
65e76b9b444a93d6b48039a1520cfa1f76253348b0bc99a2c328039f7376d3edd336cd4f6d68906b8d5d03b6368ff8c84e430e95f52331a4b552d11fa16672f2fa3e1
bfeca9de730689e1d57f1f5446a6c958cfc647f0c5e8b41f739c5830e23b96046bed86919e8025e5f83df94ce45ef1034f571dc202ecd593b9396a6c2155a7ab9
c7f167546f6adfa8a10aa5fa2e74095f00de817630ad3f053a523ea53df8049485df04505f10e4fa84ce28d3dabfe7643531510dde46145dff8558c381887bcaab6
5c0135675dca26005f7027b0b54e9c22382465d25b2cb0c8e8b4b900fadfe71695dc51f16d8a427dd99a6af406708bffd4420b696a067ac579dca44e2cfa5f099
fc468b77a1031f73acb264aef03316138f722f6d5cc0aa3f980e950de7755cb569849df5c0ce19550b7c3cfl708f42ec1634f966b3b36c8e3e4
e1805db332bcd629fb4ed37d4707b701553bc0e84bda8b1809ccad945389593ea340c7fd29a4a6d714e2dcbcb5da8ef8bc0ca582cd1136846a2e238a798da1829
add538abea5937aaeb0aec7103800e4bea019e7e088a68740c7261c404558a06332accdef5d70ca62a6e1b8b5532c965f6ac2fe2700156ba8312cfiae5a37c8365b65
875bd131e7a428fee8d4604d70de80bbe507fa39e1858b4a9773eb64222eda96645cb234894bca63a8b57ee42c415ecd1812ee0cd8466f4e4ca7a62f09f3539a579
65e76b9b444a93d6b48039a1520cfa1f76253348b0bc99a2c328039f7376d3edd336cd4f6d68906b8d5d03b6368ff8c84e430e95f52331a4b552d11fa16672f2fa3e1
bfeca9de730689e1d57f1f5446a6c958cfc647f0c5e8b41f739c5830e23b96046bed86919e8025e5f83df94ce45ef1034f571dc202ecd593b9396a6c2155a7ab9
c7f167546f6adfa8a10aa5fa2e74095f00de817630ad3f053a523ea53df8049485df04505f10e4fa84ce28d3dabfe7643531510dde46145dff8558c381887bcaab6
5c0135675dca26005f7027b0b54e9c22382465d25b2cb0c8e8b4b900fadfe71695dc51f16d8a427dd99a6af406708bffd4420b696a067ac579dca44e2cfa5f099
fc468b77a1031f73acb264aef03316138f722f6d5cc0aa3f980e950de7755cb569849df5c0ce19550b7c3cfl708f42ec1634f966b3b36c8e3e4
File Name:asyoulik.txt
1b28eb07f0c504286bb6d90993ec07a311ee72b60f37412c730a11cdd46fe32b34897f12383321597356efb879c5c0d04894d67fa447fec0fb7a8e7f4282a75fa83
fbfffd514b4b135e5cffa4f3105126a23a912b95a639f5217fb6c27b3aa8e92ad97f7cbae28a389ff7f05dbb52b6527b09dee3422df886e6b3300129687045c469387
e0520d64abe72116028b82e6a606928c24b217f5e3b794feb61cc30744732d15413ac93c0f2459ff8ce46281298d0822288d716f7a1e790e6a6d5efda4c19d6180
8e800acbb1aaecf2d6163f88f87c3f21501b34d57217b305ac5b97e7934e519bd843449fb95cfc805026506e1d8d2121b89c099aae72e44a16f6d78f849a46
5dd27fec3d9cbe1ef0rc0d0448cca18b03116b1ac521efeb16908e78c74c5ede988e003aeaa6d441b8b2807e2944a33c499f3e8d8ce480e22c512b05795695c694
63204fa4109e39029ef10053193ca1eb3ee8706b10367ac6adfab404ec137f56548d2f7797c13b214228a96ff1e3c3f725e50780eaa96d6be857cd8e80c3eaf6f3
512e17c488ae8b03feed419c14442877c7ec7f345bee7a9a5e3a15fd4e077e10091dbf2da3f2b08fb60ad6029571385fda816f81dcb2b9e87351a75006d44e40ad7
bb655a6f0af0bb3f258ef64c3043d9d3e78bdaf9a2263b0fe3db61cbbd9120a5a62d93ce475d259d6dce6893035ef5f98971dac3c334717df

```

(continued)

Table 3 (continued)

1b28eb07f0c504286bb6d90993ec07a311ee72b60f37412c730a11cd46fe32b373f112383321597356efb879c5c0d04894d67fa4d7fec0fb7a8e7f4282a75fa83
fbfffd514b4b135e5cffa4f3105126a23a912b95a639f5217fb6c27b3a8e92ad97f7cbae28a839ff7f05d5b26527b09dee3422df886e6b33001296687045c469387
e052d64abe72116028b82e6a060928c24b217fe3b794feb61cc30744732d15413ac93c0f2459ff8ece46281298d0892288dd716f7a1e790e6a6d5edfiac19d6180
e80000acb1aaedaf2d6163f88f87c3f21501b3d57217b305ac5b97e7934e4c519bd843449fb9b5c8f05026506elddd21b8b9c099aae72e44a16f6d78f8da9460
5dd27eeb3a9cbl1ef0fc0d0448cca18b03116bl1ac521efeb16908e78c74c5ede988e003aeaa6d41b8b2807e2943a3c499f3ead8ce480e22c512b05795695c694
63204fa4109e39029efl10053193ca1eb3ee8706b10367ac6adfabab404ec137f56548d2ff797c13b214228a96ffbf3c3f725e50780eaa96d6bce857cd8e80e3eaf6f3
512e17c488ae8b03feed419c14442877c7ec7f345bee7a9a5e3a15fd4e077e10091dbf72da3f2b08fb60ad6029571385fda816f81dcb2b9e87351a75006644e0ad7
bb655a60fa0bb3f258ebf64c3043d9d3e78bdaf9a2263b0fe3db61cbdb9120a5a62d93c475d259d6dc6893035ef5f98971dac3c334717df

File Name:cp.html

eb3f8938ac2030620512c6c15b3a79a5f7dac7820ff3e7532c84132d078f5df61ca798de55ceee76c20849f85f3624b9fda682ca42f97323d41239998072c62142
46262c60eb1a3e749a4877981235304e2ff2f05aac4f5134cab1d01cb10ce0db9b9a262755c01721fdeb5913d79859287af7e4bc5e78c09510645056918ffb0377
df26fb3a2e569d39ce45bae8035979562770e437089e13f5b86cf555b3486f6802e110e63cbf5cc2dc8049196b33ba0d6bd9d24bd3d7d7988cdec39875d2e33
b8953cd657f78f5bdeb69aeb6fcb0b09813332527bd6f832cafc4939725902e4b8c6de7138e74cbfa1f4c7bb7d14e71732cb4b4ad97e43aceae62f612dc022316
45ee9e9b920a993ee75ae2d24419a60ff4c85aa6a54f97038118fb220136252ff92d19c55490d507c177e9fae4b0f4758ae322631454563151d976437bf65b556
946b3d3ad317e98be567e2605c4f3447243c86ca75d1ef1a5d00fd59979f1762f9e092caaea30e4bb661e1296b371758613ce89d69f9c896a04d70cb97f3
951f814e40d743c528be5c08e7220456728ffeaaf45d6827d1f69f5b9e795370e6a3ef2a9190fe473d604c5c523abaece08a551c56970d35801fa2914ead7d57b92
bdce3e6cc56947874e5271df03bd5a5d9b19ed9cb029cbadf793906b6e1230d5b272602e437ce049909843faceed42a7baf0988e7c3fcc4ab
eb3f8938ac2030620512c6c15b3a79a5f7dac7820ff3e7532c84132d078f5df61ca798de55ceee76c20849f85f3624b9fda682ca42f97323d41239998072c62142
46262c60eb1a3e749a4877981235304e2ff2f05aac4f5134cab1d01cb10ce0db9b9a262755c01721fdeb5913d79859287af7e4bc5e78c09510645056918ffb0377
df26fb3a2e569d39ce45bae8035979562770e437089e13f5b86cf555b3486f6802e110e63cbf5cc2dc8049196b33ba0d6bd9d24bd3d7d7988cdec39875d2e33
b8953cd657f78f5bdeb69aeb6fcb0b09813332527bd6f832cafc4939725902e4b8c6de7138e74cbfa1f4c7bb7d14e71732cb4b4ad97e43aceae62f612dc022316
45ee9e9b920a993ee75ae2d24419a60ff4c85aa6a54f97038118fb220136252ff92d19c55490d507c177e9fae4b0f4758ae322631454563151d976437bf65b556
946b3d3ad317e98be567e2605c4f3447243c86ca75d1ef1a5d00fd59979f1762f9e092caaea30e4bb661e1296b371758613ce89d69f9c896a04d70cb97f3
951f814e40d743c528be5c08e7220456728ffeaaf45d6827d1f69f5b9e795370e6a3ef2a9190fe473d604c5c523abaece08a551c56970d35801fa2914ead7d57b92
bdce3e6cc56947874e5271df03bd5a5d9b19ed9cb029cbadf793906b6e1230d5b272602e437ce049909843faceed42a7baf0988e7c3fcc4ab

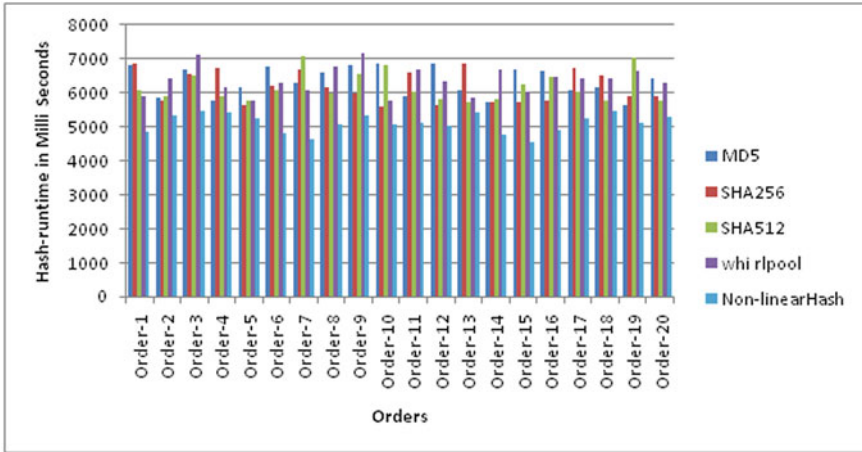


Fig. 5 Hash-runtime comparison of proposed algorithm with the traditional hashing algorithms

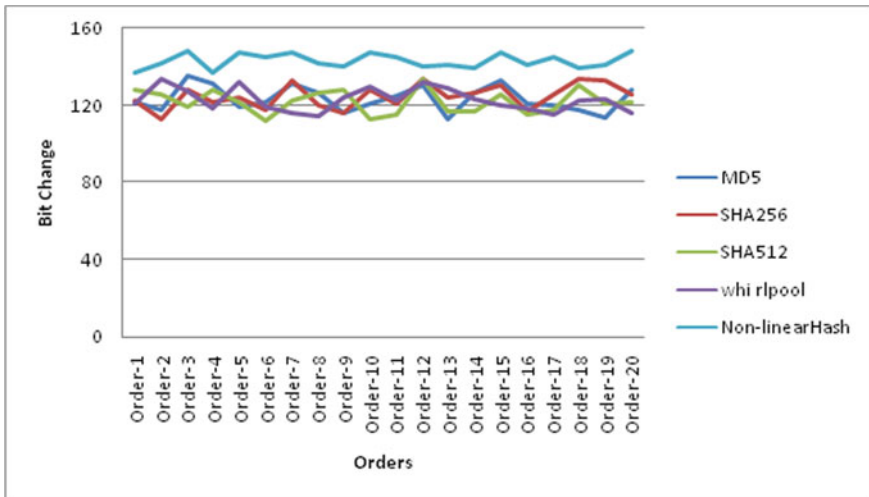


Fig. 6 Bit change comparison of proposed algorithm with the traditional hashing algorithms

That means the proposed technique is effective in terms of security and collision compared with the other traditional techniques.

5 Conclusion

There is considerable space to further develop supply chains as far as start to finish detectability, speed of item conveyance, co-ordination, and financing. Blockchain can be a useful asset for addressing the traditional SCMissues. It is presently an ideal opportunity for SCM network supervisors who are remaining uninvolved to survey the capability of blockchain for their organizations. Main difficulty in building a blockchain is generating hash of the transactions with the good use of the system resources. In this paper, a novel non-linear integrity computational algorithm is designed and implemented on supply chain datasets. Experiment results proved that the proposed algorithm works effectively compared with the traditional algorithms.

References

1. Guercini, Simone, and Andrea Runfola. "The integration between marketing and purchasing in the traceability process." *Industrial Marketing Management* 38.8 (2009): 883-891
2. Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." *Decentralized Business Review* (2008): 21260
3. Choi, Tsan-Ming. "Creating all-win by blockchain technology in supply chains: Impacts of agents' risk attitudes towards cryptocurrency." *Journal of the Operational Research Society* 72.11 (2021): 2580-2595
4. Chang SE, Chen YC, Lu MF (2019) Supply chain re-engineering using blockchain technology: a case of smart contract based tracking process. *Technol Forecast Soc Change* 144:1–11
5. Choi T-M, et al (2019) The mean-variance approach for global supply chain risk analysis with air logistics in the blockchain technology era. *Transport Res Part E Logist Transp Rev* 127:178–191
6. Liang Y-C (2020) Blockchain for dynamic spectrum management. In: *Dynamic Spectrum Management*. Springer, Singapore, pp 121–146
7. Azzi, Rita, Rima Kilany Chamoun, and Maria Sokhn. "The power of a blockchain-based supply chain." *Computers & industrial engineering* 135 (2019): 582–592.
8. Frizzo-Barker, Julie, et al. "Blockchain as a disruptive technology for business: A systematic review." *International Journal of Information Management* 51 (2020): 102029.
9. Peck, Morgen E. "Blockchain world-Do you need a blockchain? This chart will tell you if the technology can solve your problem." *IEEE Spectrum* 54.10 (2017): 38-60.
10. Firica, Oana. "Blockchain technology: Promises and realities of the year 2017." *Quality-Access to Success* (2017).
11. Choi, Tsan-Ming. "Supply chain financing using blockchain: Impacts on supply chains selling fashionable products." *Annals of Operations Research* (2020): 1–23.
12. Dietrich, Fabian, et al. "Review and analysis of blockchain projects in supply chain management." *Procedia Computer Science* 180 (2021): 724–733.
13. Agrawal, Tarun Kumar, et al. "Blockchain-based framework for supply chain traceability: A case example of textile and clothing industry." *Computers & industrial engineering* 154 (2021): 107130.
14. M. D. Karumanchi, J. I. Sheeba and S. P. Devaneyan, "Cloud Based Supply Chain Management System Using Blockchain," 2019 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), 2019, pp. 390–395, doi: <https://doi.org/10.1109/ICEECCOT46775.2019.9114692>.
15. Kshetri, Nir. "1 Blockchain's roles in meeting key supply chain management objectives." *International Journal of Information Management* 39 (2018): 80-89.

16. Wang, Yingli, Jeong Hugh Han, and Paul Beynon-Davies. "Understanding blockchain technology for future supply chains: a systematic literature review and research agenda." *Supply Chain Management: An International Journal* (2019).
17. Tönnissen, Stefan, and Frank Teuteberg. "Analysing the impact of blockchain-technology for operations and supply chain management: An explanatory model drawn from multiple case studies." *International Journal of Information Management* 52 (2020): 101953.
18. Lee, Jong-Hyouk, and Marc Pilkington. "How the blockchain revolution will reshape the consumer electronics industry [future directions]." *IEEE Consumer Electronics Magazine* 6.3 (2017): 19-23.
19. Agrawal, Divyansh, et al. "A robust drug recall supply chain management system using hyperledger blockchain ecosystem." *Computers in biology and medicine* 140 (2022): 105100.
20. Smys S, and Haoxiang Wang. "Security enhancement in smart vehicle using blockchain-based architectural framework." *Journal of Artificial Intelligence* 3.02 (2021): 90–100.

Topic Identification and Prediction Using Sanskrit Hysynset



Prafulla B. Bafna and Jatinderkumar R. Saini

Abstract The topic model implementation is not a new concept for English corpus due to the availability of plenty of resources, but developing a topic model for Sanskrit is comparatively an untouched area. The proposed approach is a 4 phased. The first phase constructs Hysynset followed by building a topic model that acts as a second phase. In the third phase, clustering is applied and the approach completes with classification and prediction that is the fourth phase. Hypernyms-hyponyms and synonyms are grouped in the first phase to reduce the dimensions and creates semantic space. The topic model is built using Latent Dirichlet Allocation (LDA) which shows very specific and informative topics as it uses Hysynset vector space model for Sanskrit (HSVSMS). The dataset belongs to more than 1100 Sanskrit stories. The documents' wise topics are presented using dendrogram obtained after applying HAC and then supervised model that is random forest is used to predict the topic of the test/new document and evaluated using classification error and accuracy. In the absence of the availability of standard experiments, current work could not be compared with other existing work in case of prediction of stories. Comparative analysis of topic identification using existing technique Vector Space Model for Sanskrit (VSMS) proves that betterment of the proposed technique that is (HSVSMS) in the form of the accuracy, misclassification error of classification, coherence score, entropy and purity and topic titles.

Keywords Sanskrit · Topic model · Vector space model · Latent Dirichlet Allocation (LDA) · Classification · Clustering · Entropy · Accuracy

P. B. Bafna (✉) · J. R. Saini

Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, India

e-mail: prafulla.bafna@sicsr.ac.in

1 Introduction

For businesses, exponentially generated data has become a challenge. It can be converted to an opportunity if required information is extracted from the data for decision-making. For example, reviews data is useful for identifying flaws in a product/service and rectify it. Many applications like recommender systems [1] topic extractions are benefited due to decision-making capacity. Text mining plays an important role by processing and analyzing data in the form reviews, comments, and documents to reach a particular decision. It converts unstructured (text) data into actionable form. Different algorithms are used to generate important insights of data. Unstructured data may be present in different languages including English [2]. Processing data in the English language is comparatively easy than processing in other regional languages that is Sanskrit [3] Marathi [4] and so on. Several pre-processing steps are needed to analyze the text data. The collection of documents/text is termed as a corpus [5]. The documents are composed of several paragraphs, which are composed of several sentences. The sentences are broken into tokens. For e.g., for the sentence ‘उदारचरितानां तु वसुधैव कुटुम्बकम्’, tokens are, ‘उदार’ ‘चरितानां’, ‘तु’,.The second step after tokenization is the removal of stop words or considering only important words. PoS tagging carried out to identify nouns, adjectives and so on. Pronouns, numeric, special characters are considered as stop words and are ignored For e.g., words like ‘तु’ reduce noisy features and allows dimension reduction to improve the performance of the algorithm. Lemmatization is carried out to reduce the token into its meaningful root form. Unique lemmas are identified which again gives the exact number of tokens present in the corpus. The most recurrent tokens or concepts can be identified using word frequency. For e.g., stories can be processed and their topics can be identified. To extract topics, a bag of words based on the frequency of the words can be used which is one of the vectorization models. As there is ambiguity in the language used by humans, the single word may be used in various contexts. To extract the semantics or context of the word concordance is used. Exploring the concordance of a concept/word results to identify the exact meaning based on context. Context identification experiments are done using the TF-IDF [4] followed by a vector space model for Sanskrit (VSMS) Latent Dirichlet allocation (LDA) is applied on a dataset like Marathi [6], Sanskrit stories, poems, etc. In LDA group of words are mapped with topics and topics are associated with documents. For e.g., ‘Jatakkatha’ are one of the types of stories/literature consisting of previous exists of the Buddha and also teaches lessons of life. Clustering is used to group similar objects/stories. Hierarchical clustering is one of the popularly used techniques. Agglomerative is the type of hierarchical clustering and known as hierarchical agglomerative clustering (HAC). It is initiated with one object representing one cluster [7].

It successively merges the clusters based upon the distance between the documents. Dendrogram that is a tree like representation of clusters is used for visualization. Entropy [8] and purity are used for performance evaluation. Minimum entropy means the minimum error and is evaluated using the clustering process and

maximum purity ensures the maximum distance amongst clusters. Classification is an unsupervised technique in which labeled data is divided into training and testing datasets. The classification model is built on a training dataset and its prediction is carried out using a testing dataset. The confusion matrix is produced which uses known true and false values. Several evaluation parameters like accuracy, error are used to evaluate the performance of the classifier which is based on correctly classified entities.

Topic modeling [9] is unsupervised learning and represents topics of the corpus. It explains the information of included documents. Clustering of words is done to represent topic. Topic is represented by specific set of documents. LDA is one of the techniques to implement topic model. It uses two matrices, which consists of distribution of topics in the document and distribution words across the topics. Topic coherence is used to evaluate LDA. The degree of semantic relevance is measured by topic coherence [6]. The top scoring words in the topic are used for calculations. It interprets and differentiates topic semantically using statistical inference. example “Football is a popular game”. “Sports need fitness” are coherent statements. Purity assesses the cluster quality which is an external evaluation measure. It is based on correctly classified entities and has value between [0.0,1]. The value near to 1 indicates more pure clusters. Entropy shows disordering that is incorrectly classified points. The clusters with lower entropy are termed as pure clusters. It depicts that objects in the same cluster should not show different behavior. β constant, $\beta(j, k)$ represents the probability of the j the topic containing the k th word. β is used to show Dirichlet distribution.

Hysynset is used to reduce dimensions and involve semantics of the terms in the process of classification and clustering. It is the group of inter relevant terms that is hypernyms-hyponyms and synonyms and termed as hysynset. Sanskrit Wordnet has all components of hysynset. For e.g., ‘पशु’, ‘प्राणनि’, ‘मृग’ are present in a one group. ‘पशु’, ‘प्राणनि’ meaning animal and ‘मृग’ means deer.

Random forest is a useful classifier when the problem is multiclass. Most of the other classifiers produce good results for binary class problems that is why the random forest is suitable and popular for text classification. It produces different decision trees and error gets stabilized as algorithm converges [2, 3, 10, 11].

This research is unique because

1. Hysynset is used as dimension reduction technique which in turn results into involving semantics and context.
2. Supervised learning and topic modeling (unsupervised) is carried out using hysynset for Sanskrit corpus.
3. Comparative analysis of VSMS and Hysynset vector space model for Sanskrit (HSVSMS) techniques is performed using entropy, purity, coherence measure, number of topics, contents of topics, accuracy and classification error.

2 Literature Review

A method based on a novel feature extension to classify text by considering graph of topic-keyword and analysis of the link is proposed. Biterm topic model (BTM) is used which extracts the topics and ranks them. The topics and their key words are reranked. The most relevant key terms are detected using the K-L divergence technique. It is added with the structural based similarity of the corpus. The related keywords are appended in the short text to perform classification on two corpora with validation parameters [12]. The contextual analysis of tweets is carried out by identifying the context by using Wikipedia. It is a challenging task due to several parameters belonging to tweets like noisy, ambiguous features, etc. Classification of tweets is carried out using Twitter Name Entity Recognition. Specific entities are selected and a new dataset is constructed using Wikipedia. LDA was implemented to extract the topics. The effect of representation of the model, topics nature, coherence score, perplexity measure is used for both datasets to compare and prove the efficacy of the proposed technique. The dataset was collected from News, Reuters, and CNBC using API [13].

Sentiment analysis [14, 15] supports many applications [16–18] in a social domain such as customer segmentation, market understanding and predictions, and so on. The sentiment analysis is carried out on big data. The detection of sentiment polarity is carried out, also existing technique of sentiment analysis is explained in the form of reviews. The open challenges in the field of sentiment analysis performed on the big data are stated [19].

Sanskrit texts containing a 3.5 million-tokens is prepared to carry out document as well as word-level analysis. Latent Dirichlet Allocation (LDA) is used to build a topic model. Corpus consistency problems are solved. Sanskrit pramāṇa texts are considered for the experimental purpose. Also, issues faced with LDA implementation of pramāṇacorpora are focused. Supervised methods with collaborative and online annotation mechanisms can improve the processing of the text considering corpus-linguistic and result into optimized information retrieval [12, 20, 21].

Less availability of resources has resulted into less work on Indic languages like Sanskrit [22–24] For e.g., Grouping the corpus based historical or cultural importance etc. Two topics are focused that is, extracting specific words in the scripts using pragmatic knowledge and the need of having a digital library to deliver access to a corpus with multiple scripts. Script free Keyword extracting is proposed which fulfills the need of intelligent frame to scale through diverse scripts. Script specific approach uses a Block Adjacency Graph (BAG) representation and another approach uses an image moment that processes multi-script corpus [25].

Support Vector Machine, Naïve Bayes and Random Forest are implemented along with Rapid Miner with a dataset of 500 scientific records. Misclassification error and accuracy of algorithms are compared to depict the best algorithm. The Density plot is different for all algorithms and SVM is more dense indicates more accuracy than other two algorithms Rapid miner allows faster and accurate execution and hence combined with all algorithms. With the different algorithms for

getting faster and accurate results. Topic human topic grading/ranking and topic coherence are assessed using LDA on the full text. Vocabulary size, word length of document and frequency of document affect the results. The journal articles are classified using various parameters like their output, year of introduction and so on. Pragmatic knowledge is used for optimal LDA models. The less dataset has more noise as compared to big dataset. Topics extracted from abstract data are completely different than that of full text. To analyze the reasons could be future direction of research [5].

The proposed techniques uses coherence of document to extract the information. Both models use document coherence which uses discourse entities. It uses either subject or object of a sentence. In the first model, text is used as Markov process. It generates sequences of entities (n-grams). The entropy is used to control the rate of newly appearing text. It benefits to minimize text drift and handle text coherence issues. In the second model, the text is represented as discourse entities graph. It is connected by various relations that is adjacency in the text. Graph topology metrics is used to incorporate various aspects of the discourse run to specify coherence, that is average clustering or separation of discourse entities in the data. Comparative analysis is performed using text coherence and document relevance. The proposed model can be used to compute document ranking using comprehensibility and cohesiveness of text [2]. Being popular technique topic modeling is comparatively untouched for clustering text documents. There are various issues, such as computing optimal parameters of model and ensuring semantic stability, etc. Entropy is used as an evaluation parameter. Gibbs sampling is used along with LDA and PLSA implementation on multilingual datasets. Entropy is only a single metrics which assures all evaluations simultaneously that is hyper-parameters and number of topics as well as semantic stability. Statistical concept is used to assure machine learning behavior [21].

3 Research Methodology

The experiments carried out to mine different topics from the corpus of stories are explained in the current section. The corpus creation and its preprocessing are followed by measuring TF-IDF measures of the tokens. VSMS [4] and HSVSMS based on hysynset are constructed. It includes hypernyms-hyponyms and synonyms relationship. In Fig. 1, 'पशु' (animal) and 'प्राणनि' (Animal) are synonyms, whereas 'मृग' (deer) and 'प्राणी' (animal) are termed as hypernyms-hyponyms. This entire group is known as hysynset. The terms present in the hysynset are weighted using TF-IDF measures. For e.g., 'पशु', 'प्राणनि', 'मृग' has TF-IDF measures as 0.7, 0.1, and 0.1 respectively. The total feature weight of a hysynset is $0.7 + 0.1 + 0.1 = 0.9$. LDA is built using VSMS and HSVSMS. The comparative analysis is carried out using several parameters such as topic coherence, number of topics, words present in the topics, dendrogram structure and cluster validation parameters, etc. Clustering is followed by classification to predict the class of test stories. The prediction is evaluated using several parameters like accuracy,

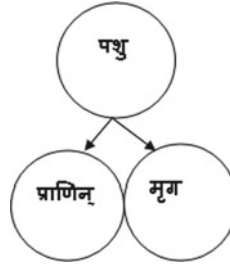


Fig. 1 Words in hysynset

misclassification errors, etc. The steps performed to carry out experimental analysis are shown in Fig. 2.

Sanskrit stories are processed using library `udpipe` (<http://cran.r-project.org/web/packages/udpipe/vignettes/udpipe-annotation.html>). The open-source ‘R’ programming language is used for the creation of the corpus. The Sanskrit language model is used by downloading, loading and calling the model using the commands `model_sanskrit ← udpipes_download_model(language = “Sanskrit”)`, `udmodel_sanskrit ← udpipes_load_model()`, and `file = “Sanskrit-ufal-ud-2.4–190,531.udpipe”` respectively. Modules used for preprocessing corpus, corpus annotation are executed using the same library for e.g., `udpipes_annotate(udmodel_sanskrit)`. The steps in research methodology are described in detail.

3.1 Constructing a Corpus and Preprocessing

Data is downloaded from different websites (www.matrubharti.com/novels/sanskrit, <https://sanskrit.pratilipi.com/sanskrit-short-stories-pdf-free-download> <https://www.hindujagruti.org/hinduism-for-kids-sanskrit/category/>) in the form of stories. A total

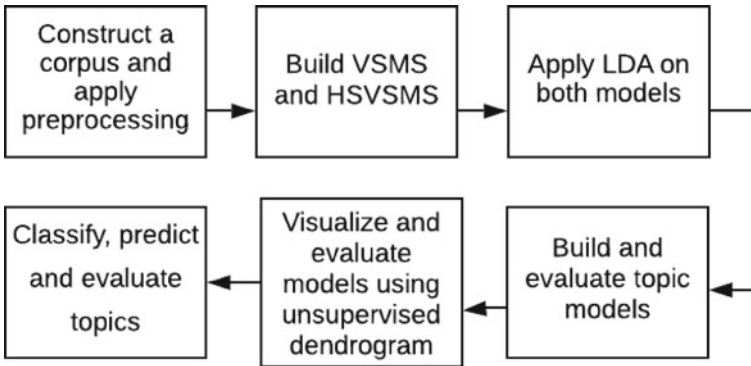


Fig. 2 Steps in research methodology

of 1127 stories are considered for experimental purposes. Encoding style needs to be used to store Sanskrit text is 'UTF8'. Tokenization is carried out which resulted into 1,91,234 tokens. Library udpipe available in 'R' provides a POS tagger to identify nouns, adverbs and so on. Including adverbs, adjectives, nouns, etc. and ignoring pronouns, numeric etc., automatically removes the noisy features in the form of stop words and special characters. The total number of tokens retrieved is 92,218 at this stage. The next step is lemmatization and a total count of lemmas is 46,189.

3.2 Construct a Vector Space Model for Sanskrit (VSMS) and HSVSMS

The TF-IDF measure is computed for all the lemmas. The terms having weight more than 0.75 are considered for the next step. Total 7,123 significant terms are retrieved which are used to build VSMS. LDA uses VSMS to build topic model. Alsohysynset vector space model for Sanskrit (HSVSMS) is built using the hysynset. It is a group of similar terms which include hyponyms-hypernyms and synonyms. Using initials, the group is it is said to be as hysynset. TF-IDF measures of all relevant terms existing in a hysynset are summed up, the addition is called as hysynset frequency. All the terms and hysynset groups having more than 0.75 are input to HSVSMS and called as features of HSVSMS.

3.3 Topic model's Intrinsic Evaluation

To evaluate topics retrieved using the topic model, the coherence measure is computed. Total 3 topics are discovered using both of the models that is VSMS and HSVSMS. For VSMS the coherence score is 0.7 and for HSVSMS it is 0.9. VSMS and HSVSMS act as topic model parameters and finds out the topic of each story. The probability of feature vector with respect to each topic is found out and the maximum value of probability is used to decide the topic of the document. To generate top informative words for each topic, probabilities are used. Both topic models that is using VSMS and HSVSMS come up with different features. Both the models derive 3 topics, but features generated using HSVSMS are more specific and results to identify the exact and specific topic of story. These words being significant are ignored by VSMS but are included by HSVSMS as they exist in hysynset.

3.4 Visualization Using Dendrogram

HAC generates 3 clusters using both models. Each cluster represents the topic. To verify the topic assignment of the documents, entropy and purity used. Entropy and purity are calculated for both models and are compared for varied data size. Both evaluation parameters that are entropy and purity are consistently better for HSVSMS.

3.5 Apply Classifier Using HSVSMS Prediction and Performance Evaluation

The algorithm is applied on HSVSMS as the label of each document is known in the form of a detected topic. The dataset is divided into 70–30% as training and testing data. Random forest classifier is applied on the labeled stories. The topic of the test stories is predicted. The confusion matrix is generated to show misclassified stories. Accuracy of the classifier is 0.7 and misclassification error is 0.27.

4 Results and Discussions

Table 1 presents different NLP steps carried out initiated from tokenization of sentence to formation of hysynset. There are total of 13,347 sentences in the corpus. The sentence as a sample is shown in column 1. The second column depicts the sample tokens extracted from the sentences. The third column specifies only significant PoS, which automatically results into stop words removal. For e.g., ‘वनेसहिःगज्जती!’ ‘वने’, ‘सहिः’, ‘गज्जती’ ‘!’ are tokens retrieved after tokenization step. ‘!’ being special character, other stop words and punctuations are discarded. Sample unique lemmas obtained after lemmatization are specified in the 4th column. Lemmatization converts ‘वनम्’ into ‘वने’. Hysynset is formed using the relationship between different lemmas. Semantic relativity between the words is identified using Wordnet (<http://www.cfilt.iitb.ac.in/wordnet>). The relationship between ‘मृगम्’ and ‘हरिणः’ is synonym. Thus they are grouped. Similarly, hypernyms-hyponyms are also grouped. It results into dimension reduction and called as hysynset. Each hysynset/token is called a feature of vector.

85,145 tokens. Lemmatization results into. Due to multiple synonyms and hypo-hypernyms final features in the form of Hysynset are 33,121. The total number of tokens are 1,91,234. Stop word removal has resulted into 92,218 tokens. 46,189 unique lemmas are retrieved after lemmatization. Total 7,123 hysynset/features are formed.

TF-IDF measure is computed for every token/feature of hysynset. If a hysynset has two or more words, their TF-IDF weights are summed up and defined as

Table 1 NLP steps on corpus sentences

| Sr. no | Sentence (32,183) | Tokenization | Noise removal | Unique lemmas | Hysynset |
|--------|--|--|--------------------|---|----------------------------------|
| 1 | कदावनेसहिःगर्जति (once lion was roaring in the forest) | कदा,वने, सहिः,, मृगम्,.. (Once, lion, forest, roaring..) | वनम्.. (forest) | मृगम् हरणिक सहिः (deer, lion) | मृगम् सहिः (deer, lion) |

feature weight. Table 2 presents the TF-IDF measure for the extracted tokens. The feature weight of hysynset is reflected in Table 3. In Table 2 ‘कनक’ and ‘अभ्र’ are treated separately and their TF-IDF weights are 0.45 and 0.43 respectively. Table 3 represents the feature weight and features considered by HSVSMS which considers {‘कनक’ and ‘अभ्र’} as hysynset. Its frequency is 0.83. It means, both of the terms are significant with respect to corpus and are chosen to include HSVSMS. These terms are not selected by VSMS due to their low TF-IDF weight than the threshold. Thus being significant, the terms are discarded by VSMS. Also, dimension reduction is observed in HSVSMS due to groups of terms.

To choose significant features threshold value is used. The features having more than 0.75 measure are presented in Table 3. The total significant features are 7123.

Hysynset is used to represent the feature vectors of stories, the matrix of feature vectors is formulated for all stories. The significant features like ‘वृत्त’ etc. are represented in columns. The feature weight of ‘वृत्त’ with respect to story1 is 0.87. In hysynset based vector space, the story1 is represented as {0.67, 0.87... 0.11}. Table 4 represents HSVSMS using significant features.

LDA models are constructed using VSMS and HSVSMS. The coherence score is used in both models for deciding number of topics. Table 5 shows various coherent scores for a different number of topics using HSVMM and VSMS respectively. The number of topics retrieved using both models that is HSVSMS and VSMS is 3. The maximum coherence score observed for using VSMS is 0.6 and using HSVSMS is 0.9.

To represent all the documents present in each topic dendrogram is generated using HAC. One topic is represented by a cluster. Documents belonging to one topic are similar and share the features. Figure 3 shows the topics extracted by HSVSMS using dendrogram for 33 stories. For example, topic 2 includes documents having story ids as 1, 7, 6. The dendrogram is presented having 33 stories. In Table 6 and Fig. 4 improved values of entropy and purity retrieved by using HSVSMS can be clearly seen respectively. The improvement is consistent

Table 2 Term and TF-IDF count used by HSVSMS

| | | | | |
|--------|--------------|-------------|----|--------------|
| Term | मृगम् (deer) | सहिः (lion) | .. | वृत्त (news) |
| TF-IDF | 0.09 | 0.45 | .. | 0.81 |

Table 3 Significant features and their weights selected by HSVSMS

| | | | | |
|---------------|------------------|--------------|---|----------------|
| Term | कनक, अभ्र (gold) | वृत्त (news) | . | भोजनम् (lunch) |
| TF-IDF weight | 0.83 | 0.081 | . | 0.76 |

Table 4 Features of HSVSMS

| | | | | |
|---------|-----------|-------|----|--------|
| Stories | कनक, अभ्र | वृत्त | .. | भोजनम् |
| ST 1 | 0.67 | 0.87 | .. | 0.11 |
| ST 2 | 0.36 | 0.45 | .. | 0.44 |
| | | | .. | |
| ST1127 | 0.11 | 0.11 | .. | 0.76 |

Table 5 Comparison of coherence score

| Number of topics | Coherence score | |
|------------------|-----------------|--------|
| | VSMS | HSVSMS |
| 1 | 0.3 | 0.4 |
| 2 | 0.4 | 0.5 |
| 3 | 0.6 | 0.9 |
| 4 | 0.3 | 0.4 |
| 5 | 0.1 | 0.3 |

throughout the dataset size. Performance evaluation of both models is performed using entropy and purity. To validate the consistency of the performance of the proposed technique, different data sets of different sizes are considered. The betterment of the HSMS is observed consistently for all data sets. It assures more topic independence and more similar topic features are obtained using HSVSMS.

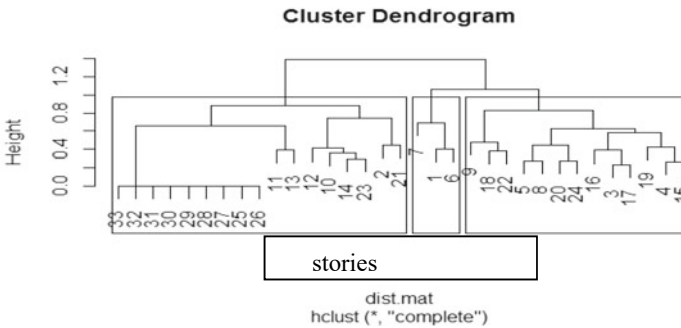


Fig. 3 Dendrogram for 33 stories

Table 6 Entropy comparison for varied datasets

| Dataset size | Entropy | |
|--------------|---------|--------|
| | VSMS | HSVSMS |
| 35 | 0.12 | 0.09 |
| 68 | 0.22 | 0.13 |
| 104 | 0.31 | 0.22 |
| 511 | 0.32 | 0.31 |
| 1127 | 0.41 | 0.39 |

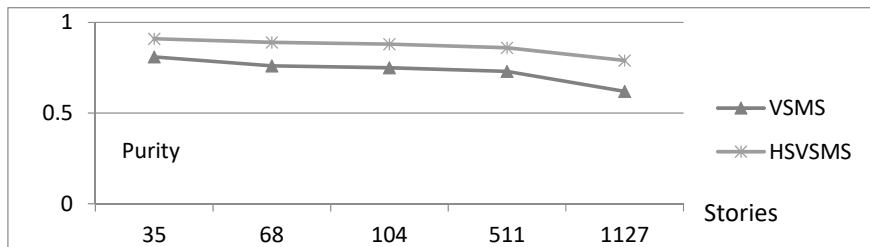
**Fig. 4** Purity trend by both models for different datasets

Table 7 shows the topic, corresponding words and their using VSMS and HSVSMS. The variation between words extracted using VSMS and HSVSMS can be easily understood by making use of pragmatic knowledge. For example, the hysynsets of topic 1, extracted using HSVSMS show the values of living as {धर्म: कर्मअहिसा that is non violence}, considering other tokens present in topic 1, it can be identified as a group of ‘Jatak Katha’. Similarly for topic 3 words detected by HSVSMS shows that the group of moral stories considering animal characters whereas in VSMS no such words are detected to decide the domain. Due to informative and relevant features generated by HSVSMS, the topics can be labeled as ‘Jatakkatha’, ‘VikramVetal’, ‘Moral stories’ and topics for VSMS can be labeled as ‘Animal stories’, ‘VikramVetal stories’, ‘Kids stories’.

β constant of the features retrieved by both models is presented in Table 7. It is clear that with proposed model membership values of features are higher than

Table 7 Comparative analysis of proposed and existing techniques

| Topics | | | |
|--------------|------------------|----------------|------------------|
| VSMS | β constant | HSVSMS | β Constant |
| अरण्यम् | 0.8 | अहिसा | 0.8 |
| सहिः | 0.6 | धर्म: कर्म | 0.76 |
| सरोवर (lake) | 0.5 | भोजनम् | 0.73 |
| भोजनम् | 0.4 | शषिय (student) | 0.64 |
| मृगम् | 0.3 | वपिर (money) | 0.56 |

| Reference Prediction | Jatak Moral | Stories | Vikram Vetal |
|-------------------------|-------------|---------|--------------|
| Jatak | 60 | 20 | 22 |
| Moral Stories | 5 | 109 | 7 |
| Vikram Vetal | 19 | 24 | 74 |

| Overall Statistics | |
|------------------------|--------------------|
| Accuracy | : 0.7229 |
| 95% CI | : (0.6714, 0.7704) |
| No Information Rate | : 0.4548 |
| P-value [Acc > NIR] | : < 2e-16 |
| Kappa | : 0.5809 |
| McNemar's Test P-value | : 7.6e-05 |

| Statistics by Class: | | | | |
|----------------------|--------------|--------------|----------------|---------------------|
| | Class: Jatak | Class: Moral | Class: Stories | Class: Vikram Vetal |
| Sensitivity | 0.7037 | 0.7219 | 0.7400 | 0.7400 |
| Specificity | 0.8327 | 0.9503 | 0.8233 | 0.8233 |
| Pos Pred Value | 0.5758 | 0.9237 | 0.6435 | 0.6435 |
| Neg Pred Value | 0.8970 | 0.8037 | 0.8802 | 0.8802 |
| Prevalence | 0.2440 | 0.4548 | 0.3012 | 0.3012 |
| Detection Rate | 0.1717 | 0.3283 | 0.2229 | 0.2229 |

Fig. 5 Confusion matrix

existing. That is ‘मृतम्’ and ‘वपिर्’ which are the top 5 th tokens of VSMS and HSVSMS has constant as 0.3, and 0.56 respectively. Similar results are observed for all topics.

To predict the topic of the story, HSVSMS is used as it’s proved to be accurate than VSMS. Once the documents are labeled with topics, classification and prediction is applied. A classifier is built using train data which is considered as 70% of total data. The remaining 30% data is used as test data. Random forest algorithm is used which produced 500 trees to classify the stories and to predict the topic of the story. Figure 5 shows misclassified and correctly classified observations in the form of matrix. for 338 stories which is 30% of 1125 (total stories) and the highest values present in the diagonal of matrix assures the true predictions. The accuracy of the algorithm is 0.72 and the error value is less than 0.28. The minimum error is reached till 0.1. Figure 6 represents generated trees for a corpus. Error is shown on Y axis which gets reduced as number of trees are increased. Stories belonging to ‘Moral’ domain are being predicted more correctly and error is also been consistently low after 350 trees.

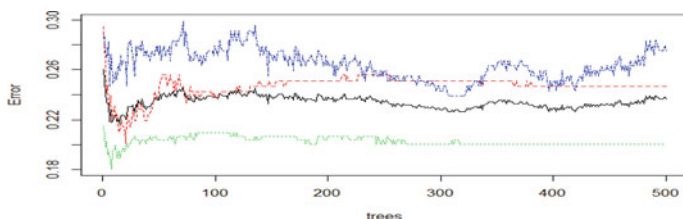


Fig. 6 Trees generated by random forest algorithm

5 Conclusions

The proposed 4-phased approach not only identifies the topics of the Sanskrit corpus but also evaluates, visualizes and identifies the topic of newly arrived story. The performance evaluation is performed at each phase that is for topic building, clustering and classification using coherence score, entropy, purity. All these evaluation parameters are better in case of proposed technique. The use of hysynset while constructing VSMS converts it into HSVSMS and avoids dimension curse their by improving accuracy. LDA is built using HSVSMS. The topics retrieved by the proposed approach are more specific due to more relevant topic features. Topics are represented using dendrogram and dendrogram is evaluated using entropy and purity. lower entropy and higherpurity proves betterment of HSVSMS than VSMS. HSVSMS being better is used for predicting topic of new story which can be 'Jatakkatha', 'VikramVetal' or 'Moral stories'. Random forest is used and evaluated using classification error and accuracy. In the absence of availability of standard experiments, current work could not be compared with other existing work in case of prediction of stories.

References

1. Al-Sultany, G. A., & Aleqabie, H. J. (2019). Enriching Tweets for Topic Modeling via Linking to the Wikipedia. *International Journal of Engineering & Technology*, 8(1.5), 144–150.
2. Syed, S., & Spruit, M. (2017, October). Full-text or abstract? Examining topic coherence scores using latent dirichlet allocation. In *2017 IEEE International conference on data science and advanced analytics (DSAA)* (pp. 165–174). IEEE.
3. Raulji, J. K., & Saini, J. R. (2016). Stop-word removal algorithm and its implementation for Sanskrit language. *International Journal of Computer Applications*, 150(2), 15-17.
4. Bafna P.B., Saini J.R., 2020, "Marathi Text Analysis using Unsupervised Learning and Word Cloud", *International Journal of Engineering and Advanced Technology*, 9(3)
5. Xu, Songhua. "Recommending personally interested contents by text mining, filtering, and interfaces." U.S. Patent 9,171,068, issued October 27, 2015.
6. Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 1(1), 17-35.
7. Bhardwaj, A., Setlur, S., & Govindaraju, V. (2007). Keyword spotting techniques for sanskrit documents. In *Sanskrit Computational Linguistics* (pp. 403–416). Springer, Berlin, Heidelberg.
8. Koltcov, S., Ignatenko, V., & Koltsova, O. (2019). Estimating Topic Modeling Performance with Sharma–Mittal Entropy. *Entropy*, 21(7), 660.
9. Neill, T. (2019). LDA Topic Modeling for pramāṇa Texts: A Case Study in Sanskrit NLP Corpus Building. In *Proceedings of the 6th International Sanskrit Computational Linguistics Symposium* (pp. 52–67).
10. Saini, Jatinderkumar & B., Prafulla. (2020). Measuring the Similarity between the Sanskrit Documents using the Context of the Corpus. *International Journal of Advanced Computer Science and Applications*. 11. <https://doi.org/10.14569/IJACSA.2020.0110521>.
11. Sharef, N. M., Zin, H. M., & Nadali, S. (2016). Overview and Future Opportunities of Sentiment Analysis Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data. *JCS*, 12(3), 153-168.

12. Wang, P., Zhang, H., Xu, B., Liu, C., & Hao, H. (2014). Short text feature enrichment using link analysis on topic-keyword graph. In *Natural Language Processing and Chinese Computing* (pp. 79–90). Springer, Berlin, Heidelberg.
13. Petersen, C., Lioma, C., Simonsen, J. G., & Larsen, B. (2015, September). Entropy and graph based modelling of document coherence using discourse entities: An application to IR. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval* (pp. 191–200).
14. Chandra, R., & Kulkarni, V. (2022). Semantic and sentiment analysis of selected Bhagavad Gita translations using BERT-based language framework. arXiv preprint [arXiv:2201.03115](https://arxiv.org/abs/2201.03115).
15. Egger, R., & Gokce, E. (2022). Natural language processing (NLP): An introduction. In *Applied Data Science in Tourism* (pp. 307–334). Springer, Cham.
16. Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. *Automation in Construction*, 134, 104059.
17. Kaddoura, S., & D. Ahmed, R. (2022). A comprehensive review on Arabic word sense disambiguation for natural language processing applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, e1447.
18. Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to deep learning in natural language processing: models, techniques, and tools. *Neurocomputing*, 470, 443–456.
19. Punitha, S. C., P. Ranjith Jeba Thangaiah, and M. Punithavalli. “Performance analysis of clustering using partitioning and hierarchical clustering techniques.” *International Journal of Database Theory and Application* 7, no. 6 (2014): 233–240.
20. Vispute, S. R., & Potey, M. A. (2013, September). Automatic text categorization of Marathi documents using clustering technique. In *2013 15th International Conference on Advanced Computing Technologies (ICACT)* (pp. 1–5). IEEE.
21. Xu, B., Guo, X., Ye, Y., & Cheng, J. (2012). An Improved Random Forest Classifier for Text Categorization. *JCP*, 7(12), 2913–2920.
22. Moon, Seonghyeon, Gitaek Lee, and Seokho Chi. “Automated system for construction specification review using natural language processing.” *Advanced Engineering Informatics* 51 (2022): 101495.
23. Nair, J., Nair, S. S., & Abhishek, U. (2022). Sanskrit Stemmer Design: A Literature Perspective. In *International Conference on Innovative Computing and Communications* (pp. 117–128). Springer, Singapore.
24. Pontillo, T., & Candotti, M. P. (2022). Dispensing with ellipsis devices in the analysis of Sanskrit ba-huvrīhi. *Journal of South Asian Linguistics*, 12(1), 1–22.
25. Yang, N. Y., Kim, S. G., & Kang, J. Y. (2018). Researcher and research area recommendation system for promoting convergence research using text mining and messenger UI. *The Journal of Information Systems*, 27(4), 71–96.

Filling the Knowledge Base of the Filtering System Algorithm Development



Elshod Khaydarov

Abstract In this article, the development of an algorithm for filling the knowledge base of the filtering system was considered, in particular, a mathematical model for the formation of a knowledge base, a hierarchy for the formation of a knowledge base of a multi-level spam detection system, an algorithm for filling the knowledge base and a flowchart. Statistical methods and data that were previously taken from studies were used for the study.

Keywords Knowledge base · Filtering · Spam · e-mail · System · Message

1 Introduction

It is necessary to constantly add new types of knowledge to the database used to detect spam messages and to integrate this database into the international database of spam messages. The integration of the existing knowledge base in the organization into the international spam database and the IP address database that sends spam messages is based on the internal regulations of each organization. Therefore, the issue of integration may be different for different organizations. However, in order to constantly add new types of knowledge to the developed knowledge base, it is necessary to form the knowledge base in accordance with the information system in the organization. Because the spam message filtering system in an organization is hierarchical, the knowledge base is also organized hierarchically. The following issue should be taken into account. In addition to spam message databases, it is necessary to create a knowledge base that contains information about the spam message in order to identify spam messages that do not exist in the database of spam messages on the organization's email servers. There are many ways to do this, but the most widely used among them is the simple intersection method.

E. Khaydarov (✉)
Amir Temur, 108, Tashkent, Uzbekistan
e-mail: elshodhaydarov1881@gmail.com

2 Main Part

A simple intersection uses messages identified by all users of the mail system, such as spam messages, to form a knowledge base. In this case, the error rate of the first category increases, while the error rate of the second category decreases. In addition to spam message databases, it is necessary to create a knowledge base that includes knowledge about the spam message in order to identify spam messages in a new view.

$$B = b_1 \cap b_2 \cap \dots \cap b_n = \bigcap_{k=1}^n b_k \quad (1)$$

Here:—organizational knowledge base;—postal system—user knowledge base;—The number of users of the e-mail system in the organization. B, b_k, n

In a complex intersection, the decision to add messages identified by a particular user, such as spam messages, to the database is made by the mail system user, which in turn creates a sequence of how many current users mark the message correctly and how many incorrectly. In many cases, the assessment of each user's authority is abstract. Recipients who receive a single message independently of the system compare it to spam messages. In this case, the error rate of the first category increases, while the error rate of the second category decreases. That is why this method is not widely used [1].

This way of forming a knowledge base is done in a simplified form as follows.

$$\begin{aligned} B &= (b_1 \cap b_2) \cup (b_1 \cap b_3) \cup \dots \cup (b_1 \cap b_n) \cup (b_2 \cap b_3) \cup \dots \cup (b_2 \cap b_n) \cup \dots \cup (b_{n-1} \cap b_n) \\ &== \bigcup_{\substack{k=n, j=n, k \neq j \\ k=1, j=1}} (b_k \cap b_j) \end{aligned} \quad (2)$$

In merging, the email system uses messages detected by at least one user, such as spam messages, to form a knowledge base. In this case, the error rate of the first category decreases and the error rate of the second category increases. This method is also rarely used [2].

This way of forming a knowledge base can be expressed as follows:

$$B = b_1 \cup b_2 \cup \dots \cup b_n = \bigcup_{k=1}^n b_k \quad (3)$$

While the filters used to detect spam messages are common to all users, they can be installed on a single server or based on the user's personal filters provided to users to account for a specific user area of interest and perform accurate filtering of incoming messages. But one of the disadvantages of users' personal spam filters is that they can't be used from other users' experiences. That is, when receiving spam messages that have passed through a personal filter, each user must then point an

error to the filter itself so that the filter can detect and capture them when receiving spam messages of this category.

A voluntary organization that typically supports an anti-spam system is divided into three levels:

- end user level;
- department or subdivision level;
- overall organizational level.

At the lower level, each user separates all the received ones into useful messages and spam messages. In parallel with this process, it can also display folders where documents that are useful to the user are stored in order to fill the knowledge base with knowledge that is high priority for itself. The system then clearly distinguishes between words specific to spam messages and words used in useful messages. This is done based on the method of detecting spam messages based on keywords. The process of forming a database system to combat spam messages is shown in Fig. 1.

The keywords are then assigned to the department level where all the lists come together, as the professional interests of all employees in a single department are

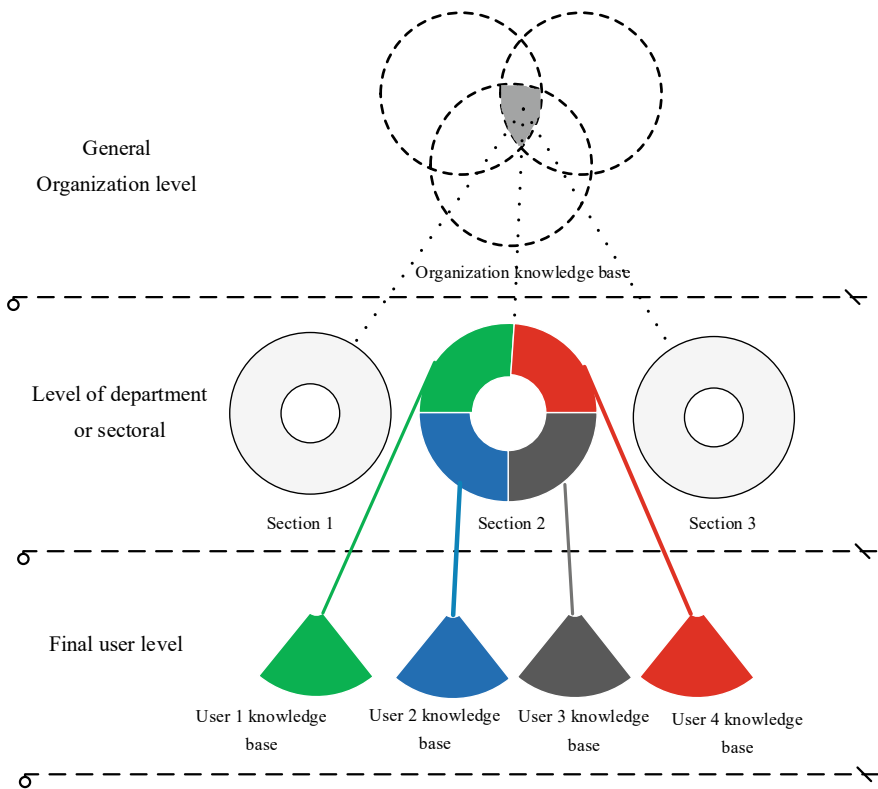


Fig. 1 Spam message detection is a hierarchy of multi-level system knowledge base formation

almost matched. In the final stage of knowledge base formation, the system separates keywords separated by all sections and removes words not separated by at least one section from the knowledge base. This is due to the possibility that different departments of the same organization have different areas of interest [3].

In such a scheme, high efficiency of spam filters can be achieved by increasing the size of the training selection (creating a knowledge base at the department level) while maintaining a low level of error by removing randomly dropped keywords (at the organization level) from the training selection [4].

This way of forming KB can be expressed as follows [5]:

$$B_1 = b_{11} \cup b_{12} \cup \dots \cup b_{1n} = \bigcup_{k=1}^n b_{1k} \quad (4)$$

$$B_B = B_1 \cap B_2 \cap \dots \cap B_p = \bigcap_{k=1}^p B_k \quad (5)$$

$$\begin{aligned} B_B &= (b_{11} \cup b_{12} \cup \dots \cup b_{1n}) \cap (b_{21} \cup b_{22} \cup \dots \cup b_{2m}) \cap \dots \cap (b_{p1} \cup b_{p2} \cup \dots \cup b_{po}) \\ &= \bigcap_{k=1}^p \left(\bigcup_{j=1, j \neq k}^r b_{k,j} \right) \end{aligned} \quad (6)$$

Here:

p -number of departments in the organization;

n -The number of employees in the first department of the organization;

m -the number of employees in the second department of the organization;

o -The number of employees in the department of the organization;

b_{po} -organization p -knowledge base formed by the employee of the department;

B_k -knowledge base of the organization-department;

B_B -The whole organizational knowledge base.

In particular, each unsolicited e-mail is intended for one user only:

$$X - E_K = (X_1 - E_K) + (X_2 - E_K) + \dots + (X_n - E_K) = \sum_{k=1}^n (X_k - E_K) \quad (7)$$

$$\begin{aligned} X_k - E_K &= (X_{k,1} - E_K - F_k) + (X_{k,2} - E_K - F_k) + \dots + (X_{k,m} - E_K - F_k) \\ &= \sum_{j=1}^m (X_{k,j} - E_K - F_k) \end{aligned} \quad (8)$$

From (7) and (8) we obtain:

$$X - E_K = \sum_{k=1}^n \left(\sum_{j=1}^m (X_{k,j} - E_K - F_k) \right) \quad (9)$$

$$A = A_1 + A_2 + \dots + A_n = \sum_{k=1}^n A_k \quad (10)$$

$$A_k = A_{k,1} + A_{k,2} + \dots + A_{k,m} = \sum_{j=1}^m A_{k,j} \quad (11)$$

From (10) and (11) we obtain:

$$A = \sum_{k=1}^n \left(\sum_{j=1}^m A_{k,j} \right) \quad (12)$$

$$X - E_K = (X_1 - E_K) \cup (X_2 - E_K) \cup \dots \cup (X_n - E_K) = \bigcup_{k=1}^n (X_k - E_K) \quad (13)$$

$$\begin{aligned} X_k - E_K &= (X_{k,1} - E_K - F_k) \cup (X_{k,2} - E_K - F_k) \cup \dots \cup (X_{k,m} - E_K - F_k) \\ &= \bigcup_{j=1}^m (X_{k,j} - E_K - F_k) \end{aligned} \quad (14)$$

From (1) and (2) we obtain:

$$X - E_K = \bigcup_{k=1}^n \left(\bigcup_{j=1}^m (X_{k,j} - E_K - F_k) \right) \quad (15)$$

$$A = A_1 \cap A_2 \cap \dots \cap A_m = \bigcap_{k=1}^n A_k \quad (16)$$

$$A_k = A_{k,1} \cup A_{k,2} \cup \dots \cup A_{k,m} = \bigcup_{j=1}^m A_{k,j} \quad (17)$$

$$A = \bigcap_{k=1}^n \left(\bigcup_{j=1}^m A_{k,j} \right) \quad (18)$$

Here:

X -Stream of incoming e-mails to individual users or to all users;

S -mail server;

E_K -server classifier;

F_n -Section classifier set and trained by users of the department;

n -Stream of incoming e-mails filtered by the server classifier and intended for users of the department; $X_n - E_K F_k n$

$X_{n,m} - E_K - F_k$ -first by the server classifier E_K

F_n -incoming stream of incoming e-messages filtered by the section filter and intended for the section -user; nnm

A_{nm} -section is intended for teaching classifier and -class-message by -user of -section; $F_n n n m$

$A_n E_{Kn}$ -a message classified by all users of the section designed to teach the server classifier.

The following steps are performed to complete this formed knowledge base. First of all, special attention is paid to filling the knowledge base of individual users, i.e., the content of messages that are useful for the user for whom the knowledge base is formed. The algorithm for filling the database of individual user knowledge is shown in Fig. 2.

At the beginning of the work, the user must specify a folder containing text and tables of documents that the system can use to populate the database of useful messages. The system then begins to form a knowledge base. Each document is reviewed separately to save the user's computer computing resources.

The system in which a single document index is built then moves on to building the next document index, and so on until the entire document index is built. In the process of constructing a document index, the system must calculate the frequency of occurrence of words separated within the document being analyzed and, in the current case, separate the lexical unit of words. When separating individual words from the document being analyzed, the system removes suffixes to generalize words that are used in different ways [3, 5].

It is possible to create a dictionary showing words with all their roots and suffixes in each word, but since it takes a long time to search for them in the database, the

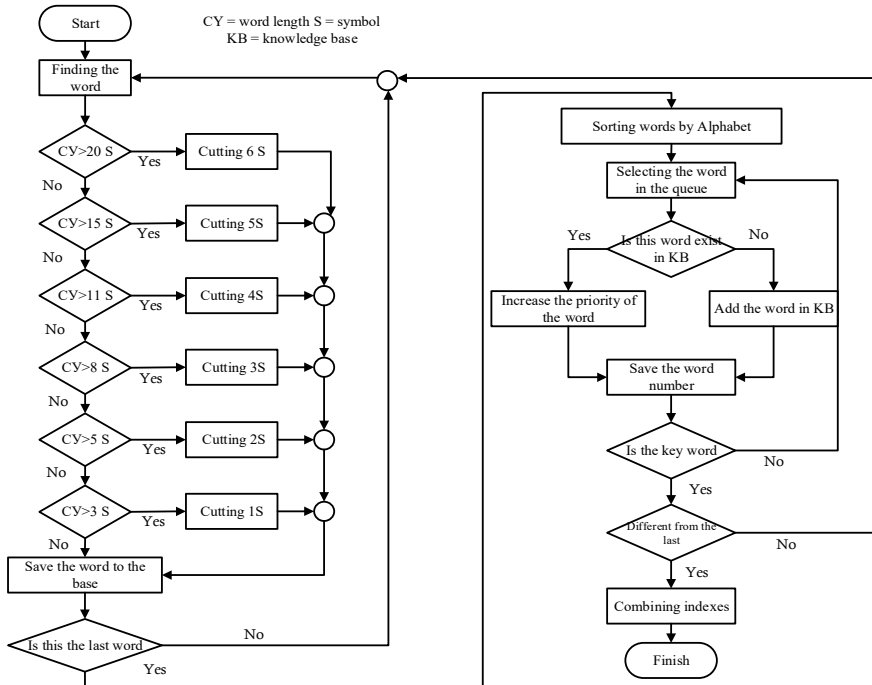


Fig. 2 Block diagram of the algorithm for filling the knowledge base

system estimates the additional length based on the length of the words. The main goal is to find the core of the keyword. The longer the word, the more characters are removed from the end.

If it is placed directly in the knowledge base without removing the suffixes in the words, the priority of the knowledge in the knowledge base decreases. Because it is clear that the meaning of a word changes with the help of an appendix, but the first type of error occurs as a result of an increase in the number of words consisting of the same stem.

Therefore, keywords are mainly formed from words with a single core without the addition of suffixes, and keywords are included in the knowledge base.

3 Conclusion

In order to calculate the frequency of occurrence of words within a document, ie all documents, all words are sorted in alphabetical order, and then the cases of word repetition are calculated. When keywords occur in repetition cases, the priority of these keywords increases. By identifying the existing content (documents) in the organization and the keywords that can be found in these content, it will be possible to filter the documents that come to the e-mail messages throughout the organization after adding them to the knowledge base.

If the keywords that occur among the existing documents in the organization are not combined into the same indexes, there will be an increase in the first and second types of errors in the process of filtering documents. This in turn affects the efficiency and reliability of the email message filtering system.

References

1. Bekmuratov T.F., Botirov F.B., Neural Network Method and Algorithm for Document Detection Based on Signaling Analysis, 11th World Conference “Intelligent System for Industrial Automation” (WCIS-2020), Tashkent, pp. 26–32.
2. Botirov F.B., Gafurov S.R., Gafurov A.A., Identification of key persons in the information security incident management process and distribution of roles between them, Chemical technology, control and management, 2021, 3(99), pp. 72–77.
3. Zaynutdinova, M., Sayfullaev, S. Analysis and Modeling of Information Security Systems in Industry 4.0. Advances in Intelligent Systems and Computing, 2021, 1323 AISC, с.р. 10–17.
4. Abdukhalil, G., Sherzod, S. Principles of adaptation in protecting corporate systems 2020 International Conference on Information Science and Communications Technologies, ICISCT 2020, 2020, 9351461
5. Schafer, J. B., Konstan, J., and Riedl, J. (1999). Recommender Systems in E-Commerce. In Proceedings of ACM E-Commerce 1999 conference.
6. Sungeetha, Akey, and Rajesh Sharma. “Transcapsule model for sentiment classification.” Journal of Artificial Intelligence 2, 03 (2020): 163-169.

7. Kottursamy, Kottilingam. "A review on finding efficient approach to detect customer emotion analysis using deep learning analysis." *Journal of Trends in Computer Science and Smart Technology* 3, no. 2 (2021): 95-113.
8. Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl, J. (2000). Application of Dimensionality Reduction in Recommender System—A Case Study. In *ACM WebKDD 2000 Workshop*.
9. Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl, J. (2000). Analysis of Recommendation Algorithms for E-Commerce. In *Proceedings of the ACM EC'00 Conference*. Minneapolis, MN. pp. 158–167

Development of a Linear Scale Consensus Method of the Blockchain System



Gennady Shvachych, Boris Moroz, Maksym Khyliko, Andrii Matviichuk, Vladyslava Kozenkova, and Volodymyr Busygin

Abstract The research highlights the development of a new and reliable blockchain system consensus method focused on applying a linearly scalable consensus mechanism. The new system is based on an analysis of available consensus mechanisms, sharding, and distributed randomness generation. The proposed approach allows the development of a blockchain with the following advantages: full scalability, security, energy efficiency, and fast consensus. Compared to known methods, the proposed shard option performs network connection and transaction verification and shows the current state of the blockchain. The decision threshold is low enough to allow small validators to participate in the network and receive appropriate rewards. The sharding process runs safely by applying a distributed randomness generation process that is unpredictable unbiased. In order to prevent adaptive Byzantine malicious validators, a continuous network reboot process is provided. The developed approach is resistant to fluctuations in the number of validators. It does not set a lower limit against the number of validators in every fragment, as in other solutions such as *Zilliqa*. Instead, an adaptive Proof-of-Stake-based model was adopted so that attackers can never occupy more than one-third of the shares that vote in a single shard that makes it reliable. The methods for creating the proposed blockchain improve available mechanisms for the blockchain functioning and have practical value for application in various digital economy sectors.

G. Shvachych
Ukrainian State University of Science and Technology, Dnipro, Ukraine

B. Moroz
Dnipro University of Technology, Dnipro, Ukraine

M. Khyliko
National Academy of Science of Ukraine, Kyiv, Ukraine

A. Matviichuk
Vernadsky National Library of Ukraine, Kyiv, Ukraine

V. Kozenkova
State Agrarian and Economic University, Dnipro, Ukraine

V. Busygin (✉)
VUZF University (Higher School of Insurance and Finance), Sofia, Bulgaria
e-mail: busygin2009@gmail.com

Keywords Blockchain · Transactions · Stability · Security · Consensus mechanism · Reliability · Validation · Shard · Algorithm · Protocol · Distributed randomness

1 Introduction

Problem Statement. Blockchain technology is still under development. Therefore, the natural outcome is that the “correct” consensus protocol issue is still under discussion and debate. Many critical considerations, such as the degree of decentralization, underlie the spirit of blockchain as a technology. At least right now, there is no consensus on the “correct” consensus algorithm, which means that competition will only intensify in the future.

This paper proposes and explores a new blockchain system, operating on a linearly scalable consensus mechanism, a selection method, confirming a shard by voting shares, and scalable randomness generation using *VRF (Verifiable Random Function)* and *VDF (Verifiable Delay Function)* delays. The system is based on analyzing available consensus mechanisms, sharding, and distributed randomness generation. It is fully scalable, secure, energy-efficient, and with fast consensus.

Compared with available methods [1–4], the improved shard method performs network connection and transaction verification and exposes the state to the blockchain. The threshold is low enough for small validators to participate in the network and get rewards. The proposed sharding process occurs safely through a distributed randomness (*DRG*) process that is unpredictable, unbiased, and proven. The network is overloaded continuously to prevent slow adaptive byzantine malicious validators. Unlike other sharded blockchains requiring *Proof of Work (PoW)* model to select validators, the proposed consensus is based on *Proof of Stake (PoS)* and is therefore energy-efficient.

Proof of Stake (PoS) is a consensus model introduced in 2011 as an alternative to the *Proof of Work (PoW)* model. It aims to overcome the scalability limitations of *PoW* networks. Note that the goal of *PoW* and *PoS* is the same—to achieve consensus in the blockchain, however, the *PoS* model implements a different way of determining participants who verify transaction blocks. There are no miners on *PoS* blockchains. The participant priority according to the rules of the *PoS* algorithm does not depend on its computing power, but on the amount of cryptocurrency the participant possesses. Thus, the *PoS* model is devoid of the main disadvantages of the *Proof of Work* model:

- Mining consumes a huge amount of electricity power and hardware wears out quickly.
- The *Proof of Work* model provides a sufficient security level only if there is a large group of miners competing for block rewards. If the network is small a hacker can get a simple majority of the processing power and reorganize the blocks as he/she sees fit.

At the same time, validators, the cryptocurrency owners, support the *PoS* blockchain model functionality. They check user transactions, and if at least 2/3 of the validators agree that the transaction is correct, it is included in a new blockchain block. Simultaneously, in order to be eligible for block verification, participants need to block several coins in a specific blockchain smart contract. That process is known as staking. After that, the *PoS* protocol can choose a participant to validate the next block. Depending on the network, the choice may be random or according to the amount of cryptocurrency staked. As a reward, the selected validator receives a transaction fee from the verified block. As a rule, the more coins one blocks, the higher the chance of being selected.

Obviously, here the validators only do useful work (validation), and not numbers enumeration, so they do not have a race for performance, like miners. However, for the network to work efficiently and quickly, validators must run software on very powerful hardware, with a constant 24/7 network connection and a wide Internet channel.

Note that the main advantage of the *PoS* model is speed. Many *PoW* blockchains (such as *Bitcoin*) will never be able to process transactions as fast as *PoS* blockchains. And speed is a key factor for a network.

As a rule, the following arguments are given in *PoS* favor:

- significant resources to carry out an attack, which makes it impractical from a financial point of view;
- if the attacker has a large number of tokens, he/she will suffer from the attack, as that will violate the cryptocurrency stability.

Let us note the arguments that cause concern:

- *PoS* provides additional motivation for the resources accumulation in one hand, which may adversely affect the network decentralization;
- if a small group is formed that collects large enough resources, it can impose own rules for the network on other participants.

Thus, the *PoS* mechanism aims to improve blockchain technology and gain significant adoption in the industry. The lower barrier to entry into the mining process, no need for infrastructure, and relatively less vulnerability to attacks are currently attracting the community to have the *PoS*-based application and coins.

In the research in question, the consensus is achieved by a linearly scalable *Byzantine Fault Tolerance (BFT)* algorithm, faster than *Practical Byzantine Fault Tolerance (PBFT)* [5].

Analysis of Recent Research and Publications. The consensus protocol is a major element of a blockchain. It discovers how reliably and quickly the blockchain is verified to reach a consensus on the next block. Achieving consensus in a distributed environment comes down to solving the problem of Byzantine generals. The problem formulation is as follows. The Byzantine army includes n legions, each of which is commanded by its general, and the army also has a commander-in-chief, to whom the generals of the legions are subordinate. The army surrounds the city intending to

attack. The favorable outcome of the war depends on the actions of all the generals, who need to communicate to come to a unified agreement on whether attack the city or not. However, there may be traitors among the generals, including the in-chief commander. A traitor can send orders of different content to different generals [6, 7]. Reaching consensus in such an environment is a quite difficult task. The problem's solution served as the basis for a mathematical model for the construction and development of blockchain technology. There is no central authority to provide for the same work on remote nodes in blockchain networks. Thus, there is a need for consensus protocols among distributed nodes. The crypto industry is not limited to applying well-known consensus mechanisms such as *PoW* or *PoS*. Researchers increasingly demonstrate new consensus protocols; in most cases, some modifications to the main available protocols. Let us consider some alternative mechanisms for achieving consensus in distributed environments.

The *BFT* algorithms are distributed network algorithms that allow consensus to be achieved even if some network nodes do not respond or give incorrect information. The *BFT* mechanism aims to protect against system failures by collective decision-making (both healthy and defective nodes) and reduce the impact of failed nodes.

The *Verifiable Byzantine Fault Tolerance (VBFT)* algorithm is a new consensus algorithm that combines traditional *PoS*, *Verifiable Random Function (VRF)*, and *BFT*. This mechanism was developed specifically for the needs of the *ONTology* platform. *VBFT* can hold the consensus groups' scalability by *VRF*; it provides the randomness and fairness of the generation of the consensus set and ensures that the final state is reached quickly. In this algorithm, the centralization risk problem is eliminated. Conforming to the *ONTology*'s plan, the *VBFT* consensus algorithm could hold up to 2401 nodes onward while reaching consensus in <10 s [7]. The main advantage of such a protocol is the absence of the risk of centralization scalability, and hence it is resistant to attacks. The algorithm's disadvantage is that its application is limited by the *ONchain* company and the *ONTology* project.

A new class of consensus protocols was proposed to increase performance, the Byzantine Fault Tolerance (*PBFT*), wherein anonymity is not an essential aspect, i.e., nodes know some information about each other (initially nodes are authenticated). That allows consensus algorithms to be optimized with far higher throughput. The speed can increase tenfold, and data processing from hundreds to thousands of transactions per second, which is great for corporate realities.

This algorithm is used in the Hyperledger Fabric project as a consensus protocol, as it can process up to one-third of illegitimate transactions [8]. The protocol operation principle is as follows: the validator receives a message at the input, according to which the validator needs to decide whether to consider it true or not. To do this, the validator performs internal checks, followed by polling the other nodes one by one whether the transaction is valid. If two-thirds of the participants vote for this transaction as correct, the validator accepts it and transfers its decision to the blockchain network to other validators. It should also be noted that there is no hashing procedure in the *PBFT* algorithm [9].

Research Purpose. Based on the literature review and the analysis results of the current state of blockchain technologies development develop a linearly scalable blockchain consensus method. The new approach should be fully scalable, evidence-based secure, and energy-efficient blockchain; explore the functionality and features of the blockchain system based on next-generation sharding, which solves many problems of available blockchains; explore the stability and reliability of the developed blockchain system consensus method.

2 Statement of the Main Research Material

2.1 *The Idea and Essence of the Blockchain System New Consensus Method*

As the *PBFT* protocol improvement, the paper proposes a linearly scalable consensus mechanism for communication complexity. Instead of asking all validators to post their votes, the leader instigates the process of multi-signature signing to gather the validators' votes into $O(1)$ —a multi-signature, and then relay it. Therefore, instead of receiving $O(N)$ signatures, every validator gets just one multi-signature, thereby reducing the communicating complexity with $O(N)^2$ to $O(N)$.

The idea of multi-signature $O(1)$ multi-signature improves the *BFT* method from the *ByzCoin* [10] blockchain, using the Schnorr signature scheme to aggregate constant size multi-valued signals and form a multicast tree between validators to further message transmission. Nevertheless, a multi-valued Schnorr signature requires a secret round of commitments, resulting in a total of two round-trip requests for a single multi-signature.

The paper proposes the method to improve available one by *BLS* (*Boneh-Lynn-Shacham*) multi-signature [11]; it requires only one round-trip request. Therefore, the method is designed to be at least 50% faster than the *BFT ByzCoin* method. Figure 1 presents the developed network communication method during one consensus round.

The developed method for conducting a consensus procedure involves the next steps:

1. The leader creates a brand new block and passes the header of the block to all validators. Concurrently, the leader relays the block contents with erasure encoding (Fig. 1—the Announce phase).
2. The validators validate the header of the block, signs it with the *BLS* signature, and send the signature back to the leader (Fig. 1—Prepare phase).
3. The leader waits for at least $2f + 1$ valid signatures from validators (including leader's ones) and combines them into a *BLS* multi-signature. The leader then relays the aggregated multi-signature with the bitmap with the changes signed by the validators. Along with step 3, the *PBFT* "preparation" phase is completed.

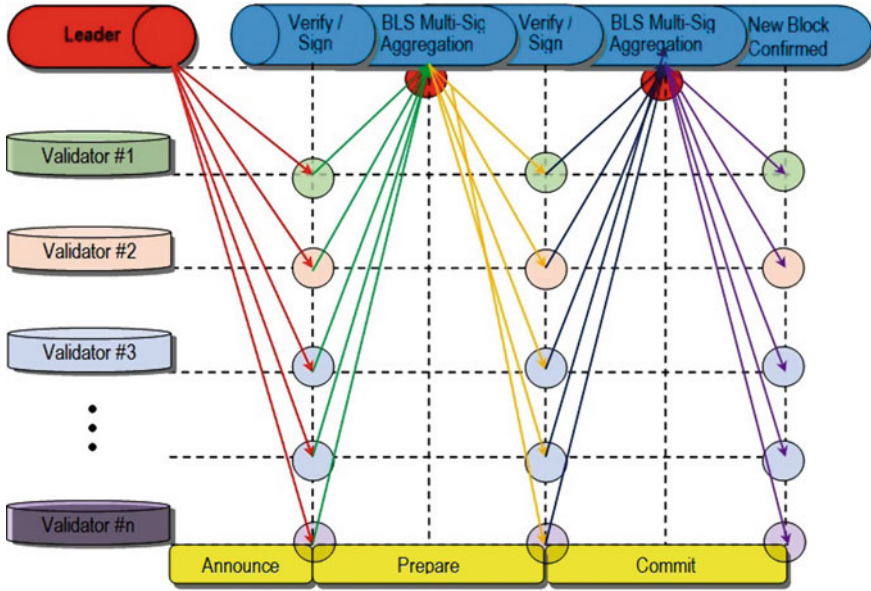


Fig. 1 Network communication of the developed method during one round of consensus (Source Authors' elaboration)

- Validators check if it has at least multi-signature $2f + 1$ signers, verifies transactions with block content relays from the leader in step 1, signs the received message from step 3, and returns it to the leader.
- The leader then awaits at least $2f + 1$ valid signatures and, starting at step 4, unites them into a *BLS* multi-signature creating a bitmap that registers everyone who signed. The leader then does a new block with all signed, multi-signed, and bitmaps and passes the new block to all validators. Along with step 5, the “fixation” phase of the *PBFT* ends (Fig. 1—the Commit phase).

Consensus validators are elected on a Proof-of-Stake basis. The proposed protocol differs from available *PBFT* wherein a validator with more voting shares has more votes instead of having one signature—one vote. Contrary to waiting for $2f + 1$ signatures from validators, the leader awaits signatures from validators who collectively possess minimum $2f + 1$ voting shares.

Note that the common procedure to download the history of blockchain and reconstruct present state is extremely slow to allow re-introducing changes (it takes several days for the *Ethereum* blockchain to synchronize the history fully), given that the current state is much smaller than the entire history of the blockchain. To optimize the state synchronization process, making the blockchain state as small as possible is proposed. Loading the current state by epoch period is possible compared to loading the entire history.

In *Ethereum*, many accounts are empty and waste state-space on the blockchain. Empty accounts cannot be deleted due to possible replay errors when old transactions

are resubmitted to a deleted account [12]. The problem can be solved by preventing replay attacks by allowing transactions to specify the current block hash: a transaction is valid only up to a certain number (e.g., 300) of blocks right after the block of the specified hash. Thus, old accounts can be deleted securely, which will greatly speed up the verification of the blockchain state.

New validators that join a shard initially download present state of that shard to enable a quick start validating transactions. The new node must make an appropriate check to ensure that the currently loaded state is valid. Instead of downloading the entire blockchain history to verify present state, the new node downloads the headers of the history block. It verifies them by affirming their signatures as cryptographic evidence (e.g. signatures and hash pointers) from present state to the genesis block. Signature verification is not computationally expensive, and it takes a significant time to verify all signatures onward from the genesis block. When it comes to mitigating this, the first block of each epoch is proposed to include an additional hash pointer to the first block of the last epoch. Thus, a new node can move over blocks during an epoch when it tracks hash pointers to the genesis block, which could greatly speed up checking the current state of the blockchain.

2.2 *Stability and Reliability Research of the Developed Blockchain System Consensus Method*

The proposed consensus method uses a different approach than the previously discussed one, with *Proof-of-Stake*, as a validator registration mechanism or a Sybil attack avoidance mechanism. So that one may become a validator, prospective participants (or interested persons) must wager a certain number of coins to become eligible. The number of pledged tokens determines the number of voting shares destined for the validator. That method contains the main chain and many shards. The main chain serves as the ledger of identity, while the shard chains store the blockchain individual states and simultaneously process transactions. That algorithm uses randomness generation by combining a *VRF* and a *VDF* and incorporates *PoS* into the sharding process, shifting fragment protection concerns from a minimum number of nodes to a minimum number of voting shares.

Each voting particle represents one vote in the *BFT* consensus. Stackers receive voting shares proportional to their tokens. Voting shares are further taken up sharding randomly. Figure 2 depicts the sharding procedure interpretation by voting shares. Tokens become validators for the fragments they vote for.

A voting share is a virtual ticket allowing a validator to cast one consensus vote. Validators can purchase voting shares by betting tokens. The number of tokens required to vote is algorithmically adjusted. At the start of each phase, new shares with validator voting rights will be taken up to shards randomly. New validators join the shard where voting shares are assigned to them. Consensus in a given fragment is achieved by validators who, at least, must jointly have $2f + 1$ voting shares to sign

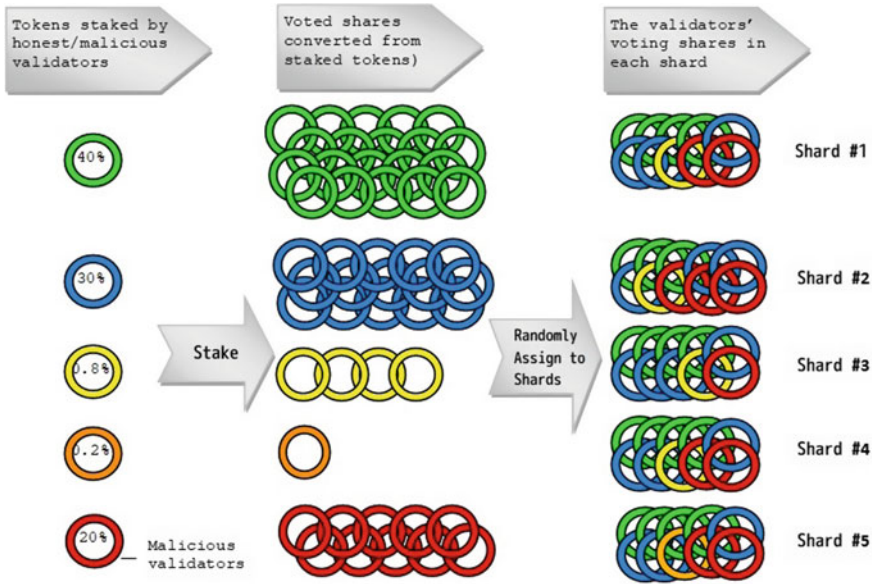


Fig. 2 Sharding procedure via shares voting (Source Authors' elaboration)

the block. To guarantee the safety of one shard, the number of voting shares in malicious validators needs to be less than one-third of all voting shares of this shard. The adaptive *PoS* of the proposed consensus method guarantees security requirements by adaptively adjusting the share price voting rights and taking up individual voting shares to shards instead of individual auditors.

The sharding security by voting shares is that even if one-fourth of all pledged tokens are malicious validators, one shard is assigned to one validator. Thus, the worst, when one malicious validator holds all the tokens (voting shares), less than one-third of the voting tokens in this shard. The reason is that the rates for each shard are m times less than the rates for the entire network, where m is shard's number. Thus, preventing the attack of malicious validators instead of sharding by validators, voting shares are split (one share has the right to vote, one shard is assigned).

In this method, the share price has a vote is set in algorithmical order so it is small enough to prevent malicious participants concentrate their voting power in one shard. The share price has the right to vote is set in tokens P_{vs} :

$$P_{vs} = \frac{Ts_{e-1}}{N_{sh} \cdot \lambda} \tag{1}$$

wherein: λ is security parameter;
 N_{sh} is the shard number;
 Ts_{e-1} is a total number of phase tokens $e - 1$.

Let us prove that when $\lambda > 600$, the probability of failure is negligible. In this case, the chance of a single shard should have over one third of the malicious voting shares.

Given the definition P_{vs} , the total number of voting shares is:

$$N_{sh} = \frac{Tse-1}{\lambda \cdot P_{vs}} \quad (2)$$

Given the credible source of randomness and the randomness-based sharding process, the probability distribution of the malicious voting shares in every part can be represented as a hypergeometric distribution (random sampling without replacement).

$$P(X = k) = \frac{C_n^N - C_k^K}{C_n^N} \quad (3)$$

wherein N is overall quantity of voting shares;

C_n^N is the number of combinations of N to n ;

N is total number of voting shares;

$K = \frac{N}{4}$ is maximum quantity of malicious voting shares;

$n = \frac{N}{N_{sh}}$ is the quantity of voting shares in every shard;

k is the quantity of malicious shares with voting rights in the shard.

Actual shard bounce rate $P(X \leq k)$ follows from the cumulative hypergeometric distribution (N, K, n, k) , which is in case if N is large, reduces to a binomial distribution (i.e., random sample with replacement):

$$P(X \leq k) = \sum_{i=0}^k n_i \cdot p_i \cdot (1 - p)^{n-i} \quad (4)$$

When n is high enough, the shard's probability that contains more than one third of malicious tokens is negligible.

When $n = 600$, the probability that the shard contains more than one third of malicious shares with voting rights $P(X \leq 200) = 0,999,997$, which shows the failure of such a shard, i.e., no consensus can be reached.

To ensure high security of the shard $\lambda = 600$. Parameter λ regulates the minimum number of voting shares that one shard must contain.

This solution is functional like the least quantity of nodes in a shard, stated in other *PoW*-based solutions [13–15]. This approach is robust against fluctuations in the number of validators. It does not set a lower limit on quantity of validators in each part like other solutions, e.g. *Zilliqa*. On the other hand, an adaptive *PoS*-based model was adopted so that attackers could never hold over one-third of the voting shares in one shard, making it reliable.

Peculiarity analysis of scalable randomness generation by *VRF* and *VDF* functions. The approach to the new blockchain formation developed in the thesis combines the advantages of the considered solutions. The proposed method uses the *BFT* consensus to ensure the random number finality. In particular, the protocol covers next steps:

1. The leader transmits a message to all validators with the last block hash $H(B_{n-1})$.
2. After receiving the message, the *VRF* is computed for every i validator to generate a random number r_i , and the proof $p_i: (r_i, p_i) = VRF(sk_i, H(B_{n-1}))$, v , wherein sk_i is the validator's secret key i, v is the current consensus number. Followed by each validator returning (r_i, p_i) to the leader.
3. The leader waits until he gets at least $f + 1$ real random numbers and combines them to get the preimage of *pRnd* ultimate randomness.
4. Leader provides *BFT* consensus of all validators to get consensus on *pRnd* and fix it to the B_n block.
5. After *pRnd* is executed, the leader begins to compute the actual probability $pRnd = VDF(pRnd, T)$, where T is the *VDF* complexity and is set algorithmically so that the probability can be computed only after k blocks. When the computation *pRnd* is in progress, the leader initiates *BFT* among all validators to reconcile reality *pRnd* and generate a probability on the blockchain (Fig. 3).

Thus, the *VDF* function is used to delay the expansion of *Rnd* evidentially and to prevent the malicious leader from preempting randomness by choosing a value for a subset of the *VRF* random numbers.

Owing to the *VDF* function, the leader cannot know the actual final randomness of *pRnd* until it is added to the blockchain. While *VDF* evaluates *rnd*, *pRnd* was sent to the previous block, so the leader can no longer manipulate it.

The worst that a malicious leader can do is either to add randomness to *pRnd*, or stop the protocol without fixing *pRnd*. That does not harm as the waiting mechanism switches the leader and restarts the protocol. In the long-term future, there could be invented *ASICs* (application-specific integrated circuits) to compute a *VDF* function that can find vulnerable nodes and compute result ahead other true nodes. However, at present, such robust circuits are not invented.

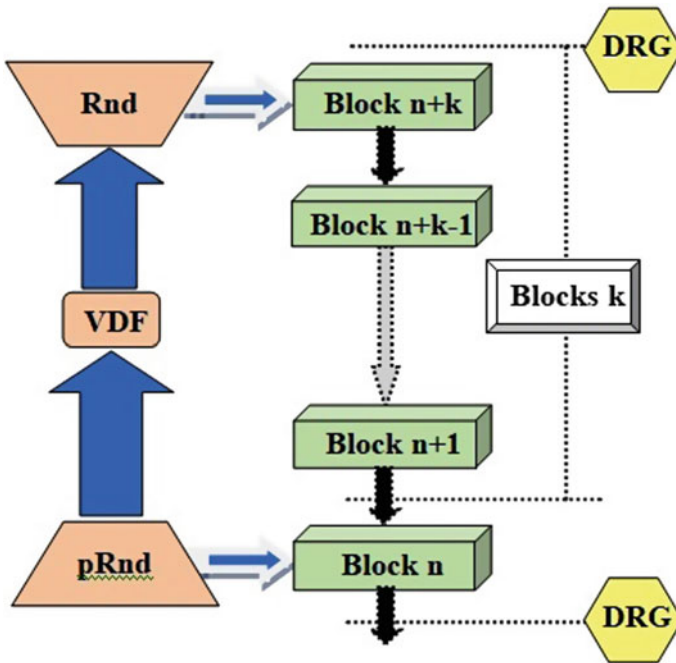


Fig. 3 A mechanism for delaying the final randomness detection by verifiable delay function (Source Authors' elaboration)

3 Conclusions

This paper's research is devoted to the development of a new and reliable blockchain system consensus method focused on the linearly scalable consensus mechanism. The proposed approach is based on the analysis of available consensus mechanisms, sharding and generation of distributed randomness. The proposed consensus mechanisms allow the development of a blockchain with the following advantages: full scalability, security, energy-efficient, and with fast consensus.

The proposed consensus method improves available ones by *BLS (Boneh-Lynn-Shacham)* multi-signature. However, it requires only one round-trip request. In this regard, the developed method is at least 50% faster than the *ByzCoin BFT* method. The paper presents an algorithm for conducting the consensus procedure. Consensus validators are elected based on an adaptive *Proof of Stake* model.

The proposed protocol differs from available *PBFT* in that a validator with more voting shares has more votes than others rather than one signature one vote. In order to become a validator, prospective participants (or interested parties) must stake a certain number of tokens to become eligible. The number of pledged tokens determines the number of voting shares. That method has the main chain and many shards. The main chain serves as the ledger of identity, while the shard chains store

the individual states of the blockchain and simultaneously process transactions. That algorithm uses randomness generation by combining *Verifiable Random Function* and *Verifiable Delay Function* and incorporates a *PoS* model into the sharding process that shifts fragment protection concerns from minimum nodes to a minimum voting shares. The number of tokens required to vote is algorithmically adjusted. At the start of each phase, for new validator voting shares there will be randomly assigned shards.

To guarantee the security of one shard, the number of voting shares in malicious validators needs to be less than one third of all voting shares of this shard. The adaptive *PoS* of the proposed consensus method guarantees security requirements by adaptively adjusting the share's price, has the right to vote, and assigns individual voting shares to shards rather than individual verifiers.

The sharding security by voting shares lies in the fact that even if for all pledged tokens 1/4 are harmful validators, then one shard is assigned to one validator. Then in the worst case, where a single malicious validator holds all the tokens (voting shares), it will have less than one third of the voting tokens in that shard.

To ensure high shard security, the network security parameter regulates the minimum voting shares that one shard must hold. Such a solution functionally corresponds to the minimum nodes in a certain network, described in some other solutions based on the *PoW* model. Thus, the presented approach is resistant to fluctuations in the number of validators. Moreover, it does not set a lower limit on the number of validators in each fragment, as in other solutions such as *Zilliqa*. Instead, an adaptive *PoS*-based model was adopted so that attackers can never occupy over one third of the voting shares in one shard, which makes it reliable.

The methods for creating the proposed blockchain improve available mechanisms for the functioning of the blockchain and have practical value for use in various digital economy sectors.

References

1. Centralized Decision Making Helps Kill Bad Products (2021) Harvard Business Review, August 31. <https://hbr.org/2016/10/centralized-decision-making-helps-kill-bad-products>
2. Krupskiy OP (2014) Modern management decision-making methods and their connection with the organizational culture of the tourism enterprises in Ukraine. *Econ Ann-XXI* 1(7–8):95–98
3. Campbell A, Kunisch S, Müller-Stewens G (2021) To centralize or not to centralize? McKinsey & Company, March 1. <https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/to-centralize-or-not-to-centralize>
4. Zavorotny A (2018) Analysis of practice of blockchain technology application in financial management. *Politech Stud J* 27. <https://doi.org/10.18698/2541-8009-2018-10-391>
5. Treiblmaier H (2019) Toward more rigorous blockchain research: recommendations for writing blockchain case studies. *Front Blockchain* 2. <https://doi.org/10.3389/fbloc.2019.00003>
6. Javed MU, Rehman M, Javaid N, Aldegheishem A, Alrajeh N, Tahir M (2020) Blockchain-based secure data storage for distributed vehicular networks. *Appl Sci* 10(6):2011. <https://doi.org/10.3390/app10062011>
7. Nick A, Hoenig L (2020) Consensus mechanisms in blockchain technology. *Lexology*, May 7. <https://www.lexology.com/library/detail.aspx?g=e30e7d54-3c7f-4ca0-8a22-478227a9b5ec>

8. The Truth about Blockchain (2019). Harvard Business Review, August 21. <https://hbr.org/2017/01/the-truth-about-blockchain>
9. Cain D (2019) Big data synchronization: 5 ways to ensure big data accuracy. Medium, July 10. <https://towardsdatascience.com/big-data-synchronization-5-ways-to-ensure-big-data-accuracy-4c4801b021ad>
10. Gimenez-Aguilar M, de Fuentes JM, Gonzalez-Manzano L, Arroyo D (2021) Achieving cybersecurity in blockchain-based systems: a survey. *Futur Gener Comput Syst* 124:91–118. <https://doi.org/10.1016/j.future.2021.05.007>
11. Tyagi N (2020) Top 10 big data technologies/analytics steps. Analyticssteps.com. <https://www.analyticssteps.com/blogs/top-10-big-data-technologies-2020>
12. Gour R (2019) Big data architecture—the art of handling big data. Medium, September 18. <https://towardsdatascience.com/big-data-architecture-the-art-of-handling-big-data-bc565c3a7295>
13. Wroughton J, Cole T (2013) Distinguishing between binomial, hypergeometric and negative binomial distributions. *J Stat Educ* 21(1). <https://doi.org/10.1080/10691898.2013.11889663>
14. Shvachych G, Busygin V, Zaporozhchenko O, Sazonova M (2020) Some aspects of the practical implementation of the blockchain technology. In: Savchuk L, Bandorinaya L (eds) Monograph: state, industries, enterprises, business: realities and trends of economic, informational and technical development. Porogy, pp 284–304
15. Shvachych G, Ivanov R, Busygin V (2019) Blockchain technology as a means of improving enterprise efficiency. In Shebeko K (ed) Collection of scientific papers of the tenth international scientific and practical conference on banking economics, October 2019. PolesGU, pp 364–368

Acquisition of a CNC Router for a Joinery in Brazil: An Approach from VFT, SAPEVO-M and WASPAS Methods



Lucas Ramon dos Santos Hermogenes, Igor Pinheiro de Araújo Costa, Marcos dos Santos, and Carlos Francisco Simões Gomes

Abstract The work demonstrates the application of Value-Focused Thinking (VFT) and multicriteria methods to assist in the process of choosing a Computer Numeric Control (CNC) Router. We analyzed a company that needed to hire professionals such as an engineer, carpenters, and assistants to be able to deliver the necessary demand. However, the company has the goal of producing a limited number of pieces per month and the limiting factor for these deliveries is specifically the lack of skilled labor since there is a difficulty in finding carpenters with experience in solid wood and with basic skills such as reading and interpreting technical drawings. The VFT was used to structure the problem and then the decision support methods SAPEVO-M to calculate the weights of the criteria and WASPAS to order the alternatives, indicating which equipment is the most advantageous for the strategic decisions of the organization.

Keywords Simple aggregation of preferences expressed by ordinal vectors · Weighted aggregated sum product assessment · Computer numeric control router · Woodwork · Furniture · Multicriteria · Decision-making

1 Introduction

Brazil, as shown by the survey conducted by [1], had an unemployment rate of 14.1% in the first quarter of 2021, and has been showing constant growth since the fourth quarter of 2019, when the rate was at 11%. One of the factors that caused the increase in the unemployment rate was the COVID-19 pandemic, affecting the economy, education, scientific research, political relations, and unemployment. Maêda et al.

L. R. dos Santos Hermogenes · I. P. de Araújo Costa (✉) · C. F. S. Gomes
Fluminense Federal University (UFF), Niterói 24210-346, Brazil
e-mail: costa_igor@id.uff.br

I. P. de Araújo Costa · M. dos Santos
Naval Systems Analysis Center (CASNAV), Rio de Janeiro 20091-000, Brazil

M. dos Santos
Military Institute of Engineering (IME), Urca 22290-270, Brazil

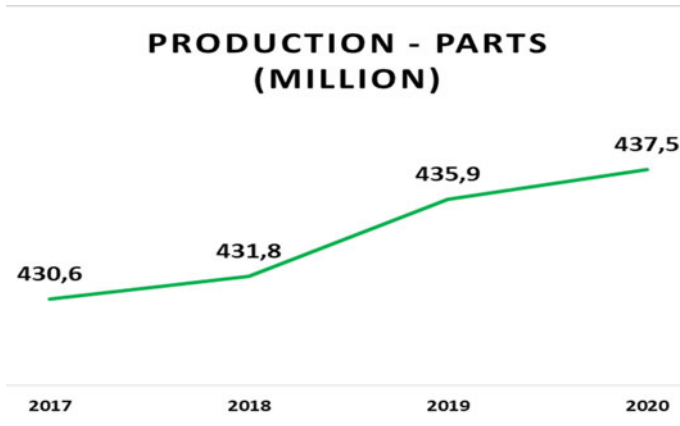


Fig. 1 Increased furniture production. Adapted from [4]

[2] state that the consequences of the pandemic caused by SARS-CoV-2 (COVID-19) are overwhelming in the most diverse sectors of the Brazilian economy, and the long-term effects are still difficult to fully measure. For [3], companies had to adapt by adopting the home office to continue delivering results, to face the pandemic and reduce the number of layoffs. However, for some companies the reality was different, and bankruptcy was inevitable. According to [4], the furniture sector presented a different result and had an increase in the production of pieces, as shown in Fig. 1.

One of the factors for the increase in production was social isolation, which motivated the change in the decoration of some Brazilian homes. This trend is observed especially for people who consume luxury furniture, and the analysed company was positively impacted by this external factor, since the average of production and sales increased 85.7% from 2020 to 2021, going from an average of 70 pieces per month to 130. Figure 2 shows that in 2017 and 2018 the sector had a drop in job creation, but precisely in the pandemic, there was a resumption in hiring for the years 2019 and 2021.

In 2021, the company adopted investment measures to guarantee deliveries, with the hiring of personnel, the search for the outsourcing of production stages in specific parts and the research of automated machines, which are fundamental steps to alleviate internal problems observed with the increase in demand. With that, the company wants to acquire a Computer Numeric Control (CNC) Router, but the process of choosing the best model is very complex, due to conflicting criteria and alternatives available in the market.

Numerous Criteria Decision Analysis (MCDA) is used as an umbrella word in this context to denote a range of formal methodologies that strive to take explicit account of multiple criteria in assisting stakeholders and groups in exploring important decisions [5].

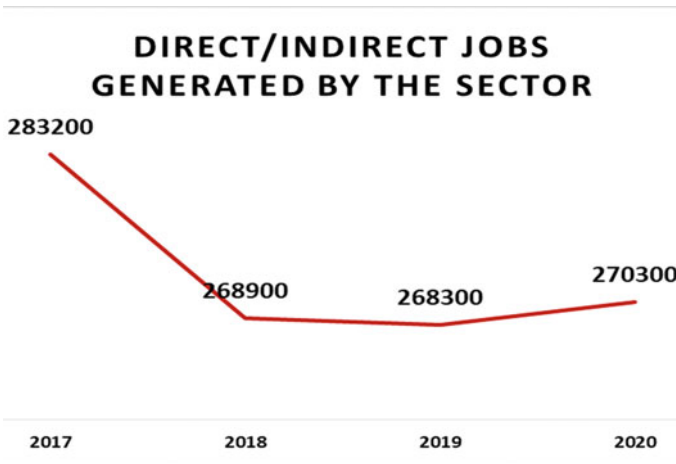


Fig. 2 Jobs generated in the sector. Adapted from [4]

Almost always, these judgments contain many conflicting objectives, different forms of non-repeatable uncertainty, costs and advantages accruing to various individuals, enterprises, groups, and other organisations [6].

Despite of the variety of MCDA strategies, methods, and techniques, the fundamental components of MCDA are a limited or infinite set of actions (alternatives, solutions, courses of action, etc.), at least two criteria, and at least one Decision Maker (DM). When confronted with a complex circumstance, the DM strives to select the optimal alternative [7]. Given these fundamental characteristics, MCDA is a decision-making activity that aids in the selection, ranking, or sorting of actions [8].

Among the existing methods, the research was carried out using the Value-Focused Thinking (VFT) approach, SAPEVO-M and WASPAS methods. The VFT approach was chosen because, according to [9], it is an approach that represents value-focused thinking for structuring problems. The SAPEVO-M and WASPAS methods were used because of all the uncertainty present in the decision-making process [10].

2 Problem

The company Luxury Handcrafted Furniture, a fictitious name due to authorization, faces a problem to getting qualified labour. Its manufacturing processes and steps are completely manual, and this is part of the essence of the organization, however, the company also has a significant variety of parts, and now added to an unexpected increase in demand, where this factor together with handbook work has also been causing delays in deliveries.

Manual work in solid wood requires technical skills that are currently not being much explored by carpentry professionals, which is observed is a greater dedication to develop skills to work with plywood and MDF sheets—Medium Density Fiberboard, or Fiberboard medium density.

To mitigate the risk of not being able to deliver products due to a possible reduction in the woodworking team, one of the solutions, in addition to the internal development of professionals, was to research to buy a CNC Router. This equipment provides greater safety and contributes to increasing the current production by over 23.07%, from 130 to 160 pieces per month, in addition to ensuring higher quality in the machining stages (cutting and drilling).

Another relevant point about the acquisition of equipment is a more fluid production and more directed towards the assembly of furniture since strategically the machining steps provide approximately 30% of the need for operating hours. Thus, this time can be used to assemble with more care and quality of the various parts of the orders in the portfolio, without the need to work overtime.

Because of this, the study in question seeks to answer the following questions: Which CNC Router should be acquired to increase the production of joinery? And how this structuring should be based so that the choice can be assisted from the Operational Research?

3 Background

3.1 Multicriteria Decision Analysis

As indicated by [11], in many real-world problems, decision-makers are faced with complex decisions, so multicriteria methods are used to help in the decision-making process, where axiomatic models provide greater security in decision making.

This process generally involves a choice between several alternatives. The feasible alternatives of meeting the objective, and selected for evaluation, are compared according to criteria and under the influence of attributes [12]. The MCDA methods are very useful to support the decision-making process in these cases because they consider value judgments and not only technical issues, to evaluate alternatives to solve real problems, presenting a high multidisciplinary [13]. The MCDA methods have been employed to support the decision-making process in several recent complex problems [10, 14–25].

3.2 Value-Focused Thinking (VFT) Approach

VFT was developed by [9]. The methodology is used to structure problems in such a way that the value involved in the decision process needs to be the basis for decision

making. It is understood that the value mentioned is not monetary, but contextual. Values are principles used for evaluation.

They range from ethical principles that must be maintained to guidelines for preferences between choices and as indicated by [9]: “VFT essentially consists of two activities: first deciding what you want and then, find out how to get it”. Figure 3 demonstrates a clear organization of value-focused thinking with some ramifications.

Values must be the driving force for decision making, they must be the basis of time and effort that are spent thinking about decisions. “The fundamental notion in decision making should be values, not alternatives” [9]. Two objectives are part of the structuring of problems: the fundamental one, understood as the main one and all the others are actions that lead to it; and the objectives, which are kind of essential for the structuring of the problem.

For [26], it is possible to structure the problem by evaluating the inputs into the problematic situation, where tools such as Brainstorming, meetings with stakeholders and structuring from the VFT contribute with action plans to define goals, criteria, alternatives and the method to be applied in the process decision, where this phase can be identified with the convergent one.

Strategically, the DMs involved in the process understand that the increase in production is also related to training and qualification, partnerships with educational

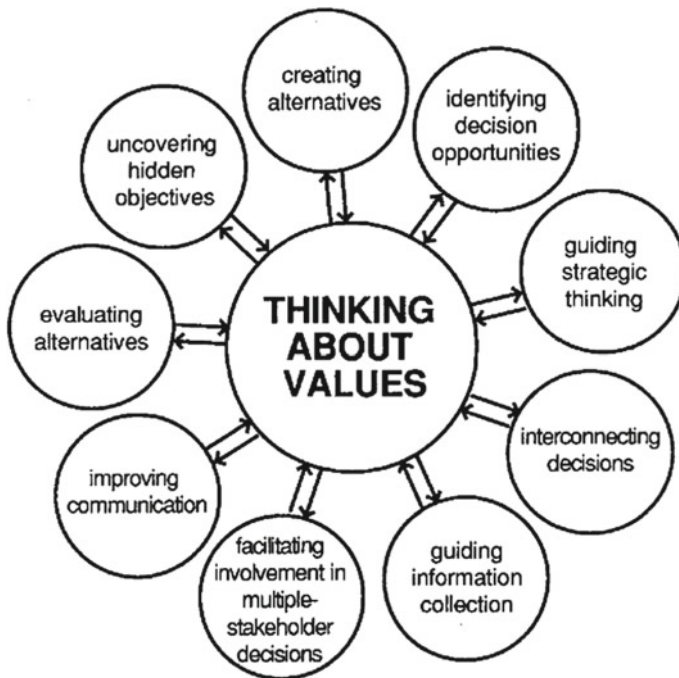


Fig. 3 Thinking about values. Adapted from [9]

centres to seek specialized labour and the acquisition of machinery and equipment that ensure a more automated process in certain stages of manufacturing.

3.3 SAPEVO-M

Gomes et al. [10] demonstrate the SAPEVO-M method (Simple Aggregation of Preferences Expressed by Ordinal Vectors–Multi-decision Makers) which represents an improvement on the original version of SAPEVO, initially generated by [27], where evolution allowed the application of multiple decision-makers, in addition to introducing a process of standardization of evaluation matrices.

As shown by [10], the SAPEVO-M method can be divided into the following steps:

- Transforms the ordinal preferences of the criteria into a weight vector for the criteria;
- Transforms the alternatives’ ordinal preferences into a weight vector for the alternatives;
- Determines the overall weights for the alternatives.

The method has been widely used in solving complex problems: Berriel et al. [28] used the SAPEVO-M and VIKOR method in decision making about the economic recovery in the municipality of Nilópolis during the covid-19 pandemic; Costa et al. [29] used the method to choose a Brazilian navy ship to act as a field hospital in the fight of the COVID-19 pandemic.

To make the comparisons, the relative importance between the criteria is used, as shown in Table 1.

Table 1 Relative importance between the criteria. Adapted from [10]

| Scale 1 | Verbal representation | Scale 2 |
|---------|---------------------------|---------|
| <<<1 | Absolutely less important | −3 |
| <<1 | Much less important | −2 |
| <1 | Less important | −1 |
| 1 | Equally important | 0 |
| >1 | More important | 1 |
| >>1 | Much more important | 2 |
| >>>1 | Absolutely more important | 3 |

3.4 WASPAS

The Weighted Aggregated Sum Product Assessment (WASPAS) method was created by [30], being the combination of the Weighted Sum Method and Weighted Product Method.

The method is considered compensatory and is designed to increase the accuracy of the final ordering. Its applicability occurs initially through the elaboration of the decision matrix, where it is necessary to insert the data of each alternative for each criterion and the weights, attributed by the decision-makers or calculated from other multicriteria methods.

The first step after assembling the decision matrix is to normalize the matrix, and for this two equations are used, in which one applies Eq. (1) when the criteria are profit, benefit, or maximization, and Eq. (2) for cost or minimization criteria [30].

$$\bar{x}_{ij} = \frac{x_{ij}}{\max_i x_{ij}} \quad (1)$$

$$\bar{x}_{ij} = \frac{\min_i x_{ij}}{x_{ij}} \quad (2)$$

Afterward, it is necessary to calculate the total relative importance based on the WSM method, according to Eq. (3).

$$Q_i^{(1)} = \sum_{j=1}^n \bar{x}_{ij} w_j \quad (3)$$

With the total relative importance calculated the next step is to calculate the total relative importance based on the WPM method, as indicated in Eq. (4).

$$Q_i^{(2)} = \prod_{j=1}^n (\bar{x}_{ij})^{w_j} \quad (4)$$

With the two calculated amounts, the total relative importance is verified, from the generalized equation of the set (Q), Eq. (5).

$$Q_i = \lambda Q_i^{(1)} + (1 - \lambda) Q_i^{(2)}, \lambda = 0, \dots, 1. \quad (5)$$

In the end, with the calculated values, the highest-ranking values are observed, as the ordering of the method indicates that the higher the final value, the better the alternative.

4 Practical Application of SAPEVO-M WASPAS Methods when Choosing a Router

To solve the problem observed, which was to increase production capacity and minimize labour shortages, the divergent and convergent phases diamond was applied. A brainstorming was carried out between the production coordination, the operational directors, the financial directors, and the consulting company that assists in financial and industrial processes.

In the end, the group realized that the company needed to automate some processes, but without losing the essence of manual work, it was also discussed which actions can be applied to train and qualify the current team to improve the quality of the workforce.

It could be outsourced, and machines can contribute to the company’s productive capacity to increase, all with the thought focused on the maximum preservation of manual work. These are differentiating points in the company’s marketing campaigns and one of the most important values of the organization, as shown in Fig. 4.

Upon entering the convergent phase, the company left a problematic situation and entered a situation of developing action plans, where there was a clear objective, which was to purchase a CNC Router to automate crucial steps in the manufacturing process, increasing production capacity and at the same time preserving the manual work for the steps following those processed in the equipment.

The equipment criteria and the MCDA method used to assist in the decision-making process with the ordering of alternatives were also defined. The acquisition of equipment has contributed to the company’s fundamental objective, which is to increase production capacity, and this can be seen in Fig. 5.

The next step was to define the equipment criteria, namely:

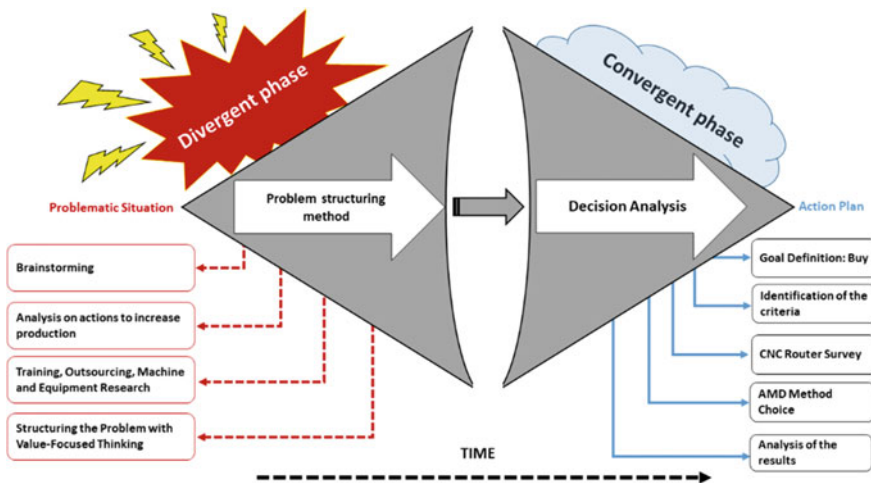


Fig. 4 Divergent and convergent phases in the decision problem. Adapted from [26]

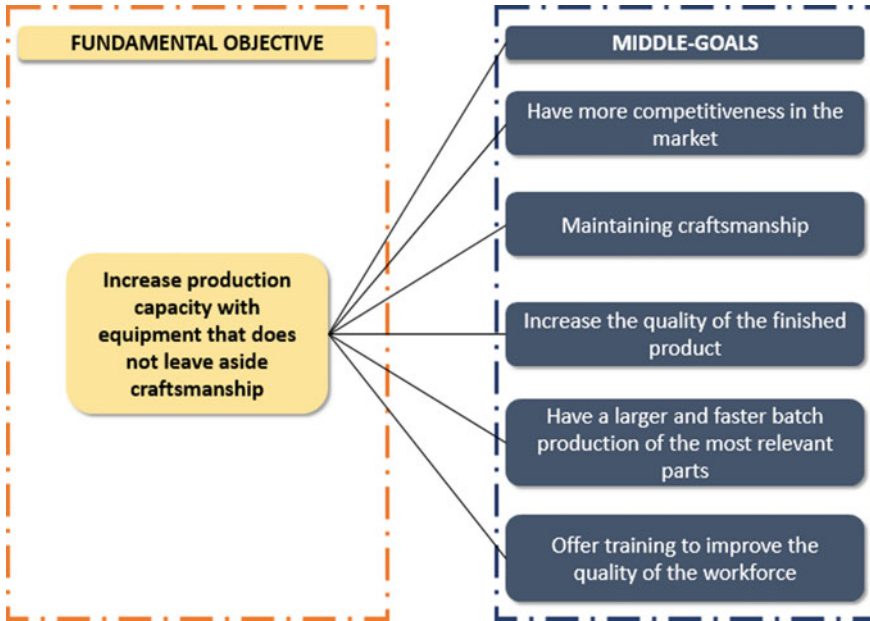


Fig. 5 Goals network. Adapted from [9]

- Work area (m²): The greater the useful working area of the equipment, the better, because in addition to the machine working with solid wood, it will also be necessary to cut MDF and Plywood sheets;
- Delivery time (days): The faster the delivery, the better;
- Z-axis (mm): indicates the operating limit height, the shaft with too short height makes it impossible to operate wood boards with thicker thicknesses;
- Cutting motors (unit): Usually, the equipment has only one cutting motor (Spindle), but some machines are sold with two motors, which has a certain advantage over the others;
- Cutting motor power—Spindle (hp): The more power of the cutting motor, the better;
- Cutting speed (m/min): The faster the cut, the better, as this directly affects the production capacity of the equipment;
- Cost (BRL): Equipment sales price, the lower the better;

The next step was the definition of the MCDA method, where SAPEVO-M was used to calculate the weights through the computational tool [31] and the WASPAS method was used to sort the alternatives, using an excel spreadsheet.

After defining the criteria, a decision matrix was built, with each alternative and each criterion, where the data were collected through interviews with equipment sellers and machine catalogs (Table 2).

Table 2 Decision matrix

| Alternatives | Work area (m ²) | Term (days) | Z axis (mm) | Cutting motors (unit) | Spindle power (hp) | Cutting speed (m/min) | Cost (BRL) |
|-----------------------------|-----------------------------|-------------|-------------|-----------------------|--------------------|-----------------------|------------|
| SCRIBA 2013 H1S0TR | 2.6 | 120 | 190 | 1 | 3.3 | 4.2 | 210,000.00 |
| Mobile J/A-CNC28182-MBL-TAF | 5.18 | 70 | 200 | 1 | 6 | 12 | 136,970.00 |
| Solid J/A-CNC28182-TAF | 5.18 | 70 | 200 | 1 | 4.7 | 15 | 185,000.00 |
| RMC 3000 PLUS 2S | 5.09 | 50 | 150 | 2 | 4 | 7 | 82,419.00 |

To participate in the decision-making process, two DMs made the comparisons to obtain the weights—operational and financial managers. Each DM compared pair by pair each criterion with each other and then each alternative considering each criterion. At the end of the comparison, the weights were calculated as shown in Fig. 6.

It was observed that the most important criteria are speed and cost, respectively. After calculating the weights, they were used in the WASPAS method as shown in Table 3.

| Criteria | weights |
|-----------------------------|---------|
| Work area (m ²) | 0,74 |
| Term (days) | 0,64 |
| Z axis (mm) | 0,49 |
| Cutting motors (unit) | 0,82 |
| Spindle power (cv) | 0,70 |
| Cutting speed (m/min) | 1,54 |
| Cost (BRL) | 1,00 |

Fig. 6 Weights calculated from the SAPEVO-M method

Table 3 Decision matrix and weights of the criteria

| Alternatives | 0.74 | 0.64 | 0.49 | 0.82 | 0.70 | 1.54 | 1.00 | λ |
|-----------------------------|------|------|------|------|------|------|------------|-----|
| | Max | Min | Max | Max | Max | Max | Min | |
| SCRIBA 2013 H1S0TR | 2.6 | 120 | 190 | 1 | 3.3 | 4.2 | 210,000.00 | 0.5 |
| Mobile J/A-CNC28182-MBL-TAF | 5.18 | 70 | 200 | 1 | 6 | 12 | 136,970.00 | |
| Solid J/A-CNC28182-TAF | 5.18 | 70 | 200 | 1 | 4.7 | 15 | 185,000.00 | |
| RMC 3000 PLUS 2S | 5.09 | 50 | 150 | Two | 4 | 7 | 82,419.00 | |

After identifying the cost and benefit criteria, it is necessary to normalize the matrix according to Eqs. (1) and (2), and then verify the relative importance from Eq. (3), according to Table 4.

Soon after, Eq. (4) is used to calculate the weighted multiplicative relative importance, and Eq. (5) to verify the final ordering of the set (Q), as shown in Table 5.

At the end of all the procedures, it is verified that the option RMC 3000 PLUS 2S is the most suitable for the company to be able to increase its production, the machine can be used in the machining and drilling stages of various parts.

When performing the comparisons in the SAPEVO WEB tool, it was also possible to observe the ordering by the SAPEVO-M method, Fig. 7, and the alternative RMC 3000 PLUS 2S was also indicated as the best option.

Table 4 Weighted additive relative importance matrix

| | Work area (m ²) | Term (days) | Z axis (mm) | Cutting motors (unit) | Spindle power (cv) | Cutting speed (m/min) | Cost (BRL) | Output 1 |
|-----------------------------|-----------------------------|-------------|-------------|-----------------------|--------------------|-----------------------|------------|----------|
| SCRIBA 2013 H1S0TR | 0.3726 | 0.2660 | 0.4640 | 0.4077 | 0.3852 | 0.4308 | 0.3929 | 2.7192 |
| Mobile J/A-CNC28182-MBL-TAF | 0.7423 | 0.4560 | 0.4885 | 0.4077 | 0.7004 | 1.2308 | 0.6023 | 46280 |
| Solid J/A-CNC28182-TAF | 0.7423 | 0.4560 | 0.4885 | 0.4077 | 0.5486 | 1.5385 | 0.4460 | 46276 |
| RMC 3000 PLUS 2S | 0.7291 | 0.6385 | 0.3663 | 0.8154 | 0.4669 | 0.7179 | 1.0010 | 47351 |

Table 5 Final ranking of alternatives

| | Work area (m ²) | Term (days) | Z axis (mm) | Cutting motors (unit) | Spindle power (cv) | Cutting speed (m/min) | Cost (BRL) | Output 2 | Output (Q) | Ranking |
|-----------------------------|-----------------------------|-------------|-------------|-----------------------|--------------------|-----------------------|------------|----------|------------|----------|
| SCRIBA 2013 H1S0TR | 0.5995 | 0.5718 | 0.9753 | 0.5683 | 0.6579 | 0.1411 | 0.3921 | 0.0069 | 1.3631 | 4 |
| Mobile J/A-CNC28182-MBL-TAF | 1,000 | 0.8067 | 1,000 | 0.5683 | 1,000 | 0.7094 | 0.6014 | 0.1956 | 2.4118 | 2 |
| Solid J/A-CNC28182-TAF | 1,000 | 0.8067 | 1,000 | 0.5683 | 0.8428 | 1,000 | 0.4451 | 0.1720 | 2.3998 | 3 |
| RMC 3000 PLUS 2S | 0.9867 | 1,000 | 0.8689 | 1,000 | 0.7528 | 0.3096 | 1000 | 0.1998 | 2.4675 | 1 |



Fig. 7 SAPEVO-M ordering

Another relevant point is that in the ordering of the WASPAS and SAPEVO-M methods, the alternative SCRIBA 2013 H1S0TR was in last place, and, for logical reasons, it is not the indicated equipment to be purchased.

5 Final Considerations

The application of the methods to calculate the weights and to verify the final ordering of the alternatives proved to be very relevant, and the results obtained may contribute to the decision-makers to be able to make the most suitable acquisition for the company.

The application of the VFT method guaranteed an adequate structuring, maintaining the main value that the company seeks to preserve, which is handmade work.

With the application of VFT, decision-makers were not only looking for an automated machine to increase production, but they were looking for a machine to automate a part of the process. In this way, other parts such as assembly and certain more refined operations can still be carried out traditionally.

Another relevant factor is that to avoid the scenario of lack of labour, in addition to purchasing the equipment, the company is also focused on training and qualifying internal employees so that they develop as professionals and the company has more and more quality in their manufacturing processes.

The machine recommended for purchase is equipment that will certainly contribute to the fundamental objective defined in the VFT and the ordering both by the WASPAS method and by the SAPEVO-M method provide greater security in the decision process.

This combined methodology made it possible to obtain the weights of the criteria considering the opinion of several experts, which served as the basis for decision-making process through the application of WASPAS.

As a result, it was possible to choose the best equipment and knowing which machine should be avoided, which is also extremely relevant to make the decision much more concrete.

As a limitation of this study, we highlight the analysis of a reduced number of machines, but we emphasize that the methodology can be expanded to cases with more alternatives to be evaluated.

References

1. IBGE (2021) Unemployment [Internet]. Cited 2022 Jan 27. <https://www.ibge.gov.br/indicadores/#desemprego>
2. Maêda SM do N, Basílio MP, Costa IP de A, Moreira MÂL, dos Santos M, Gomes CFS et al (2021) Investments in times of pandemics: an approach by the SAPEVO-M-NC method. In:

- Modern management based on big data II and machine learning and intelligent systems III [Internet], pp 162–168. <https://ebooks.iospress.nl/doi/10.3233/FAIA210244>
3. Hermogenes LR, Santos M, Nascimento PF, Teixeira LF (2020) A IMPORTÂNCIA DAS DIGITAL SKILLS EM TEMPOS DE CRISE: alguns aplicativos utilizados durante o isolamento social devido à pandemia do covid-19. *Rev Augustus* 25(51):198–218
 4. ABIMABLE. Industry data [Internet]. Cited 2022 Jan 27. <http://abimovel.com/capa/dados-do-setor/>
 5. Belton V, Stewart T (2002) *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media
 6. Keeney RL, Raiffa H, Meyer RF (1993) *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge University Press
 7. Zhang J, Kou G, Peng Y, Zhang Y (2021) Estimating priorities from relative deviations in pairwise comparison matrices. *Inf Sci (NY)* 552:310–327
 8. Greco S, Figueira J, Ehrgott M (2016) *Multiple criteria decision analysis: state of art surveys*, vol 37. Springer
 9. Keeney RL (1992) *Value-focused thinking: a path to creative decisionmaking*. Harvard University, Cambridge, MA
 10. Gomes CFS, Santos M dos, Teixeira LFH de S de B, Sanseverino AM, Barcelos MR dos S (2020) SAPEVO-M: a group multicriteria ordinal ranking method. *Pesqui Operacional [Internet]*, p 40. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-74382020000100212&tlng=en
 11. Teófilo PLBDC, Gomes CFS, dos Santos M. Estratégias De Combate À Segunda Onda De Infecções Por Covid-19: Priorização De Ações Por Meio Do Método Multicritério Wings
 12. Cardoso RS, Xavier LH, Gomes CFS, Adissi PJ (2009) Uso de SAD no apoio à decisão na destinação de resíduos plásticos e gestão de materiais. *Pesqui Operacional [Internet]*. 2009 Apr. 29(1):67–95. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-7438200900100004&lng=pt&tlng=pt
 13. Santos M dos, Quintal RS, Paixão AC da, Gomes CFS (2015) Simulation of operation of an integrated information for emergency pre-hospital care in Rio de Janeiro Municipality. *Procedia Comput Sci [Internet]* 55:931–938. <https://linkinghub.elsevier.com/retrieve/pii/S1877050915015860>
 14. Costa IP de A, Sanseverino AM, Barcelos MR dos S, Belderrain MCN, Gomes CFS, Santos M dos (2021) Choosing flying hospitals in the fight against the COVID-19 pandemic: structuring and modeling a complex problem using the VFT and ELECTRE-MOR methods. *IEEE Lat Am Trans [Internet]* 19(6):1099–106. <https://ieeexplore.ieee.org/document/9451257/>
 15. Costa IP de A, Maêda SM do N, Teixeira LFH de S de B, Gomes CFS, Santos M dos (2020) Choosing a hospital assistance ship to fight the Covid-19 pandemic. *Rev Saude Publica [Internet]* 54. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090141296&doi=10.11606%2FS1518-8787.2020054002792&partnerID=40&md5=90355a4a86a1b09b1add8956ace15019>
 16. Santos M dos, Costa IP de A, Gomes CFS (2021) Multicriteria decision-making in the selection of warships: a new approach to the AHP method. *Int J Anal Hierarchy Process [Internet]*. 2021 May 19;13(1). <https://ijahp.org/index.php/IJAHp/article/view/833>
 17. Maêda; SM do N, Costa IP de A, Castro Junior MAP, Fávero LP, Costa AP de A, Corriça JV de P et al (2021) Multi-criteria analysis applied to aircraft selection by Brazilian Navy. *Production* 31:1–13
 18. De Almeida IDP, Corriça JV de P, Costa AP de A, Costa IP de A, Maêda SM do N, Gomes CFS, et al (2021) Study of the location of a second fleet for the Brazilian Navy: structuring and mathematical modeling using SAPEVO-M and VIKOR methods. *ICPR-Americas 2020 Commun Comput Inf Sci [Internet]* 1408:113–24. https://link.springer.com/10.1007/978-3-030-76310-7_9
 19. Moreira MÁL, Costa IP de A, Pereira MT, dos Santos M, Gomes CFS, Muradas FM (2021) PROMETHEE-SAPEVO-M1 a hybrid approach based on ordinal and cardinal inputs: multi-criteria evaluation of helicopters to support Brazilian Navy Operations. *Algorithms [Internet]*. 2021 Apr 27;14(5):140. <https://www.mdpi.com/1999-4893/14/5/140>

20. Oliveira AS, Gomes CFS, Clarkson CT, Sanseverino AM, Barcelos MRS, Costa IPA et al (2021) Multiple criteria decision making and prospective scenarios model for selection of companies to be incubated. *Algorithms* [Internet], 2021 Mar 30;14(111). <https://www.mdpi.com/1999-4893/14/4/111>
21. Tenório FM, dos Santos M, Gomes CFS, Araujo J de C (2020) Navy warship selection and multicriteria analysis: the THOR method supporting decision making. In: *International joint conference on industrial engineering and operations management* [Internet]. Springer, pp 27–39. http://link.springer.com/10.1007/978-3-030-56920-4_3
22. Costa IP de A, Moreira MÁL, Costa AP de A, Teixeira LFH de S de B, Gomes CFS, Santos M Dos (2021) Strategic study for managing the portfolio of IT courses offered by a corporate training company: an approach in the light of the ELECTRE-MOr multicriteria hybrid method. *Int J Inf Technol Decis Mak* [Internet], 2021 Sep 17;1–29. <https://www.worldscientific.com/doi/abs/10.1142/S0219622021500565>
23. Drumond P, de Araújo Costa IP, Lellis Moreira MÁ, dos Santos M, Simões Gomes CF, do Nascimento Maêda SM (2022) Strategy study to prioritize marketing criteria: an approach in the light of the DEMATEL method. *Procedia Comput Sci* [Internet], 2022;199:448–455. <https://linkinghub.elsevier.com/retrieve/pii/S1877050922000540>
24. Mellem PMN, de Araújo Costa IP, de Araújo Costa AP, Lellis Moreira MÁ, Simões Gomes CF, dos Santos M et al (2022) Prospective scenarios applied in course portfolio management: an approach in light of the Momentum and ELECTRE-MOr methods. *Procedia Comput Sci* [Internet]. 199:48–55. <https://linkinghub.elsevier.com/retrieve/pii/S1877050922000072>
25. De Paula NOB, de Araújo Costa IP, Drumond P, Lellis Moreira MÁ, Simões Gomes CF, dos Santos M et al (2022) Strategic support for the distribution of vaccines against Covid-19 to Brazilian remote areas: a multicriteria approach in the light of the ELECTRE-MOr method. *Procedia Comput Sci* [Internet] 199:40–47. <https://linkinghub.elsevier.com/retrieve/pii/S1877050922000060>
26. Abuabara L, Paucar-Caceres A, Burrows-Cromwell T (2019) Consumers' values and behaviour in the Brazilian coffee-in-capsules market: promoting circular economy. *Int J Prod Res* 57(23):7269–7288
27. Gomes L, Mury A-R, Gomes CFS (1997) Multicriteria ranking with ordinal data. *Syst Anal* 27(2):139–146
28. Berriel PC, Gomes CFS, Costa IP de A, Santos M dos (2020) Retomada econômica no município de Nilópolis durante a pandemia de COVID-19: uma análise a partir do método ELECTRE-MOr. *Congr Int XXXI EPIO–XXXIII ENDIO y RED-M IX*
29. Costa IP de A, Maêda SM do N, Teixeira LFH de S de B, Gomes CFS, Santos M dos. APOIO HUMANITÁRIO AO COMBATE À PANDEMIA DE COVID-19: uma abordagem multicritério para escolha de navio da marinha do brasil mais adequado a ser empregado. *Rev Augustus* 25(51):56–78
30. Zavadskas EK, Turskis Z, Antucheviciene J, Zakarevicius A (2012) Optimization of weighted aggregated sum product assessment. *Elektron ir Elektrotechn* 122:3–6
31. Teixeira LFHSB, Santos M dos, Gomes CFS (2018) SapevoWeb Software (v.1), sob registro INPI: BR512020000667-1 [Internet]. Cited 2020 Aug 14. <http://www.sapevoweb.com>

General Collective Intelligence as a Platform for Social Technology



Andy E. Williams

Abstract A recently developed method for systems modeling called Human-Centric Functional Modeling (HCFM) has been used to develop a model for a General Collective Intelligence or GCI platform designed to enable groups to self-assemble into massive networks of cooperation predicted to have the capacity to exponentially increase impact on collective goals. The goal of Social Technologies have been described as achieving the “ability of a networked society to resolve social problems for itself, i.e., without government intervention”. This short article explores how GCI is predicted to exponentially increase impact on that goal.

Keywords General collective intelligence · Human-centric functional modeling · Social platforms

1 Introduction

The newly emerging science of Human-Centric Functional Modeling or HCFM is targeted at significantly increasing the capacity of humans to model and navigate complexity in any system [1]. HCFM hypothesizes that the behavior of any system can be represented in terms of functional state spaces which the system moves through. In a functional state space each region represents a state that the system might occupy. Each state is separated from other regions (other states) by the processes through which that state might transition to those other states. In the case that a functional state space is “spanned” by some basic set of functions, these processes represent some combination of those functions. These states are “functional states” because they are described in terms of those processes composed of functions. As an example, HCFM represents the functional state space of the human cognitive system in terms of a space of concepts or conceptual space in which the cognition can potentially transition from one concept to another concept through reasoning processes. These reasoning processes are hypothesized to be spanned by

A. E. Williams (✉)
Nobeah Foundation, Nairobi, Kenya
e-mail: awilliams@nobeahfoundation.org

a set of four basic functions [2]. The importance of being able to apply HCFM to cognition is that it allows the application to any other system to be understood by analogy with this case in which all the function states are concepts, and in which all the processes composed of functions consist of reasoning, and so are all by definition conceptualizable.

2 Shortcomings of Current Centralized and Decentralized Platform Approaches

Censorship, surveillance and lack of privacy, lack of transparency, user coercion due to closed and proprietary platforms that remove options, and other shortcomings of centralized platforms that symbolize misalignment with user goals are well known [3]. What is less well known is that seemingly decentralized platforms also suffer from the same misalignment because unless they are governed by a decentralized governance process as well. Because any element of centralization in any process, whether user interface functionality, security, business logic, data storage and access functionality, or in any other functionality as well as in the governance of that functionality, will naturally prioritize the interests of the centralized decision-maker [4–6]. For this reason it is important to distinguish true decentralization from the mere distribution of centralized interests [7].

3 Using Functional State Spaces to Define What Complete Decentralization Means

Functional state space is hypothesized to have the capacity to represent the complete set of behaviors of any system in any given domain of behavior. Systems are described as moving through functional state space as they execute their functions. Since a problem in functional state space is defined as the lack of a path for the system to navigate from one functional state to another, then to have general problem-solving ability, that is to have the ability to solve any problem in general, the system must be able to have the potential capacity to solve the problem of navigating from any one point in functional state space to any other. HCFM hypothesizes that systems have general problem-solving ability in any given domain when their navigation in functional state space is governed by a pattern of dynamical stability in a “fitness space” that describes their fitness to execute all of their functions, where this pattern of dynamical stability allows them to sustain their navigation through functional state space until they solve the problem [1].

In terms of information technology, any functional state space potentially defines a single universal model of all the different information systems in each computing domain, so that groups might decompose all such information systems into a library

of implementations of functions within that domain. A General Collective Intelligence or GCI [8] is a hypothetical platform that leverages one or more functional state spaces to enable groups to self-assemble into massive self-sustaining networks of cooperation that increase the groups ability to solve any problem that can be represented as a problem of navigation in functional state space.

The decoupling of software into a library of functions with well-defined interfaces is intended to in turn enable groups to self-assemble into massive self-sustaining networks of cooperation that might be able to use those functions to self-assemble software platforms at radically greater speed and scale than possible today. In addition, software platforms that self-assemble in this decentralized way are predicted to also have the capacity to solve collective challenges that are otherwise not reliably solvable today due to a misalignment of incentives with the centralized owners of current platforms. These areas of misalignment represent gaps in the functional state space that the groups cannot reach without GCI. These gaps contain problems that the group cannot define and contain solutions that the groups cannot discover without GCI [10].

To understand how might might be possible to facilitate the self-assembly of a software platform through a decentralized approach consider an individual wanting to achieve some outcome “j”. This user might choose from a menu of interaction processes involving a number of roles, where this process targets that desired outcome (Fig. 1).

This user then might use a General Collective Intelligence platform to connect with individuals in their network who would like to interact through those processes, who might use GCI to connect with individuals in their networks who might like to do the same, and so on, until the network self-assembles into a platform that does exactly what the network of individuals wants it to do.

To begin modeling this scenario, consider that each user has access to a set of decentralized functionality which they might use (Fig. 2).

These users connect through some algorithm to form a network. This network might be used to determine which other users share the same outcome as their targeted goal (Fig. 3).

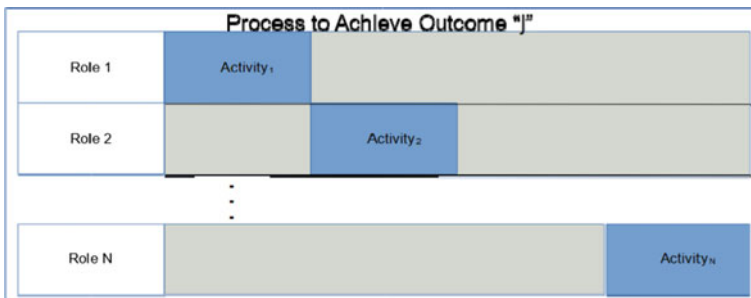


Fig. 1 Multiple users cooperate to execute a process, each having a specific role

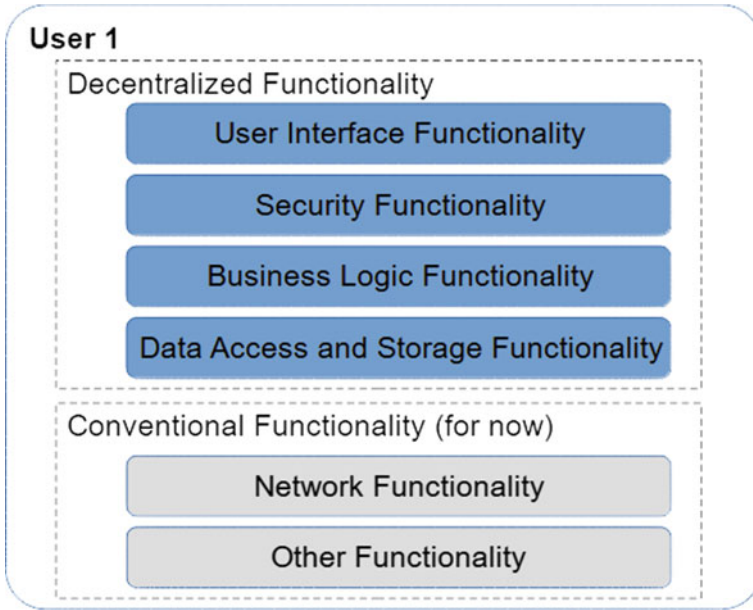


Fig. 2 Decentralized user functionality and other conventional (either centralized or decentralized at a level less than required by GCI) functionality

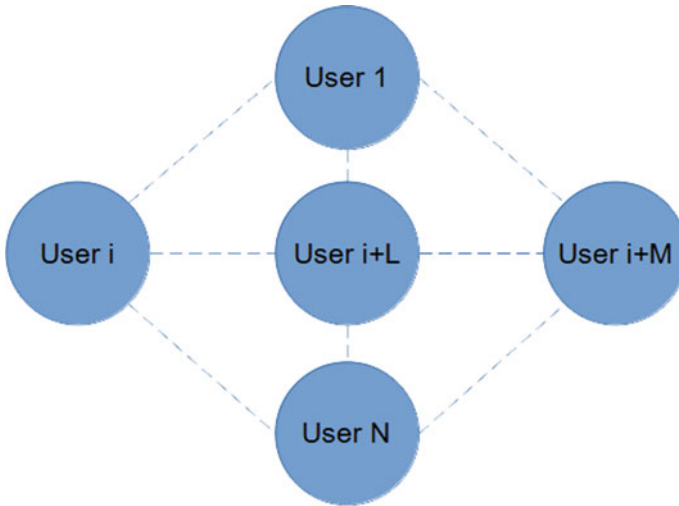


Fig. 3 A number of users are organized through General Collective Intelligence to spontaneously connect into a network of users

Each user might query their network to determine which other users are interested in participating in the platform to achieve the common targeted goal (Fig. 4).

All functionality implemented by any platform operating in that domain might potentially be decoupled into a library of functions (Fig. 5) that might be leveraged by users to achieve that shared goal.

The platform will be limited to the behavior represented by the proportion of the functional state space that has been modeled and implemented as a library function (Fig. 6).

The rules enabling this self-assembly on a self-sustaining basis are simple to understand but extremely complex to implement. These rule are that the value created by each new collaborative activity must be projected so that any activity proposed by any individual in the network might be accommodated, with a portion of that projected value being used to incentivize others to recruit for the activity as well as a portion being used to incentivize others to participate in the activity itself in order to create that value as in Fig. 7.

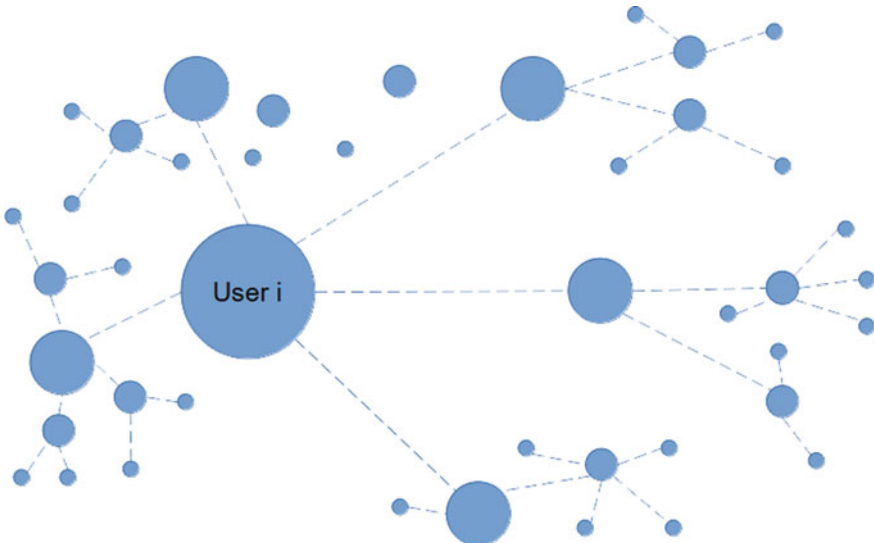


Fig. 4 Each user in the network of user “i” that is aligned in sharing the common goal of achieving outcome “j” will query their network in turn for additional individuals targeting the same goal. These will in turn query their network for additional individuals targeting the same goal, until a potentially massive network of users sharing the same goal is amassed

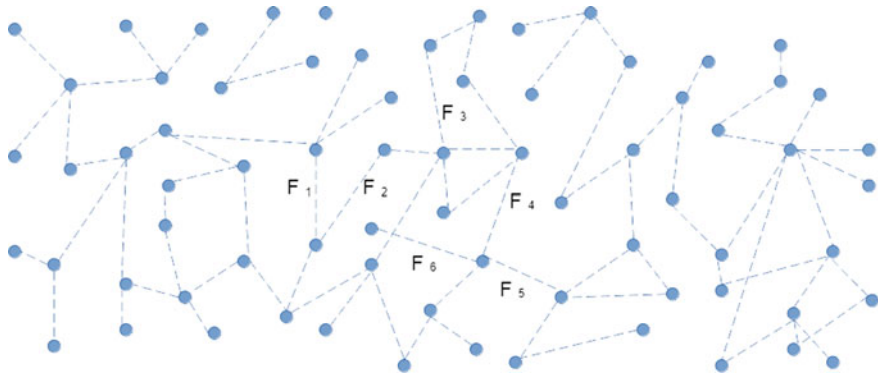


Fig. 5 Each functional state space describes a domain of behavior that the system is capable of. One domain might be user interface behavior, another might be security behavior, another might be data access and storage behavior, etc. If the functional state space describes all possible behavior in a given domain, then all functionality in that domain from any platform can potentially be decoupled into a library of functions. Each of these functions will represent a path in this functional state space. This path serves as the “unique key” identifying the function. If there are many different implementations of this function they can be compared according to this unique key to determine which is the most fit in terms of achieving the targeted outcome in a given context. Here the domain relevant to some outcome “j” is depicted.

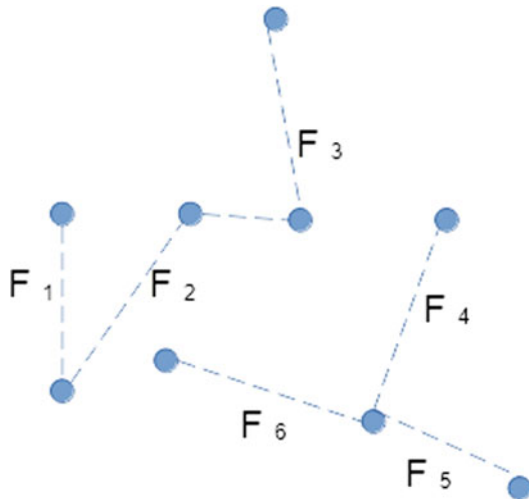


Fig. 6 In a domain relevant to an outcome “j”, it might be true that only a few of the functions have actually been modeled and/or implemented

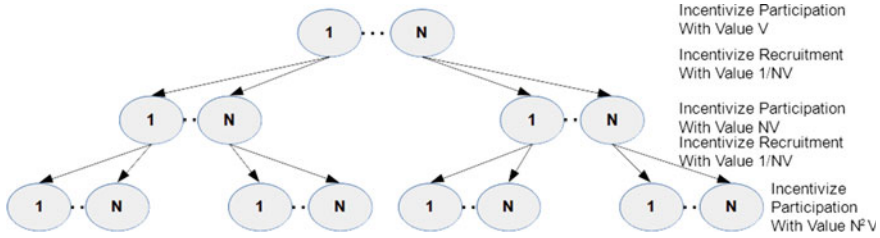


Fig. 7 An algorithm sets incentives for recruitment of others for participation and incentive for actual participation so that a viral increase in participation is reliably achievable

4 Cognitive Applications and Collective Cognitive Platforms

Intelligent agents based on some subset of functionality required for Artificial General Intelligence or AGI might create the potential to explore the fitness of every combination of different technologies in implementing every computing operation, and also might create the potential to explore interactions between computation operations that are much more complex than currently possible for humans to navigate, and to do so at vastly greater speed and scale, in order to achieve vastly greater impact on any targeted outcomes of computation. GCI (General Collective Information) as a platform [8, 9] creates the potential to collectively store information about which combination of different technologies is best suited to implementing every computing operation, so that computing operations executed by any one individual benefit from intelligence gained from the execution of any computing operation by any other individual.

5 General Collective Intelligence and Social Technology

In the case of social technology, centralized approaches have resulted in an unprecedented power to disseminate and censor information now being in the hands of a few tech giants able to manipulate the masses to serve their own agendas. This observation is in line with the hypothesis that without GCI the stable balance is for entities to compete for resources to increase their ability to compete further, and in line with the hypothesis that technology removes natural limits to the accumulation of resources and the power which accompanies them. This hypothesized natural fall of civilizations towards greater and greater centralization of power and inequality due to technological advance in the absence of General Collective Intelligence has been called the “technology gravity well”. Along with this technological advance and centralization of resources and power must necessarily come an increase in ability to censor, in ability to conduct privacy breaching surveillance, and in ability to commit any number of abuses that haven’t yet even been discovered.

This is a collective optimization problem. Decentralized, self-organizing processes are required in order to have the capacity to solve collective optimization problems. Looking at social media as an example, a social media platform might provide its users with a set of functionality decided by the social media platform owner. This functionality which the platform owner decides to provide for those users can't be separated from being aligned with his interests. That owner might decide to limit what users can do or whether they can even join the platform. In a decentralized approach, an individual wanting to interact socially might choose from a menu of social interaction processes, and then might use a General Collective Intelligence platform to connect with individuals in their network who would like to interact through those processes. These individuals might use GCI to connect with individuals in their networks who might like to do the same, and so on, until the network self-assembles into a social media application that does exactly what the network of individuals wants it to do. With intelligent agents working on each individual's behalf, this self-assembly might recreate a Facebook or a twitter with an even greater user base in microseconds. General Collective Intelligence is essentially a platform that provides this ability to self-assemble to execute any process in this decentralized way. This decentralization makes it possible to achieve "collective optimization", in which some optimal collective outcome, like ability to assemble a network for a given purpose through social media, is achieved by optimizing outcomes for every person individually. Because GCI is potentially the only system providing a model of pervasive decentralization, GCI might be the only known intervention with the potential capacity to reliably combat the hegemony of big tech companies.

6 Related Work

In cybernetics, the concept of a noosphere similar to the concept of a functional state space, and the concept of a global brain similar to the model of General Collective Intelligence (GCI) discussed in this paper, have been described by other authors [12]. In philosophy, there is an established tradition of using semantic models for representing information [13]. Bas C. van Fraassen has been attributed credit [14] for the concept of the "state space" as a semantic modeling approach, and a number of researchers citing his work have taken the concept of "state spaces" and applied them to all of science in general, including to physics and to modeling the physical world [15].

As mentioned, the difference is that this model of GCI hypothesizes a functional definition for general problem-solving ability in groups (collective intelligence), and how that collective intelligence might be exponentially increased through a GCI platform. Without the ability of a group to leverage a platform to achieve this exponential increase in its general collective intelligence factor, and without the resulting ability to achieve an exponential increase in impact on collective problems from poverty, to health care, to climate change, or even "wicked problems" in basic

research, it is predicted that the increase in impact required to reliably solve such problems can't reliably be achieved.

7 Proposed Future Work

Proposed future work includes exploring how GCI might be used to implement a blockchain platform and cryptocurrency able to leverage self-assembling cooperation that is financially self-sustaining. The concept for such a hypothetical platform that enables the free flow of value to serve any need, rather than the flow of value only where permitted by centralized interests, has been called the "Internet of value" [16]. Work to implement such a "cognitive blockchain" platform is expected to include exploring how architecture level services such as network connectivity, security, scalability, and decentralization, might be self-assembled from a library of functional components by such a platform.

8 Discussion/Analysis

In the absence of GCI, any current group suffers from systematic errors in decision-making that limit its collective intelligence. As an example, what is commonly lost in discussions about what COVID policy is best is that there are two main cognitive biases individuals might have (type 1 or intuitive reasoning and type 2 or rational methodical reasoning), neither of which permits an understanding of the other. Both biases are useful however, but without a platform such as GCI to help select the optimal bias in each context we cannot as groups switch between these biases in an optimal way, the same way we can switch between these biases within our individual cognition. Furthermore, we lack a comprehensive model able to take into account all possible factors affecting the well-being of the population so that more intelligent decisions can actually be made [11].

A GCI in effect organizes groups to act like a single collective organism capable of radically more complex cooperation that is predicted to exponentially increase the general problem-solving ability (intelligence) of groups, and therefore the ability of the group to solve any problem in general. One might argue that this translates to an exponential increase in ability to impact any collective outcome with respect to any given measure, such as achieving an exponential increase in collective outcomes per program dollar.

A General Collective Intelligence based social media platform is predicted to have the capacity to form and evolve exponentially more quickly than current social media platforms. This potential pace of evolution requires fundamentally different thinking about the regulation of social media. Even without GCI the "technology gravity well" hypothesis suggests that legislation can never solve the problem with social media platforms. Legislation can at most assuage the public angst that "something

be done”. However, technology always moves more quickly than legislation. This is more clear in the case of GCI based platforms since there is clearly no possibility that any current legislative bodies could ever create legislation to manage such platforms as they would move too quickly. Luckily GCI can potentially facilitate management of such platforms even if they might be evolving at exponentially greater rates and on exponentially larger scales, since evolution would be governed by a pattern of dynamical stability that remains in balance with collective well-being regardless of the speed and scale of activities.

Each of the designs of such platforms that are currently in the works are based on a limited ontology describing their domain of behavior rather than a complete functional state space. These platforms still however are predicted to increase impact on certain outcomes by orders of magnitude while also radically increasing development speed and scale [9]. To fully implement this concept and to achieve the full magnitude of their predicted benefits such platforms must evolve from being based on an ontology capable of describing a limited number of behaviors, to being based on a functional state space capable of describing all possible behaviors.

9 Conclusions

General Collective Intelligence as a platform is intended to enable groups to exponentially increase impact on any targeted collective challenge through leveraging potentially massive networks of cooperation that self-assemble in a self-sustaining way to achieve that purpose. When directed to social technologies GCI has the potential to address a number of barriers to more effectively achieving the outcomes targeted through use of those technologies.

References

1. Williams AE (2021) General problem-solving ability in natural systems as a model for computation, June 13. <https://doi.org/10.31730/osf.io/5x3g7>
2. Williams AE (2020) A model for artificial general intelligence. In: Goertzel B, Panov A, Potapov A, Yampolskiy R (eds) Artificial general intelligence. AGI 2020. Lecture notes in computer science, vol 12177. Springer, Cham. https://doi.org/10.1007/978-3-030-52152-3_38
3. Franchi E, Tomaiuolo M (2013) Distributed social platforms for confidentiality and resilience. In: Social network engineering for secure web data and services. IGI Global, pp 114–136
4. de Filippi P, Lavyssi ere X (2020) Blockchain technology: toward a decentralized governance of digital Platforms? In: The great awakening: new modes of life amidst capitalist ruins. Punctum Book, 978-1-953035-08-0 (hal-03098502)
5. Atzori M (2015) Blockchain technology and decentralized governance: is the state still necessary? December 1, 2015. Available at SSRN: <https://ssrn.com/abstract=2709713> or <https://doi.org/10.2139/ssrn.2709713>
6. Chen Y, Pereira I, Patel PC (2021) Decentralized governance of digital platforms. *J Manag* 47(5):1305–1337. <https://doi.org/10.1177/0149206320916755>

7. Vergne J (2020) Decentralized vs. distributed organization: blockchain, machine learning and the future of the digital platform. *organization theory*, October 2020. <https://doi.org/10.1177/2631787720977052>
8. Williams AE (2021) Defining a continuum from individual, to swarm, to collective intelligence, and to general collective intelligence. *Int. J. Collaborative Intell.* 2(3):205–209
9. Williams AE (2021) Approximating an artificial general intelligence or a general collective intelligence. *Int J Collaborat Intell*, in print
10. Williams AE (2022) Is Sustainability a pattern of solution in a functional state space?, February 1 Retrieved from osf.io/preprints/africarxiv/cb2xh
11. Henriques G, Williams AE (2021) General collective intelligence and systemic error in groups. <https://www.psychologytoday.com/us/blog/theory-knowledge/202106/general-collective-intelligence-and-systemic-error-in-groups>. Accessed 19 July 2021
12. Beigi S, Heylighen F (Accepted/In press). Collective consciousness supported by the web: healthy or toxic? In: *Computational collective intelligence: 13th international conference, ICCCI 2021, Proceedings (LNAI)*. Springer
13. Suppe F (1989) *The semantic conception of theories and scientific realism*. University of Illinois Press
14. Lloyd EA (1984) A semantic approach to the structure of population genetics. *Philos Sci* 51(2):242–264
15. Niiniluoto I (1987) Quantities, state spaces, and laws. In: *Truthlikeness*. synthese library (Studies in epistemology, logic, methodology, and philosophy of science), vol 185. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-3739-0_3
16. Williams AE, Visconti RM (2020) The application of artificial general intelligence to the cognitive blockchain and the internet of value

Risk Analysis Related to the Corrosion of Atmospheric Storage Tanks: A Multimethodological Approach Using the FRAM and the ELECTRE-MOR Methods



Érick Pinto Moreira, Igor Pinheiro de Araújo Costa,
Fernando Benedicto Maimier, Carlos Francisco Simões Gomes,
and Marcos dos Santos

Abstract The storage activity of oil and its derivatives is fundamental in the production chain of the O&G (Oil and Gas) industry, in which ASTs (Aboveground Storage Tanks), commonly constructed using carbon steel materials, are used for this purpose. During the operational life of these assets, corrosion is a factor that, if not controlled, generates risks for industrial environments where the tanks are present. Therefore, this work aims to propose a new methodology for assessing risks related to atmospheric corrosion in storage tanks, by applying this framework in ASTs located in a tank farm and sorting the tanks into risk classes. A multi-methodological approach was adopted, in which the FRAM was used to structure the problem and with the literature review, they served as the basis for the definition of evaluation criteria. The ELECTRE-MOR method was applied, using collected data referring to the defined criteria, and using the ELECTRE-MOR WEB software the tanks were classified into three risk classes. It was observed that the method was consistent in pointing out the tanks with the most critical criteria in the “high risk” class, with one special surprise in pointing out a tank that apparently did not need attention in this same class. The application proves to be feasible and can be improved and used as a basis for intervention campaigns in these assets.

Keywords Risk analysis · Storage tanks · Atmospheric corrosion · FRAM · Resilience engineering · ELECTRE-Mor · MCDM methods

É. P. Moreira (✉) · I. P. de A. Costa · F. B. Maimier · C. F. S. Gomes
Fluminense Federal University (UFF), Niterói, RJ 24210-346, Brazil
e-mail: erickmoreira@id.uff.br

M. dos Santos
Military Institute of Engineering (IME), Urca, RJ 22290-270, Brazil

1 Introduction

Corrosion is a common factor in the daily life of society and, when considered its influence on the environment from the perspective of health, safety, and the environment, as well as under economic bias, can be considered a problem, as it generates losses [1].

If we consider the data addressed by Koch et al., which estimates the global cost related to corrosion at 3.4% of world Gross Domestic Product (GDP) [2], it is possible to estimate these costs in Brazil, multiplying this percentage by the Brazilian GDP of 7.4 trillion BRL (Brazilian Real—Currency) according to Brazilian Institute of Geography and Statistics (IBGE) [3], which results in 251.6 billion BRL, where it is important noting that these values are not considered the costs arising from accidental consequences with people and the environment.

Thus, the study of corrosion and its relationships with causal factors of defects and/or operational failures, as well as the correlation between these factors and criticality conditions, are important for the development of solutions and actions that mitigate and/or prevent their occurrence and more serious consequences that are possible, such as accidents, also considering the financial impact of these actions.

In this context, the risk of corrosion-related accidents in storage tanks is an important point to be considered, as demonstrated in the study prepared by Chang and Lin, where 242 tank accidents were analyzed, which the results indicated that approximately 26% of these occurred in storage terminals and approximately 7% had the consequence of cracks, fractures or ruptures of the tank, which may have its main root cause related to corrosion [4].

Considering what Topalis addresses, regarding the importance of avoiding the waste and to ensure directing the resources and efforts to critical assets, it is evident the need for evolution in the risk management processes of these enterprises [5], and which are being addressed by world-renowned organizations on the subject of risk-based inspection (RBI), such as the American Petroleum Institute (API), the American Society of Mechanical Engineers (ASME), the Engineering Equipment and Materials Users Association (EEMUA) and the European Committee for Standardization (CEN), among other organizations.

With this into account, the use of a structured methodologies to aid decision making, considering the factors and conditions studied in the research, is an opportunity for the proposition of the technique in a context not yet applied and as an alternative to directing solutions and actions of prevention in an objective way, reducing the subjectivity of decisions taken by the technical bodies of the organizations.

Therefore, this work aims to propose a new framework for assessing risks related to atmospheric corrosion in storage tanks, and not to cover limitations of existing methods and/or systems. The methodology consists in using the Multi-criteria Decision Analysis (MCDA) approach, using the ELECTRE-MOr (ELimination Et Choix Traduisant la REalité—Multicriteria Sorting Method with Ordinal Weight Input and Multiple Decision Makers), to generate a risk classification of these industrial assets. As a way to strengthen the proposed methodology, the FRAM (Functional Resonance

Analysis Method) will be applied, which is based on the need to better understand the performance of complex socio-technical context systems [6].

2 Literature Review

2.1 Aboveground Storage Tanks

In the oil and gas industry, after the exploration and production of hydrocarbons, the oil produced is transported to the continent and refined. At some point during these phases, oil and its derivatives need to be stored, thus making this an essential activity for the oil industry's production cycle.

It is common for storage to take place in atmospheric storage tanks, equipment present in tank parks in terminals, refineries, distribution centers, and industrial plants [7].

In Brazil, according to the publication of the Technical Regulation of Terminals for The Movement and Storage of Petroleum, Derivatives, Natural Gas and Biofuels (RTT) by the National Agency of Petroleum, Natural Gas and Biofuels (ANP), in 2019, Brazil had 110 authorized terminals, 61 waterway terminals, and 49 land terminals, which together totaled 2014 tanks [8].

This equipment, for the most part, is constructed using carbon steel and one of the major problems faced by them during their life cycle is corrosion, which affects the bottom, the shell, and the roof plates, internally and externally, and can reach critical levels of thickness or even the complete loss of material at certain points, coming to stick.

To mitigate these risks, throughout the operation of this equipment, periodic inspections are necessary, usually derived from normative recommendations and, as a consequence of these inspections, when non-compliant parameters are identified, preventive and/or corrective interventions are necessary.

However, it happens that, depending on the type of intervention recommended, it is necessary to remove the service tank for this to occur, an unwanted fact due to the loss of storage and the costs involved. However, the non-performance of interventions causes risks to the operation of the equipment, which may result in accidents.

2.2 Atmospheric Corrosion

As approached by Popov and Leygraf et al., the exposure of metallic materials to different types of the atmosphere is what makes characteristic the form of corrosion known as atmospheric, which occurs by the spontaneous interaction of the material with the environment, in this case, the atmosphere, and varies according to the characteristics of this, such as temperature, relative humidity, time of exposure of the

Table 1 Categories of atmospheric corrosiveness

| Category | Corrosiveness | Category | Corrosiveness |
|----------|---------------|----------|---------------|
| C1 | Very low | C4 | High |
| C2 | Low | C5 | Very high |
| C3 | Medium | CX | Extreme |

Source ISO, 2012 [12]

surface to electrolyte film, the pH of the electrolyte and the presence of contaminants such as sodium chloride—NaCl—for example, which is characteristic of marine and coastal atmospheres [9, 10], while Gentil and Mainier, Guimarães and Merçon add that local climatic conditions should also be considered, such as wind intensity and direction, prevailing winds, cyclic variations in temperature and humidity, salinity, the incidence of rainfall and ultraviolet radiation (insolation) [1, 11].

International Organization for Standardization (ISO) classifies atmospheres into 6 categories of corrosiveness, according to Table 1.

The location of the terminal addressed in this study, because of its characteristics, from which one can highlight the industrial, coastal environment, high humidity, and salinity, of tropical climate, places the classification of corrosiveness of its atmosphere in category “C5” or “CX” [12].

2.3 FRAM

The FRAM method, as Hollnagel describes, assumes that there is not necessarily a causal relationship to the events. Unlike some methodological approaches that assume the linear, complex, and dynamic reasoning of technological systems, the FRAM, based on the principles of resilience engineering, considers the non-technical links of the systems, which are represented mainly by the human factors present in these, dealing with the relationship between the technical parts and the means to which they are inserted and considering their possible variability and reverberations, characteristic factors of complex socio-technical systems [13].

Its use for risk and security management purposes is broad and focused on industrial scenarios, as addressed by Gattola et al., based on four principles, which are [14]:

- Equivalence of failures/successes;
- Emergency;
- Approximate adjustments;
- Functional resonance.

For its application, it is important to be clear in the definition of objectives, because it is considered that it can be used both for accident investigation purposes, as well as for risk analysis and evaluation, which is the purpose of application in this research [13]. The methodological process goes through four steps that are:

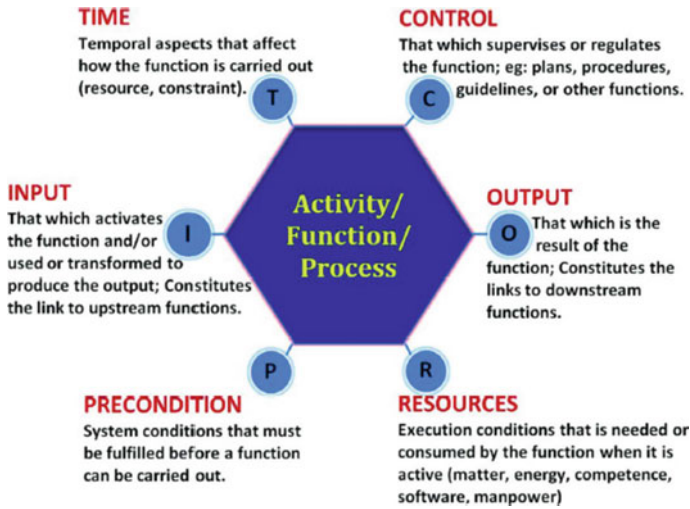


Fig. 2 Representation of an activity by the FRAM method (Source Thomas et al. [15])

1. Identify and describe the activities;
2. Identify variability;
3. Aggregate variability;
4. Develop recommendations.

The activities modeled using the FRAM are described considering 6 aspects, as shown in Fig. 2:

The whole complex system will be modeled by broken it into activities that are represented by the 6 aspects shown in Fig. 2 to support the understanding of the interactions between these activities and to help the identifications of the weakness points of the system.

As a facilitator of its application, we used the Software FRAM Model Visualizer (FMV), developed by Hill, to support the modeling of the system under study and the application of the process steps [16].

2.4 ELECTRE-MOr

Moreira et al. mention that the Multi-criteria Decision-making (MCDM) methods could be inferred as techniques to structure and analyse complex evaluation problems with more clarity making use of quantitative and qualitative criteria, allowing decision-makers (DM) have trade-offs options between them [17], these methods derive from OR, a multidisciplinary subject that adopts mathematical and logical

frameworks to solve real problems of the society, and is an area of Industrial Engineering [18, 19], so regarding this, MCDM methods become a relevant mechanism for managerial areas of the organizations [20].

The MCDM method applied will be ELECTRE-MOr. This method is particularly interesting for the problem presented because it allows the evaluation of alternatives based on multiple criteria with ordinal weight input, for separating the alternatives into predetermined classes, because it is multidecision and because it is a non-compensatory method.

For the application of the method, it is necessary to define the alternatives under evaluation, the criteria are established, the data are obtained for the established criteria related to each defined alternative, the preference relationships for each criterion are determined, the weights of the criteria are assigned and the overclassification relationship indicated.

The alternatives considered in this study will be the tanks with a volume up to 500 m³, located in the storage terminal covered, totalizing 14 tanks, and the criteria will be established based on bibliographic research and the application of the FRAM, to classify the risks related to atmospheric corrosion.

To determine the preference relationships, the method establishes three fundamental factors for comparing each criterion between the alternatives and the predefined class limits:

- Weak preference (q): There are clear and positive reasons that imply a preference in favor of one of the two alternatives, but these reasons are insufficient to assume a strict (strong) preference in favor of this [21];
- Strong preference (p): There are clear and positive reasons that justify a significant preference in favor of one of the two alternatives [21];
- Veto (v): Limit that sets a different value for each criterion, from which the overclassification between the alternatives will not be accepted [22].

In the allocation of weights, an adaptation of the SAPEVO method (Simple Aggregation of Preferences Expressed by Ordinal Vectors) is used, which requires decision-makers to make a peer comparison between the criteria [23, 24], resulting in the generation of weights.

To indicate the overclassification relationship, the cut-off level (λ) is adopted, which indicates whether or not credibility is sufficient to accept the overclassification [25]. λ is arbitrarily assigned, ranging from 0.5 (medium) to 1.0 (one), assuming the more demanding condition the closer to 1.0 it is.

The axiomatics of the method, as well as its mathematical modeling, can be found in the articles proposed by Costa et al., which applied it to classify hospital aircraft to be used in the fight against the COVID-19 pandemic, and to support the strategic process of portfolio formation of courses offered by a company in the Information Technology (IT) training sector that should be prioritized, maintained, or discarded by the company's management by classifying them [26, 27].

According to Cinneli et al., the provision of a computational tool to support the application of an MCDA method, which will perform the axiomatic and mathematical development of this, for the presentation of results and graphic representation, is

a relevant characteristic for the method to be disseminated and used, so it is emphasized that for practical application [28], in this research, the ELECTRE-MOr WEB software [29] developed by Costa et al. was used.

3 Methodology

This article comprises a methodological flow in which we seek to show a problem situation, with a relevant theme for society, which begins with exploratory research through the contextualization of the scenario, where the object of study is presented and its specificities and location are delimited, as shown in Sect. 1.

With the understanding of the problem, we advance to the descriptive research, which seeks to deepen the knowledge of the research theme related to the object of study, through a review of the available literature, which discusses storage tanks, atmospheric corrosion, FRAM method, and ELECTRE-MOr method, as presented in Sect. 2.

With the knowledge about the theme of research and the object of study having been more detailed, we will seek to raise the critical criteria that, based on the theoretical references and aided by the FRAM, are more adherent to the problem.

The data, related to the established criteria, will be obtained through field visits and the survey of historical records of assets of this type, present in a storage terminal, being limited to analyzing the volume tanks up to 500 m³.

The modeling and application of the MCDA methodology, under the theoretical framework of Operational Research (OR), will take place through the use of the ELECTRE-MOr, to distribute the tanks in three classes, respectively “High Risk”, “Medium Risk” and “Low Risk”, all these previous steps was presented in Sect. 4.

The analysis of the results, shown in Sect. 5, will be based on the understanding of the risk classification presented after the modeling and application of the ELECTRE-Mor.

To summarize the steps of this methodological research, a methodological flowchart is presented in Fig. 1.

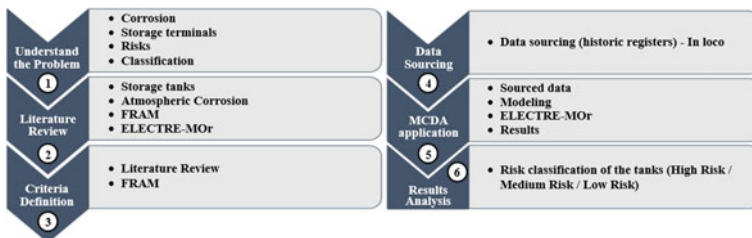


Fig. 1 Methodological flowchart

4 Case Study

To structure the classification problem, using the ELECTRE-MOR, it is necessary to define the alternatives that will be classified and the criteria under which these alternatives will be evaluated. In this study, for the classification of risks related to atmospheric corrosion of oil storage tanks and derivatives, tanks with a volume smaller than 500 m³ in operation in a storage terminal in a coastal area, totaling 14 tanks, were defined as alternatives. To select the criteria, in addition to the theoretical basis obtained by the bibliographic survey, the FRAM method was applied as a way to support the structuring of the problem.

In this study, the FRAM was applied to describe the terminal activities related to the management of tank integrity and the activities that interrelate with them and may generate positive or negative influence on management, as well as their potential sources of variability (Fig. 3).

The activities represented with a gray edge represent the necessary construction of systematic control of the tanks, which is fundamental when it is intended to manage any processes, where the product of this system is the precondition necessary for the integrity management process (orange edge), which is directly linked to inspection activities (green edge), which function as the connection link between management and the processes of operation (blue edge) and maintenance (yellow edge).

The connections between these processes and activities are represented by the lines that connect the aspects of each activity to the other with which it relates. The variability of the system is represented by activities that have a sine wave in its center,

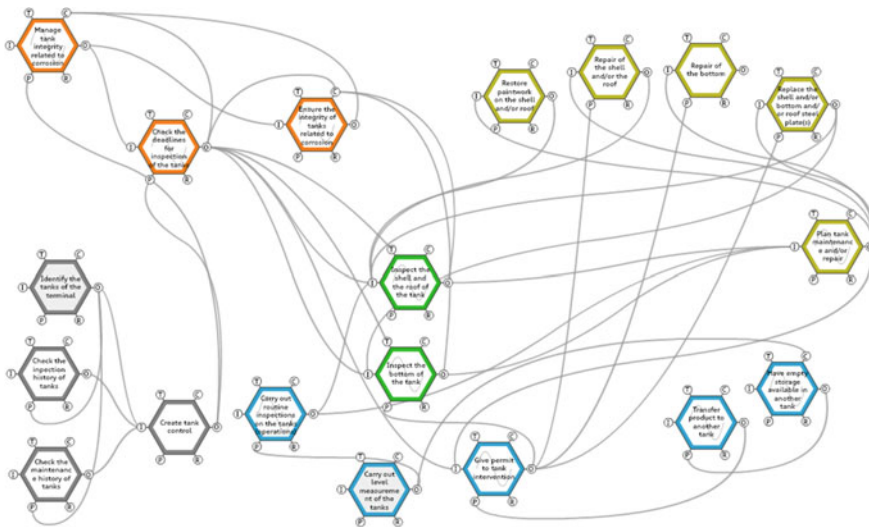


Fig. 3 Representation of the terminal as a socio-technical system by the FRAM method

Table 2 Evaluation criteria

| # | Criteria | Type | Unity | Monotonic |
|---|-------------------------|--------------|--------------------------|-----------|
| 1 | Age of operation | Quantitative | Years | Cost |
| 2 | Roof area | Quantitative | m ² | Cost |
| 3 | Shell area | Quantitative | m ² | Cost |
| 4 | Visual roof evaluation | Nominal | (1_Bad/2_Regular/3_Good) | Profit |
| 5 | Visual shell evaluation | Nominal | (1_Bad/2_Regular/3_Good) | Profit |
| 6 | Inspection expired? | Binary | (0_No/1_Yes) | Cost |

these variabilities so may be related to human, organizational, or technological factors [13].

It can be noticed that the inspection and maintenance activities are crucial to keep the integrity of these assets in a acceptable state, however they could be hugely affected by the operations activity as they are fundamental to allow the inspection and maintenance occur regularly as necessary.

With the evaluation of all these interrelationships between the processes related to the management of tank integrity, observing that 10 of the 19 activities listed have some kind of variability, and still considering what was raised in the literature review, it was suggested the evaluation of alternatives under the following criteria (Table 2).

Criteria number 1, 2, and 3 are defined by the understanding of the literature review, as atmospheric corrosion primarily attack the external side of the steel plates of the tanks and if the tank has huger areas of these plates exposed to the corrosion the we could infer more risk to atmospheric corrosion attack on that, also corrosion is a timely based action mechanism so the age of each tank is an indicator of the exposure time of the tank to atmospheric corrosion. Criteria number 4, 5, and 6 are defined by the understanding of the importance of the inspection and maintenance activities that could be affected and prejudiced by the operations activities, addressing the importance to have the inspections taken regularly and not becoming expired and to have minimally external visual inspections carried out to give aid to the risk understanding on each tank.

The classification of the criteria as monotonic cost and profit allude to the expected objectives for them, which brings reference to the intention to maximize the monotonic criteria of profit and to minimize the monotonic criteria of cost. To obtain the data related to these criteria, the survey was carried out through field visits, for criteria 4 and 5, and researched in historical records, for criteria 1, 2, 3, and 6, considering tanks with a volume smaller than 500 m³, which resulted in 14 tanks, described as letters of the alphabet, from “A” to “N”, for confidentiality reasons (Table 3).

The preference relationships (Table 4) were assigned considering the data ranges that each criterion presented, also based on the experience of the authors, who also participated in the joint evaluations between the criteria, a step that uses an adaptation of the SAPEVO method, contributing to the definition of the weights of the criteria.

Three classes were defined for the distribution of alternatives, which are from the best to the worst, respectively, “low risk”, “medium risk” and “high risk”. For the

Table 3 Data for evaluation

| Alternatives | Criteria | | | | | |
|--------------|----------|-------|--------|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| A | 81 | 88.62 | 120.14 | 2 | 2 | 0 |
| B | 61 | 45.58 | 169.93 | 3 | 3 | 0 |
| C | 61 | 45.61 | 168.54 | 1 | 1 | 1 |
| D | 61 | 45.78 | 176.04 | 3 | 3 | 0 |
| E | 61 | 45.50 | 173.12 | 1 | 1 | 1 |
| F | 81 | 36.31 | 108.73 | 2 | 3 | 0 |
| G | 81 | 36.35 | 109.00 | 2 | 1 | 1 |
| H | 51 | 36.09 | 182.92 | 1 | 1 | 1 |
| I | 41 | 21.61 | 143.69 | 1 | 2 | 1 |
| J | 71 | 36.17 | 183.13 | 1 | 1 | 1 |
| K | 71 | 21.58 | 142.10 | 3 | 2 | 0 |
| L | 45 | 12.53 | 58.96 | 2 | 2 | 1 |
| M | 45 | 12.53 | 58.71 | 2 | 2 | 1 |
| N | 45 | 12.53 | 89.07 | 2 | 2 | 1 |

Table 4 Factors of preference relationships and weights of the criteria

| Alternatives | Criteria | | | | | |
|-----------------------|----------|-------|--------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Weak preference (q) | 5 | 10.00 | 30.00 | 0.50 | 0.50 | 0.10 |
| Strong preference (p) | 10 | 20.00 | 50.00 | 1.00 | 1.00 | 0.50 |
| Veto (v) | 20 | 30.00 | 100.00 | 2.00 | 2.00 | 2.00 |
| Weight | 0.01 | 0.55 | 0.09 | 1.00 | 0.45 | 0.73 |

purpose of the research, the minimum λ of 0.5 was considered. As the application proposal is to classify the tanks into risk classes, using the minimum λ it can be assumed that a smaller number of tanks will be in the worst classes, corresponding to tanks with high risk, where it is intended to reduce the number of tanks that will focus for a more careful evaluation.

After the application of ELECTRE-MOr, through the computational tool, the characteristic classifications of the method were obtained, which considers two forms of distribution in two scenarios. Based on the experience of the authors and the purpose of the classification, the aggregation was performed between the 4 classification combinations generated, which considered the largest number of times that the best class was attributed to the alternative and, in case of a tie, it was considered the best class, to reduce the alternatives classified in the worst class. Table 5 shows the result of the already aggregated classification.

Table 5 Classification of alternatives

| Alternatives | Classification | Alternatives | Classification |
|--------------|----------------|--------------|----------------|
| A | High risk | H | High risk |
| B | Low risk | I | Medium risk |
| C | High risk | J | High risk |
| D | Low risk | K | Medium risk |
| E | High risk | L | Low risk |
| F | Medium risk | M | Low risk |
| G | High risk | N | Low risk |

The “High Risk” class represent the tanks that has more probability to have any integrity problem related to atmospheric corrosion as the “Low Risk” class represent the tanks with less probability.

When analyzing the results, crossing the classifications with the data mass, it is possible to notice that the tanks categorized as “High Risk” have common characteristics such as a long time in operation, larger areas of plates exposed to the atmosphere, results of bad visual evaluations and are with the inspection overdue. However, a particular case (Tank A) draws attention, since the tank is up to the inspection and visual inspection evaluations considered as regular, but negatively opposes these positive points the fact that it is one of the largest tanks, with the longest operating time.

5 Conclusions

It can be concluded, after analyzing the results, that the use of the ELECTRE-MOr method was effective when classifying the tanks since the assets with more critical characteristics from the point of view of the adopted criteria were highlighted in the categories of “High Risk” or “Medium Risk”. This classification can be used, as example, for the planning of inspection campaigns, to validate the evaluation performed. The use of the FRAM method proved interesting for understanding the scenario, bringing more clarity to the complex context of the terminal and the potential sources of variability that can bring disturbances to the system.

References

1. Gentil V (2011) Corrosão. Grupo Gen-LTC
2. Koch G, Varney J, Thompson N, Moghissi O, Gould M, Payer J (2016) International measures of prevention, application, and economics of corrosion technologies (IMPACT) Study. NACE International
3. IBGE. <https://www.ibge.gov.br/explica/pib.php>. Accessed 18 Apr 2021
4. Chang JI, Lin C (2006) A study of storage tank accidents. *J Loss Prev Process Ind* 19:51–59

5. Topalis P (2017) Risk based inspection methodology for atmospheric storage tanks. DNV GL
6. Patriarca R, Di Gravio G, Woltjer R, Costantino F, Praetorius G, Hollnagel E (2020) Framing the FRAM: a literature review on the functional resonance analysis method. *Saf Sci* 129:104827
7. Dantsoho AM (2015) Risk-based framework for safety management of onshore tank farm operations. PhD thesis. Liverpool John Moores University
8. ANP (2019) Publicação do Regulamento Técnico de Terminais Para Movimentação e Armazenamento de Petróleo, Derivados, Gás Natural e Biocombustíveis (RTT). Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
9. Popov BN (2015) Corrosion engineering: principles and solved problems. Elsevier
10. Leygraf C, Wallinder IO, Tidblad J, Greadel T (2016) Atmospheric corrosion. Wiley
11. Manner FB, Guimarães PIC, Merçon F (2003) O ensino de corrosão com base na avaliação crítica do mobiliário urbano de praças e logradouros públicos. COBENGE
12. International Organization for Standardization: Corrosion of metals and alloys—Corrosivity of atmospheres—Classification, determination and estimation (ISO Standard No. 9223). ISO (2012)
13. Hollnagel E (2012) FRAM, the functional resonance analysis method: modelling complex socio-technical systems. Ashgate Publishing
14. Gattola V, Patriarca R, Tomasi G, Tronci M (2018) Functional resonance in industrial operations: a case study in a manufacturing plant. *IFAC-Pap* 51:927–932
15. Thomas J, Davis A, Samuel MP (2021) Strategic quality management of aero gas turbine engines, applying functional resonance analysis method. In: Mistry C, Kumar S, Raghunandan B, Sivaramakrishna G (eds) Proceedings of the national aerospace propulsion conference. Lecture notes in mechanical engineering. Springer, Singapore, pp 65–91
16. Hill R (2021) FRAM model visualizer (v.2.1.5). <https://functionalresonance.com/the%20fram%20model%20visualiser/>. Accessed 28 June 2021
17. Moreira MAL, Gomes CFS, dos Santos M, Silva MC, Araujo JVGA (2020) PROMETHEE-SAPEVO-M1 a Hybrid modeling proposal: multicriteria evaluation of drones for use in naval warfare. In: Thomé AMT, Barbastefano RG, Scavarda LF, dos Reis JCG, Amorim MPC (eds) Proceedings in mathematics & statistics, vol 337. Industrial engineering and operations management: XXVI IJCIEOM. Springer, Switzerland, pp 381–393
18. Almeida IDP, Corriça JVP, Costa APA, Costa IPA, Maêda SMN, Gomes CFS, dos Santos M (2021) Study of the location of a second fleet for the Brazilian navy: structuring and mathematical modeling using SAPEVO-M and VIKOR methods. In: Rossit DA, Tohmé F, Mejía Delgadillo G (eds) Production research. ICPR-Americas 2020: production research. Communications in computer and information science, vol 1408. Springer, Cham, pp 113–124
19. dos Tenório FM, Santos M, Gomes CFS, Araujo JC (2020) Navy warship selection and multicriteria analysis: the THOR method supporting decision making. In: Thomé AMT, Barbastefano RG, Scavarda LF, dos Reis JCG, Amorim MPC (eds) Proceedings in mathematics & statistics, vol 337. Industrial engineering and operations management: XXVI IJCIEOM. Springer, Switzerland, pp 381–393
20. dos Santos M, Costa IPA, Gomes CFS (2021) Multicriteria decision-making in the selection of warships: a new approach to the AHP method. *Int J Anal Hierarchy Process* 13:147–169
21. Roy B, Figueira JR, Almeida-Dias J (2014) Discriminating thresholds as a tool to cope with imperfect knowledge in multiple criteria decision aiding: theoretical results and practical issues. *Omega* 43:9–20
22. Gomes LFAM, Gomes CFS (2019) Princípios e métodos para a tomada de decisão: Enfoque multicritério, Atlas
23. Gomes CFS, dos Santos M, Teixeira LFHSB, Sanseverino AM, Barcelos MRS (2020) SAPEVO-M: a group multicriteria ordinal ranking method. *Pesquisa Operacional* 40:1–23
24. Gomes L, Mury AR, Gomes CFS (1997) Multicriteria ranking with ordinal data. *Syst Anal-Model-Simul* 27:139–146
25. Rocha C, Dias LC (2008) An algorithm for ordinal sorting based on ELECTRE with categories defined by examples. *J Global Optim* 42:255–277

26. Costa IPA, Sanseverino AM, Barcelos MRS, Belderrain MCN, Gomes CFS, dos Santos M (2021) Choosing flying hospitals in the fight against the COVID-19 pandemic: structuring and modeling a complex problem using the VFT and ELECTRE-MOr methods. *IEEE Lat Am Trans* 19:1099–1106
27. Costa IPA, Moreira MAL, Costa APA, Teixeira LFHSB, Gomes CFS, dos Santos M (2021) Strategic study for managing the portfolio of IT courses offered by a corporate training company: an approach in the light of the ELECTRE-MOr multicriteria hybrid method. *Int J Inform Technol Decis Mak* 20:1–29
28. Cinelli M, Kadziński M, Gonzalez M, Słowiński R (2020) How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. *Omega* 96:102261
29. Costa IPA, Castro MAP, Gomes CFS, dos Santos M (2021) ELECTRE-MOr WEB software (v.1). <http://electremor.com/>. Accessed 28 June 2021

Implementing Braking and Acceleration Features for a Car-Following Intelligent Vehicle



Johann Carlo R. Marasigan, Gian Paolo T. Mayuga , Elmer R. Magsino , and Gerald P. Arada 

Abstract In this study, we have implemented braking and acceleration features for driver assistance in a car-following intelligent vehicle to minimize unnecessary braking and provide appropriate acceleration inputs in order to alleviate traffic buildups. Our system employs a fuzzy logic controller to process speed, obtained from an installed on-board diagnostics adapter, and distance, based from LIDAR sensors, inputs to output the necessary amount of braking or acceleration input to achieve the desired car-following targets while maintaining comfortable driving condition. Classifying driving behavior has been incorporated in the type of membership functions used to characterize the speed and distance inputs. Our hardware is easily installed in vehicles and does not require major changes to the vehicle. Finally, our experiments provide a 90% accuracy in achieving desired speed and maintaining inter-vehicular separation.

Keywords Braking and acceleration features · Driver assistance · Fuzzy logic · Intelligent vehicle · Smart city

J. C. R. Marasigan · G. P. T. Mayuga
Electronics, Computer and Communications Engineering, Ateneo de Manila University,
Quezon City, Philippines
e-mail: johann.marasigan86@gmail.com

G. P. T. Mayuga
e-mail: gmayuga@ateneo.edu

E. R. Magsino (✉) · G. P. Arada
Department of Electronics and Computer Engineering Gokongwei College of Engineering,
De La Salle University, Manila, Philippines
e-mail: elmer.magsino@dlsu.edu.ph

G. P. Arada
e-mail: gerald.arada@dlsu.edu.ph

1 Introduction

Aside from bottlenecks [2] and dynamic incidents such as phantom traffic [15] and vehicular accidents, driving behavior has played a major role in the buildup of traffic in major thoroughfares. This may be due to the experience and skills inefficiency of human drivers. Because of these driving skill limitations, cars on a single lane normally have non-uniform inter-vehicular spacing among them while running at different speeds. An abrupt halt from the car leader may result to an unwanted accident. Another problem resulting from inefficient driving behavior is the frequent changing of lanes because of slow leading vehicles [8]. Some car manufacturers have developed adaptive cruise control [13] and automatic braking [6] but a study in [5] showed that these are not enough to minimize traffic buildup.

The driving braking behavior in vehicle-bicycle conflicts were analyzed in [4] to improve existing autonomous braking emergency systems. On the other hand, Deligianni et al. [3] studied the naturalistic driving behavior in order to improve the comfort of passengers riding autonomous vehicles. Results indicated that when danger are present, drivers performed a hard braking behavior before becoming smooth. Also, speed was always a reason for braking. The same research objective was also seen in [7]. Angle sensors were placed in the throttle and brake pedals to record the driver's braking behavior on a road with four predetermined locations and initial speed. The effects of different blood alcohol concentrations on various driving conditions were examined through simulation. Drivers showed aggressive and impulsive behaviors leading to more crash probabilities [17].

In [9], the SenSafe framework utilizing smartphones is proposed. SenSafe utilized the smart phone's sensors to sense driving behavior and exchange driving information. It also reminded and alerted the users. Finally, in Wu et al. [16], an anti-rear-end collision warning device helped reduce rear-end car collisions by turning the brake lights ON in advance.

In a study done in India [1], the acceleration and deceleration characteristics of different vehicles types ranging from motorcycles to trucks were observed. It also considered the types of engines used (whether gasoline or diesel-fed). Their observations showed varying times in reaching the maximum speed of individual vehicles, even those within the same classification. This variation can also be seen in deceleration maneuvers.

One of the various methods to solve traffic buildup is through a centralized approach like data dissemination in vehicular networks [10]. However, this centralized approach requires a lot of traffic modeling, infrastructure deployment, and vehicle registrations to fully achieve a seamless traffic management system. Another approach is to allow each road user to monitor their own driving behavior and follow a set of allowable rules, e.g., car-following movements, effectively becoming decentralized and region-dependent.

In this study, we address the second approach by developing an inexpensive and easy-to-install hardware in a vehicle that will be used by a fuzzy logic controller, as opposed to a PID motor controller [11], to address the differences in driver behavior. The fuzzy logic controller's main objective is to keep an optimal car-to-car distance while cruising. Since fuzzy logic control does not rely on exact and crisp values to characterize any response, it is applicable in modeling human driving behavior. Additionally, the fuzzy logic response can easily be adjusted to suit the needs of each unique driver. The developed hardware is easily applicable to old car models, which, unlike their newer counterparts have already some of these presented features but are proprietary and requires a lot of changes to the existing vehicle.

In effect, there is a decentralized approach in remedying the traffic buildup in a vehicular network section. The goal of this study is to minimize traffic buildup and maintain a constant speed and car-to-car distance in car-following scenario involving intelligent vehicles. It is envisioned that if all vehicles have these features with a uniquely characterized driver behavior, then traffic buildups due to driver inefficiencies are minimized.

The rest of the paper is organized as follows: Sect. 2 discusses the fuzzy logic controller used in implementing the proposed braking and acceleration features for a car-following intelligent vehicle. We also discuss here the hardware setup. Our empirical experiment results are shown in Sect. 3. Finally, we conclude our work in Sect. 4.

2 Braking and Acceleration Features for Intelligent Vehicles

In this section, we present our developed hardware needed to achieve the braking and acceleration features to assist intelligent vehicles in a car-following traffic model. We also discuss the fuzzy logic controller implemented in a Raspberry Pi that will model the driver's behavior and produce the necessary acceleration/braking output.

Figure 1a shows the 2008 Toyota Vios 1.3J vehicle used in this research study where the hardware to achieve the braking and acceleration features are installed. We chose this vehicle model because it allows us to monitor the car driving status by simply attaching an on-board diagnostics adapter. To measure the distance, we placed a Garmin LIDAR-Lite V3 directly at the windshield instead at the car bumper to limit the sensing range equal to the length of the car's front body, i.e., where the engine is located. This is depicted in Fig. 1b. This Garmin LIDAR-Lite V3 has a resolution of 1 cm and a sensing range of 5 cm–40 m which are enough specifications to achieve the car-following with braking and acceleration features of our system. Additionally, it is also inexpensive and easily accessible.

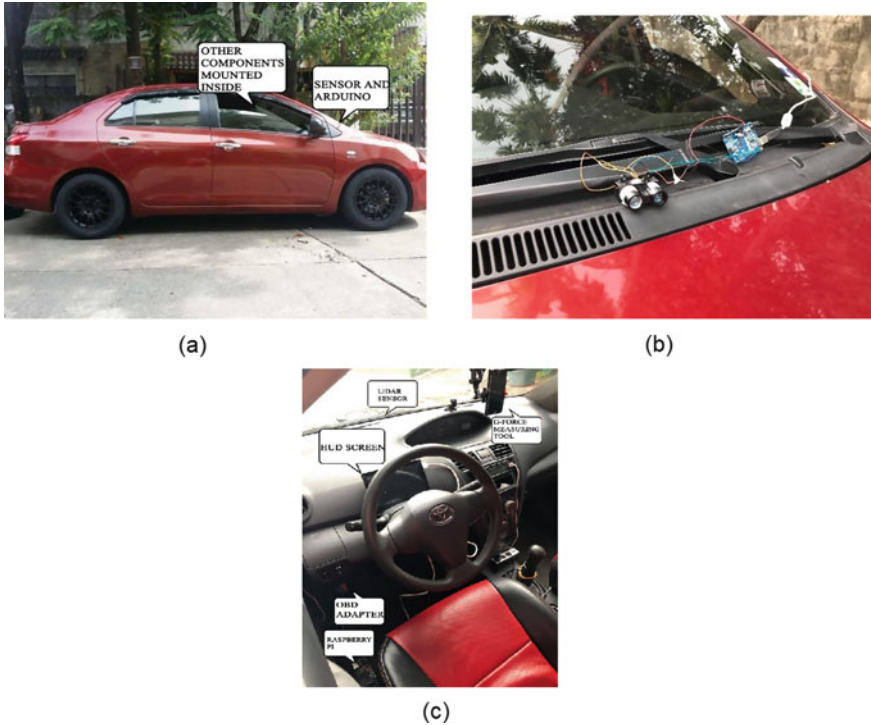


Fig. 1 **a** In this experiment, a 2008 Toyota Vios 1.3J is used to validate the braking and acceleration features for an intelligent vehicle. **b** The Garmin-LIDAR sensor is placed at the outside of the vehicle's windshield. **c** The display, smartphone and on-board diagnostics hardware are found inside the vehicle [12]

Inside the vehicle, the Raspberry Pi, Heads-Up Display, and on-board diagnostics are installed. The Raspberry Pi senses the car-follower inter-distance and speed, does the appropriate computations and provides the correct acceleration/braking output.

Since the driver still has the control of the vehicle, a Heads-Up Display informs the driver of the car's current driving status and what must be done to achieve the desired driving condition. The smartphone is also located inside to monitor the G-forces through an application. These are all seen in (Fig. 1c).

The Raspberry Pi implements a fuzzy logic control to determine the necessary braking and acceleration inputs to maintain a distance threshold between the following and leading cars while cruising at a certain speed, as shown in Fig. 2. The fuzzy logic controller input and output behaviors can easily be calibrated to suit the driver's driving attitudes, where we initially define them to be linear membership functions. Given the various combinations between the speed and distance inputs, the Mandami-based FLC provides whether the car needs to accelerate or brake.

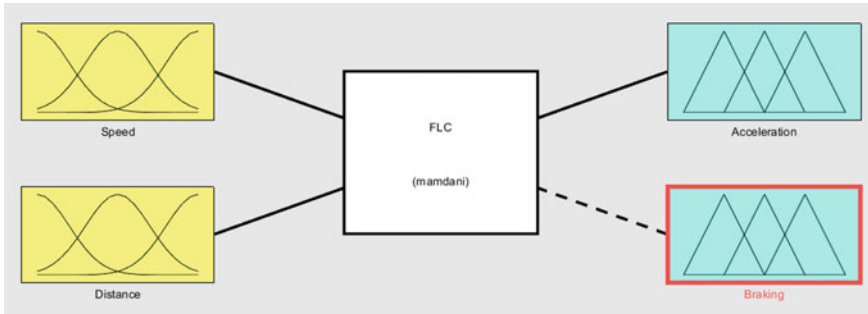


Fig. 2 Surface views of the **a** Acceleration and **b** Braking features given their fuzzy relationship with the speed and distance of the car-following intelligent vehicle

Table 1 25-fuzzy rule set for the Braking and Acceleration Features of a car-following intelligent vehicle

| Speed\distance | TC | C | A | F | VF |
|----------------|----|----|----|----|----|
| TS | FB | FB | LA | MA | HA |
| S | FB | MB | LA | MA | HA |
| N | HB | MB | LA | MA | MA |
| F | FB | MB | LB | LA | LA |
| TF | FB | HB | MB | MB | LB |

Table 1 depicts the 25 rules implemented for the braking and acceleration features where we categorized the speed range (S), $0 \leq S \leq 100$ kph, into five levels, namely, Too Slow (TS), Slow (S), Normal (N), Fast (F), and Too Fast (TF). The distance between cars, D , $0 \leq D \leq 3000$ cm, are also classified into five levels, namely, Too Close (TC), Close (C), Adequate (A), Far (F), and Very Far (VF). Finally, the acceleration and braking responses are categorized into Light (L), Moderate (M), Heavy (H), and Full (F). Note that for a more responsive or braking or acceleration options, these 25 rules can further be extended at the expense of requiring more computational power. Figure 3 depicts the surface views formed by the 25-fuzzy rule set in Table 1. This is generated in the Fuzzy Logic Control toolbox of Matlab.

For simplicity, triangular membership functions are used to assume the linear relationship between all input and output values. A more complicated modeling may implore non-linear membership functions. Given these input combinations, the Raspberry Pi, performs the necessary calculations, through an if-else conditional statements. For example, when the instantaneous speed is *NORMAL*, N , AND the current separation between cars is *ADEQUATE*, A , then, we apply an acceleration value equivalent to *LIGHT*, LA .

The fuzzy logic controller implements an implication method of the *min* operator for the inputs, while the *max* operator aggregates the outputs. To compute the necessary output, the centroid of the aggregated output is used for defuzzification [14]. An example of how the fuzzy logic controller performs this procedure is shown in Fig. 4. We note that only one output is going to be implemented in the car, i.e., either accelerate or brake, and will be dictated also by the Raspberry Pi.

3 Experimental Results and Discussion

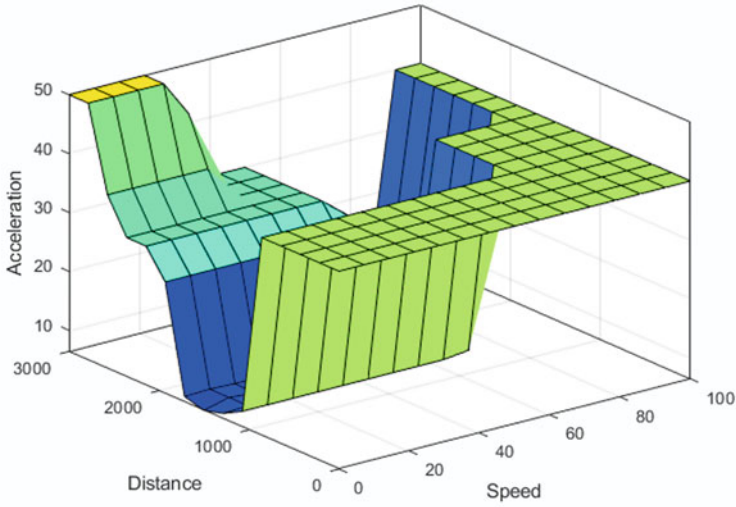
Our experimental tests have been implemented in a subdivision to initially limit the speed and car density. This has been done to ensure that everything is fully functional and operational before deploying in a more realistic setups, i.e., on a national road. The leading car can vary its speed in an independent manner while our test vehicle should be able to adaptively adjust its speed and distance. The route employed inside the subdivision is highlighted in Fig. 5.

In one of the experiments, Fig. 6 illustrates the scenarios where our test vehicle detects an obstacle (a cyclist 12 meters away) and an empty space. Given an obstacle, our system responds by suggesting an 8% acceleration input to narrow the gap between our car and the cyclist.

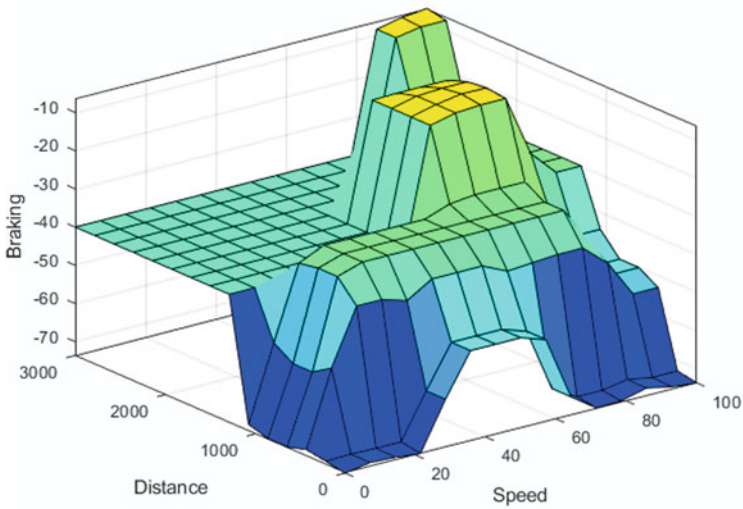
The right side panel shows the G-forces experienced by the test vehicle. The red graph is the minimal experienced lateral G-force by the car since there is a straight line motion. An increase in this value will indicate right/left turning maneuvers. The green graph is the vertical G-force and will always stay constant unless there is an upward or downward sloped road ahead. Finally, the blue graph depicts the forward and backward forces signifying that the car is either accelerating or decelerating.

In our extensive experiments in the chosen subdivision route, the performance of our proposed braking and acceleration features for a car-following intelligent vehicle is observed to approximately differ by at most 10% when comparing our computation to the vehicle's speedometer and tachometer, while making sure that the targeted inter-vehicle distance is always achieved.

In Fig. 7, we show one of our actual driving experiments. Here, our vehicle is attempting to be the car-following intelligent vehicle by trying to maintain the distance between the car leader. In the first figure, upper left, the RPM display is 1189 which denotes that it needs to hasten in order to catch the leader. In the next figure, upper right, the speed is now below that of the first picture to denote that our vehicle has already caught up. We note that our display numbers are not yet well-calibrated. Also, our system does not detect incoming vehicles and only monitors vehicles on its own lanes. This will be added to the list of improvements in the future.



(a)



(b)

Fig. 3 Surface views of the **a** Acceleration and **b** Braking features given their fuzzy relationship with the speed and distance of the car-following intelligent vehicle

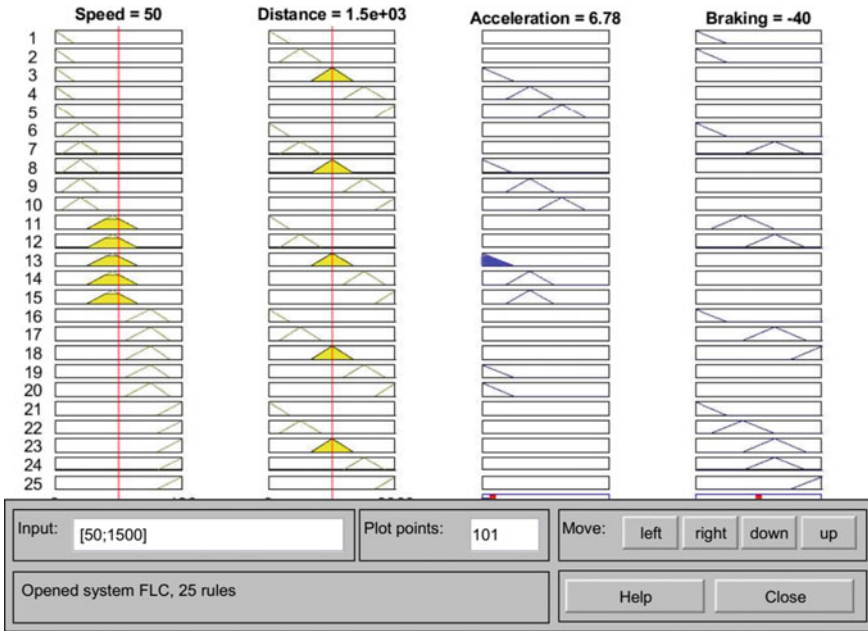


Fig. 4 The process of how fuzzy logic determines the acceleration/braking output values given the speed and distance readings. The acceleration and brake values are expressed in percentage

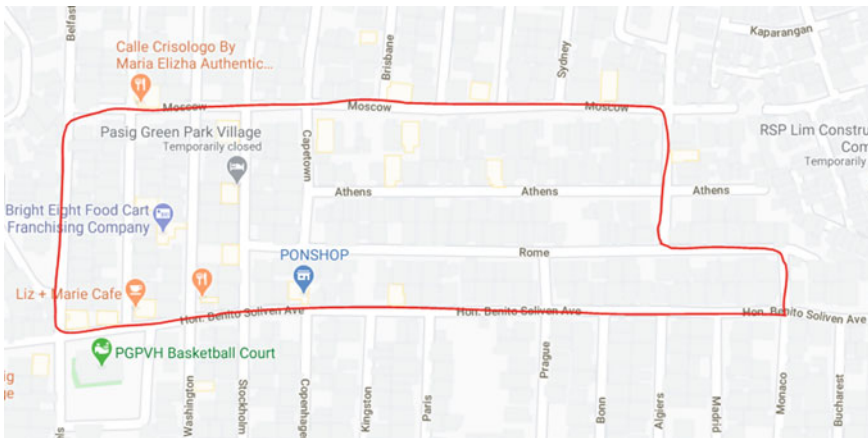


Fig. 5 Route in a subdivision to evaluate the braking and acceleration features of a car-following intelligent vehicle



Fig. 6 Readings when the test vehicle detects an obstacle in (a) and an empty space in (b). We note that incoming vehicles are not detected by system if it is found in the other lane. Also, the bicycle is not also detected by the system

As a final note regarding our system, the driver still has the total control of the vehicle. Our braking and acceleration features system only recommends the appropriate requirements needed by the vehicle to maintain an optimal distance from the leading vehicle. If the driver continuously performs unwanted braking or acceleration, the required vehicle control parameters will just be always changing as seen from the heads-up display.



Fig. 7 One of the actual driving experiments having snapshots taken every 15 s. It is viewed from left to right, then top to bottom

4 Conclusion

In this work, we have implemented a braking and acceleration feature hardware for a car-following intelligent vehicle developed from inexpensive and off-the-shelf electronic components and a fuzzy logic controller to initially model the driver's braking and acceleration behaviors. Our empirical experiments have high accuracy in maintaining the target speed and inter-vehicular distance during car-following scenarios in a constrained environment. We have integrated our hardware easily to an existing vehicle and does not require any major changes to the experimental car. In the future, we explore the various types of nonlinear fuzzy membership functions to model and use them to adjust accordingly based on the driving behaviors.

References

1. Bokare P, Maurya A (2017) Acceleration-deceleration behaviour of various vehicle types. *Transp Res Procedia* 25:4733–4749
2. Davis L (2012) Mitigation of congestion at a traffic bottleneck with diversion and lane restrictions. *Physica A: Stat Mech its Appl* 391(4):1679–1691
3. Deligianni SP, Quddus M, Morris A, Anvuur A, Reed S (2017) Analyzing and modeling drivers' deceleration behavior from normal driving. *Transp Res Rec* 2663(1):134–141
4. Duan J, Li R, Hou L, Wang W, Li G, Li SE, Cheng B, Gao H (2017) Driver braking behavior analysis to improve autonomous emergency braking systems in typical Chinese vehicle-bicycle conflicts. *Accid Ana Prevent* 108:74–82

5. Gunter G, Gloude-mans D, Stern RE, McQuade S, Bhadani R, Bunting M, Delle Monache ML, Lysecky R, Seibold B, Sprinkle J et al (2020) Are commercially implemented adaptive cruise control systems string stable? *IEEE Trans Intel Transp Syst*
6. Iizuka H, Kurami K (1994) Automatic braking system with proximity detection to a preceding vehicle. *US Patent 5,278,764*, January 11 1994
7. Ikram K, Khairunizam W, Shahriman A, Hazry D, Razlan ZM, Haris H, Halin H, Chin SZ (2018) Analysis of human behavior during braking for autonomous electric vehicles. *Eng Appl New Mater Technol* 453–459
8. Jiang P, Dong Z, Wang Y, Su Y, Li Z (2018) Empirical study of phantom traffic jams on urban expressway. In: 2018 ninth international conference on intelligent control and information processing (ICICIP), pp 99–102 (2018)
9. Liu Z, Wu M, Zhu K, Zhang L (2016) Sensafe: a smartphone-based traffic safety framework by sensing vehicle and pedestrian behaviors. *Mob Inform Sys* 2016
10. Magsino ER, Arada GP, Ramos CML (2020) Investigating data dissemination in urban cities by employing empirical mobility traces. In: 2020 IEEE 12th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM). *IEEE*, pp 1–5
11. Magsino ER, Dollosa CM, Gavinio S, Hermoso G, Laco N, Roberto LA (2014) Stabilizing quadrotor altitude and attitude through speed and torque control of bldc motors. In: 2014 13th international conference on control automation robotics & vision (ICARCV). *IEEE*, pp 438–443
12. Marasigan JC, Mayuga GP, Magsino E (2022) Development and testing of braking and acceleration features for vehicle advanced driver assistance system. *Int J Electr Comput Eng* (2088-8708) 12(2)
13. Mayuga GPT, Magsino ER (2019) Adaptive cruise control employing taillight tracking for a platoon of autonomous vehicles. *Int J Adv Trends Comput Sci Eng* 640–645
14. Obias KCU, Magsino ER (2021) Implementing a tollgate queue selector based on genetic algorithm. In: 2021 IEEE 9th region 10 humanitarian technology conference (R10-HTC). *IEEE*, pp 1–6
15. Sancar FE, Fidan B, Huissoon JP, Waslander SL (2014) Mpc based collaborative adaptive cruise control with rear end collision avoidance. In: 2014 IEEE intelligent vehicles symposium proceedings, pp 516–521
16. Wu S, Wang W, Li Z, Gao L (2017) A collision warning device based on the emergency braking behavior prediction. In: International conference on man-machine-environment system engineering, pp 415–422
17. Yadav AK, Velaga NR (2019) Effect of alcohol use on accelerating and braking behaviors of drivers. *Traff Injury Prevent* 20(4):353–358

Homogeneous Map Partitioning Employing the Effective Regions of Movement Method



Elmer R. Magsino 

Abstract In this study, we evaluate the Effective Regions of Movement (ERM) method of irregularly partitioning a location map into various homogeneous clusters. To form ERMs, the map under study is uniformly partitioned into $N \times N$ grids and their corresponding target spatiotemporal stable network characteristics, e.g., vehicular capacity, density, and speed, are calculated to present the map's spatiotemporal snapshot. From this, ERMs are formed by merging edge-adjacent grids that satisfy a set of constraints. As a consequence, ERM partitions are automatically discriminated and prioritized. We evaluate the ERM performance by employing empirical mobility traces then determine the processing time as higher resolution in terms of N is presented. Results also show that the ERM method provides fewer and homogeneous partitions when compared to DBSCAN and k -means clustering. Given such partitions, network designers or engineers can automatically perform additional procedures such as optimal RSU deployment, vehicular connectivity evaluation, and traffic management.

Keywords Effective regions of movement · Non-uniform map partitioning · Spatiotemporal stable characteristics · Intelligent transportation system · K-means clustering · DBSCAN clustering.

1 Introduction

Nowadays, public vehicles and their activities can now be tracked because of the prevalence and installation of GPS-based devices (Global Positioning System). As more public utility vehicles roam the urban and rural roads to transfer people, com-

E. R. Magsino (✉)

Department of Electronics and Computer Engineering, Gokongwei College of Engineering,
De La Salle University, Manila, Philippines
e-mail: elmer.magsino@dlsu.edu.ph

modities, and services from their origin to destination, their abundant trajectory records become more helpful in understanding mobility behavior. Each recorded GPS coordinate is the vehicle's spatial and temporal footprint in the urban/rural map. From the GPS traces of these vehicles, movement patterns, origin and destination hotspots, traffic, social, and operational dynamics, and human mobility behavior can be recognized and studied. These extracted information become beneficial in the creation of new road safety, rules, and regulations, development of new roads and infrastructure, and improvement in travel trajectory.

In understanding the empirical mobility traces, partitioning and then clustering the area under study to group related and adjacent data points becomes the initial step. The research work in [3] determined the urban city road area hotspots and mobility patterns by implementing a modified DBSCAN (density-based spatial clustering of applications with noise). Clustering has also been done in [15] to observe mobility while enhancing connectivity among dynamic vehicles. In [12] performed a uniform partitioning to model the road network by removing empty grids. As the number of grids is increased, the more visible the road networks are. On the other hand, the Manhattan model was used in [11] for studying the capital and operating expenditures in roadside unit deployment. We note that the Manhattan model is a famous urban model because of its uniform grid-type formation. In [16], a grid-growing clustering algorithm was presented to group geo-spatial points into its corresponding partitions without the need of identifying the parameter on how many clusters must be formed.

In this study, we evaluate the proposed non-uniform partitioning of vehicular maps presented in [9] by utilizing empirical taxi mobility dataset and determine the algorithm's complexity. This method of partitioning urban maps into irregular but homogeneous partitions will provide a better insight on how a vehicular network operates given a characteristic under observation. Since areas display one uniform behavior, it can be thought that these regions can behave in the same manner. Once map locations are partitioned in homogeneous ERM clusters, transportation applications such as data dissemination problem in vehicular networks [8], control of connected and autonomous vehicles [10], accident reporting [2], and traffic management [14] can be further explored. The principle of ERMs can also be used in pre-processing of image data to be used in machine learning [1, 4].

The rest of the paper is organized as follows: Sect. 2 discusses the fuzzy logic controller used in implementing the proposed braking and acceleration features for a car-following intelligent vehicle. We also discuss here the hardware setup. Our empirical experiment results are shown in Sect. 3. Finally, we conclude our work in Sect. 4.

2 Formation of Effective Regions of Movement from Stable Vehicular Grid Characteristics

The Effective Regions of Movement (ERM) concept is a partitioning method of irregularly grouping uniform grids to achieve homogeneous clusters based on a given

vehicular network characteristics, e.g., regional speed, available information, vehicular capacity, or density [9]. These homogeneous ERMs are also effectively given priority.

2.1 Computation of Spatiotemporal Stable Grid Characteristics

Figure 1 illustrates a vehicular network map with equally partitioned grids. Each grid $g_{p,q}$ contains a spatiotemporal stable network characteristic, $\zeta_{p,q,STS}$, defined by (1), where $\alpha(iT_S)$ and $\omega(iT_S)$ correspond to the feature scaling parameter and the weight

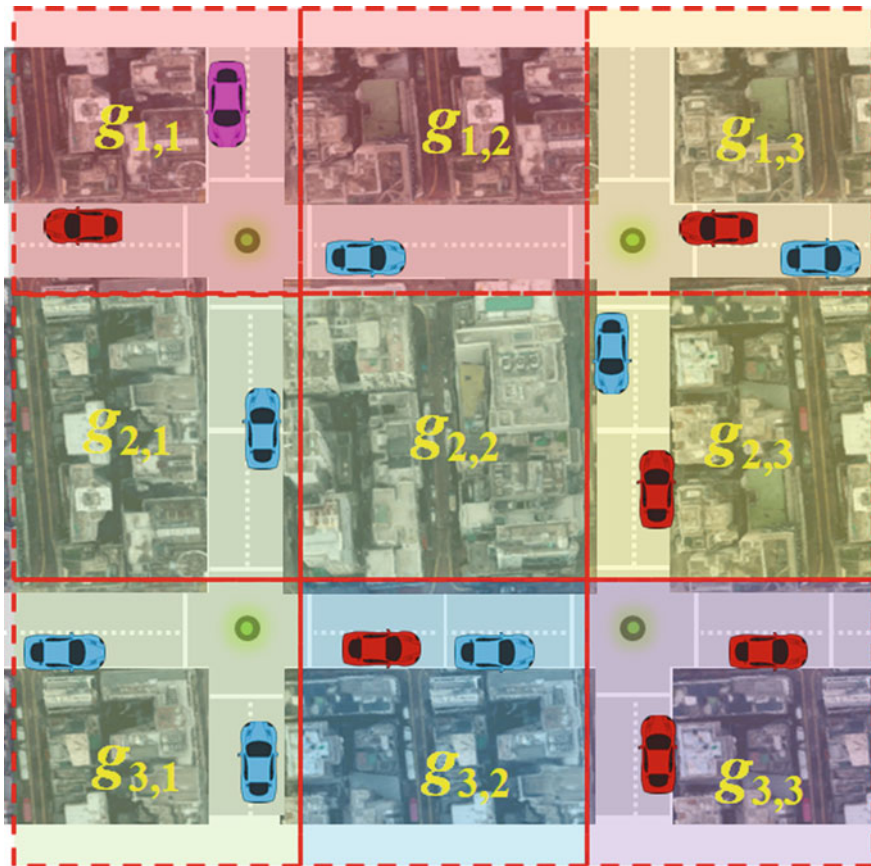


Fig. 1 A vehicular network map that is uniformly partitioned into $N \times N$, $N = 3$, equal grids. Each vehicle can represent a certain vehicular capacity in each uniform grid. Color highlights pertain to the ERM the grid belongs to.

correlating all the $\zeta_{p,q}$'s at time $t = iT_S$, respectively. T_S is denoted as the mobility trace sampling time in minutes.

$$\zeta_{p,q,STS} = \sum_{i=0}^I \alpha(iT_S)\omega(iT_S) \tag{1}$$

where

$$\alpha(iT_S) = \frac{\zeta(iT_S) - \min[\zeta(iT_S)]}{\max[\zeta(iT_S)] - \min[\zeta(iT_S)]} \tag{2}$$

$$\omega(iT_S) = \frac{\zeta(iT_S)}{\max[\zeta(i = 0, \dots, IT_S)]} \tag{3}$$

An example of how to utilize (1) is illustrated in Fig. 2 given $N = 20$ and $T_S = 10$ min. The $\zeta_{p,q,STS}$ is derived from the vehicular capacity of each grid sampled at various times of the day from 12 midnight to 11:50 PM.

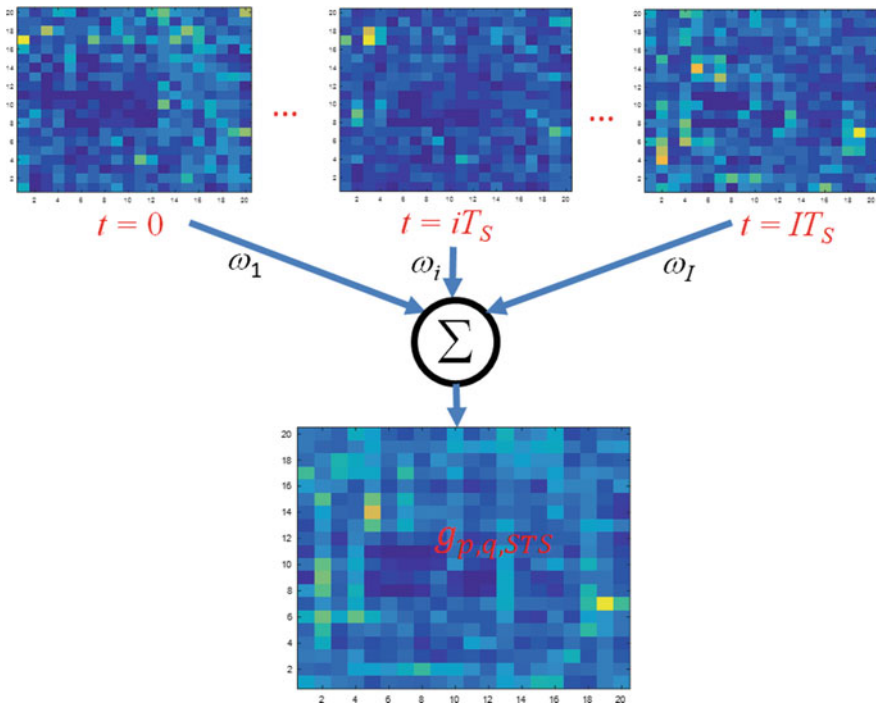


Fig. 2 An illustrative example of how to obtain the spatiotemporal stable network characteristics from a uniformly partitioned map.

From these spatiotemporal stable map grids, the formation of an ERM, ERM_e , is adopted from [16] and formulated by (4a), where the union of two uniform grids will depend on a threshold value, ζ_c , as defined by Constraint (4b). The $|\bullet|$ in the constraint defines the cardinality of the uniform grids considered for merging. Also, Constraint (4b) can also make not use of the cardinality condition, e.g., regional speed where average is best considered. We note that the ERM formation can be tight, i.e., more conditions, or loose, i.e., few constraints. $\Delta p, \Delta q \in \{-1, 0, 1\}$. Also, ERM formation in the study considers only edge-adjacent grids, which are highly fitted for grid-like road network structures (Fig. 3).

$$ERM_e \equiv g_{p,q} \cup g_{p+\Delta p, q+\Delta q} \tag{4a}$$

$$\text{subject to } |\{\zeta_{g_{p,q}}\} \cup \{\zeta_{g_{p+\Delta p, q+\Delta q}}\}| \leq \zeta_c \tag{4b}$$

2.2 Illustrative Example

Consider Fig. 1 and define $\zeta_{g_{p,q}} = c_{g_{p,q}}$, where $c_{g_{p,q}}$ is the grid's vehicular capacity and set $\zeta_c = 100$. We also set the violet car equal to 50 vehicles, red car equal to 30 vehicles, and light blue car equal to 20 vehicles. Initially, let us choose $g_{1,1}$ as the starting grid to form our ERMs. $g_{1,2}$ and $g_{2,1}$ are only considered for merging because

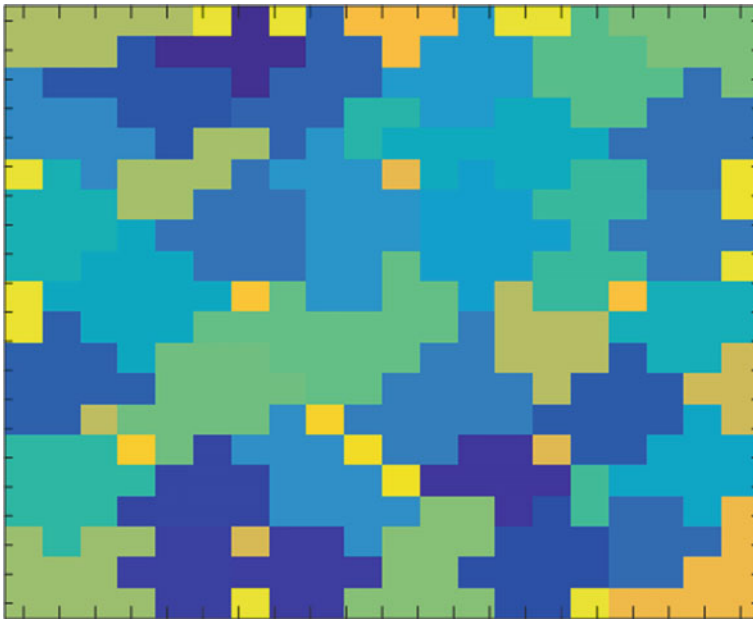


Fig. 3 Sample formed ERMs derived from Fig. 2 following the formation guidelines depicted in (4a)

of the edge-adjacency condition. $|g_{1,1}| = 80$ vehicles < 100 . Since $|g_{1,1} \cup g_{1,2}| \leq 100$, then these grids are merged forming ERM_1 . $g_{1,3}$ will be paired with $g_{2,3}$ and labeled as ERM_2 . This process goes on until all grids are associated to an ERM cluster. In Fig. 1, there are five ERMs where two are single-grid ERMs (light blue and light violet highlights). The ERMs are shown in highlighted color marks.

The formation of ERMs follow these rules.

1. The first grid considered for merging is the grid above $g_{p,q}$, i.e., $g_{p-1,q}$, and follows a clockwise direction in selecting the next grid for merging.
2. Initial grid selection seeds can follow a certain set of criteria, e.g., grid with the highest vehicular density or capacity, specific landmark, or, grid with highest road speed.

3 Simulation Results and Discussion

To evaluate the ERM non-uniform partitioning concept, the empirical taxi mobility dataset from Beijing [6] is utilized. The coverage area spans the 1st and 2nd Rings of Beijing City. We emphasize that ERM relies on the abundant and less anomalous mobility traces. The dataset attributes are given in Table 1. The Beijing taxi dataset used in the evaluation of the ERM partitioning method has the following features, namely: (1) longitude, (2) latitude, (3) time stamp, and Taxi ID. Each GPS entry in the dataset is sampled every 10 s from 00:00 to 23:59:50 daily.

Simulation is done under Matlab by following the procedures listed below. The processing time it takes to fully complete these tasks are logged for comparison as the sampling time and N are varied.

1. Extract mobility traces given an area limit of a map under study.
2. Perform uniform partitioning based on a given N .
3. For each grid, determine the vehicular capacity.
4. Perform (1) to compute for the spatiotemporal stable network characteristics of the map under study.
5. Merge edge-adjacent grids based on (4a).

Table 1 Spatiotemporal characteristics of Beijing taxi mobility dataset

| Dataset attribute | Value |
|--|--------------|
| Number of observed days | 7 days |
| Total number of taxis | 24,845 |
| Daily average number of taxi trips | 79,012 trips |
| GPS sampling Time | 10 s |
| Total number of recorded time each day | 24 h |
| N | 500 |

3.1 Examples of Spatiotemporal Stable Network Characteristics

The Beijing map is uniformly partitioned to determine the grid's vehicular capacity and location average speed. The grid car capacity is easily obtained by counting the number of vehicles present in the grid through their longitude and latitude coordinates. An additional pre-processing step is done to remove repetitive GPS traces caused by traffic slowdown or jams, passenger loading, and unloading scenarios.

The dataset does not provide the taxi's instantaneous speed, S_{ins} in kph, therefore, the speed at each sampling time, $T_S = 10$ sec, is computed in (5).

$$S_{ins} = 360d \quad (5)$$

$$d = 6371c \quad (6)$$

$$c = 2\text{atan2}(\sqrt{a}, \sqrt{1-a}) \quad (7)$$

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\Delta\lambda}{2}\right) \quad (8)$$

where $\Delta\phi$ and $\Delta\lambda$ are the latitude and longitude difference between GPS points 1 and 2, respectively. d is in kilometers.

However, this speed method computation has one drawback. The speed calculation via the Haversine method seldom provides unrealistic values brought about by incorrect GPS data logging. This happens when GPS satellites are lost, e.g., traversing tunnels or presence of high-rise buildings.

Given the uniform grid's processed vehicular capacity and calculated speed, the spatiotemporal stable network characteristics for the Beijing map are shown in Fig. 4. Here, $N = 500$ to highlight that as N increases, the road network of an urban setup under study becomes more pronounced, but at the expense of computation complexity. The vehicular capacity is normalized to the highest grid with the most cars. On the other hand, the grid speed is normalized to 100 kph. Lighter colors are locations highlighting positions where there are anomalies in the logged GPS data since the values are above 1.0. These figures show where most taxis can be found throughout the day and what is the average speed. There is a direct relationship between these two plots, i.e., in locations where there are more vehicles, the taxi speeds are normally slower than those places with few taxis.

An advantage of using the spatiotemporal stable network characteristics is we can simply remove unnecessary or unwanted grids or emphasize desired grids by applying thresholds to the network characteristics. Shown in Fig. 5 are examples of implementing higher (in (a)) and lower vehicular (in (b)) capacity threshold values. Figure 5a shows a uniform distribution of taxis across Beijing City, however, when we reduce the threshold capacity, we can easily visualize which road networks are denser.

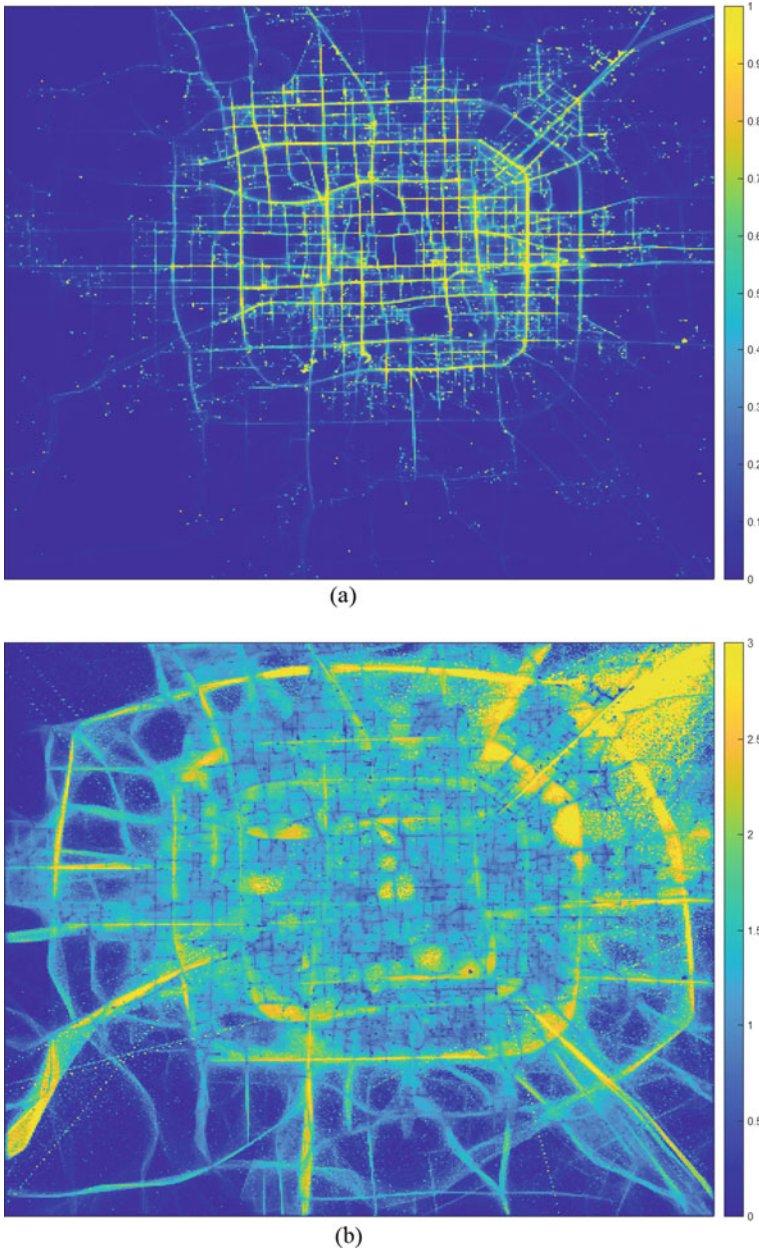


Fig. 4 The vehicular capacity (a) and speed (b) spatiotemporal stable network characteristics derived from the Beijing taxi dataset with $N = 500$. Lighter color means higher volume and faster speed

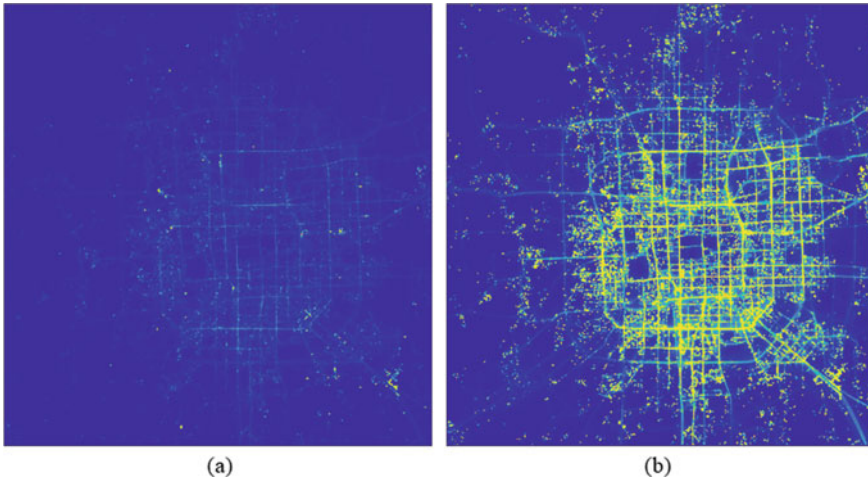


Fig. 5 Applying Thresholds to the Spatiotemporal Stable network characteristics **a** Higher and **b** Lower vehicular capacity threshold. Lighter color indicates denser grids

3.2 Processing Time of Spatiotemporal Stable Network Characteristics

Figure 6 presents the amount of processing time, in semilogy plot, needed to obtain the spatiotemporal stable network characteristic of each grid, $g_{p,q}$, as the number of uniform partitioning, N , increases and as the sampling time, T_S , is varied. We note that the processing time is equal to the following activities: (1) uniform partitioning of the map into $N \times N$, (2) calculate grid network characteristics, and (3) computing the spatiotemporal network characteristics in (1). Knowing how long the processing time it takes for the spatiotemporal network characteristics to be computed is necessary to give users the idea how scalable the procedure once time-consuming parameters are changed, e.g., N , number of involved empirical mobility traces, grid speed, etc.

As the sampling time decreases, from 30 to 2 min, the processing time increases as there are more mobility traces that are considered for each uniform grid. However, it is observed that for all sampling times, there is almost the same slope in the increase of processing time as N is increased. This result supports the scalability when the spatiotemporal network characteristics are calculated.

Figure 7 shows the percentage of empty grids as N is increased while also varying the sampling time, T_S . At lower number of partition $N \times N$, there are few to no visible empty grids because existing partitions already encompassed these grids with zero vehicular capacity. There is a similar trend of having empty grids as N is increased and as sampling time, T_S is varied. At $T_S \leq 10$ min, the decrease in empty grids becomes obvious since more GPS coordinates have been considered. We note that more mobility traces taken into account imply more movement in the area under

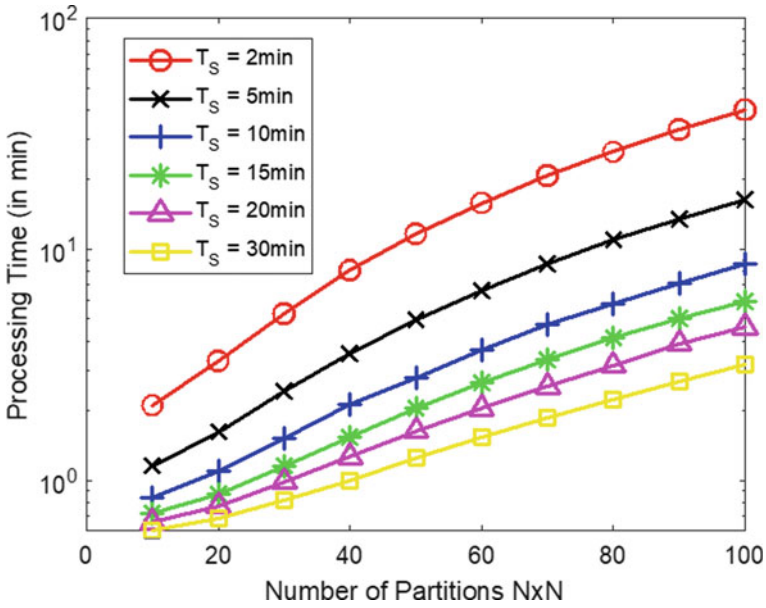


Fig. 6 The computational complexity of processing the spatiotemporal stable network characteristics as N^2 increases

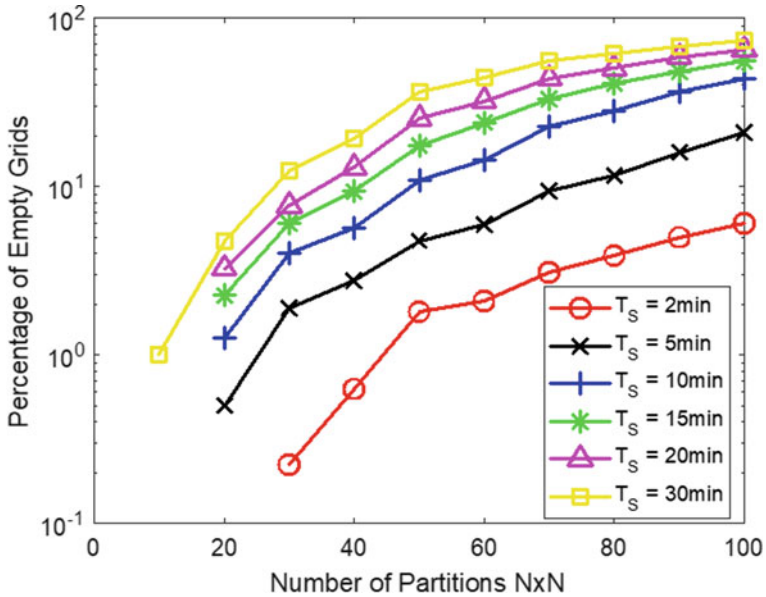


Fig. 7 The amount (in %) of empty grids as sampling time and N are varied

study, which means that there are less accidents and traffic jams happening. It may also imply that the dataset date can possibly involve holidays or special days.

3.3 Comparison with other Clustering Methods

We compare the ERM partitioning method with one of the simplest unsupervised clustering techniques, k -means [7] and DBSCAN, which stands for Density-based Spatial Clustering of Applications with Noise [5]. k -means minimizes the in-cluster distance from the designated cluster center. On the other hand, DBSCAN employs the density reachability and connectivity criteria to cluster related points. DBSCAN needs two parameters: (1) number of minimum nearby neighbors, $minPts$ and (2) radius ϵ that encloses these points [13].

Figure 8 illustrates the formed clusters when using the (a) k -means, (b) DBSCAN, and (c) ERM partitioning methods. When using the k -means technique, a total of 216 clusters are formed and discriminated by the various colors. The value of $k = 216$ comes from using the elbow method in determining the optimal value of k . In using DBSCAN, the parameters are set to $minPts = 25$ and $\epsilon = 0.0003$, thereby forming 4090 clusters in the process. The red color represent the unclustered points, therefore, considered as noise. In this result, a trial and error method has been implemented. Finally, there are 109 ERMs formed when using the edge-adjacent ERM partitioning technique. ERM produces fewer clusters compared to these two benchmarks because ERM based its homogeneity on the vehicular capacity of the network, unlike k -means that focuses on minimizing the euclidean distance in the cluster and the DBSCAN that relies on the values of $minPts$ and ϵ .

Table 2 summarizes the salient features of ERM partitioning method in comparison to the two benchmarks, k -means and DBSCAN. ERM produces fewer clusters that are prioritized and homogeneous in terms of a desired vehicular behavior.

When utilizing k -means, the clusters are almost uniform in shape and area. These partitions are inherently adjacent to each other because of the objective function governing it. The trouble with using k -means in geo-spatial data is that when there are GPS coordinates that are too close to each other. Based on experience, these points make the simulation run longer.

Table 2 Comparison between ERM and other benchmarks

| | k -means | DBSCAN | ERM |
|------------------|----------------------------|--------------|--------------------------------|
| # of Clusters | 216 | 4090 | 109 |
| Cluster Priority | w/o priority | w/o priority | w/ priority |
| Homogeneity | Homogeneous inter-distance | None | Homogeneous vehicular capacity |

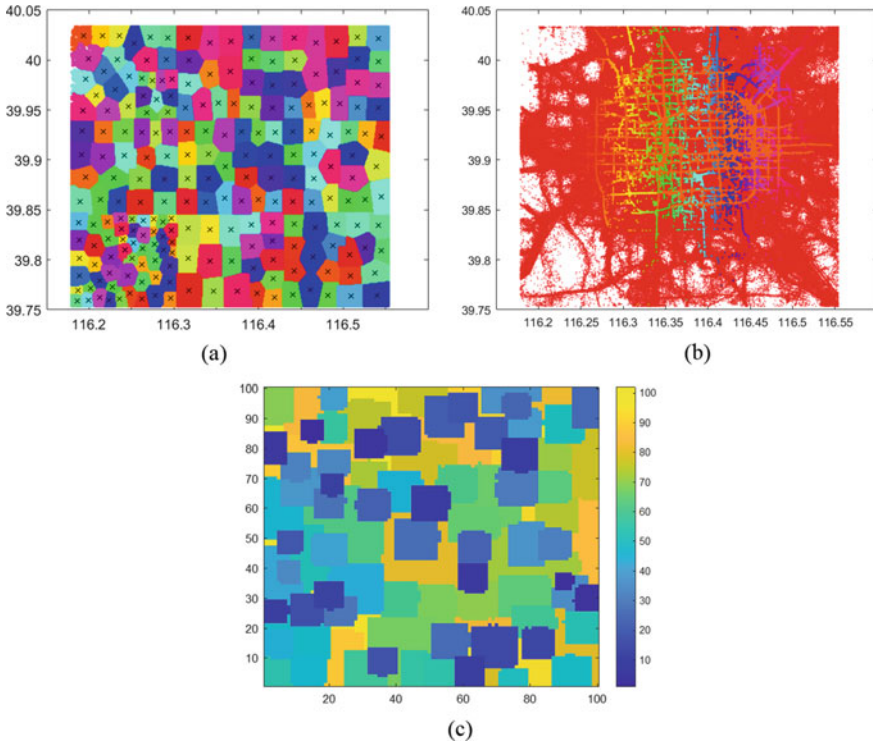


Fig. 8 Beijing map partitioning using the **a** K-means and **b** DBSCAN, and **c** ERM (with 100×100 grids) methods

Visually, DBSCAN provides clusters that follow road networks based on trajectories. From Fig. 8b, the partitions highlight the Beijing roads highly traversed by taxis, which cannot be seen from *k*-means and ERM. If a certain spatial location is empty of traces, DBSCAN will neglect these locations.

Unlike *k*-means and DBSCAN, the ERM partitioning method provides irregular partitions that can also include the empty grids and its corresponding priority level, as dictated by its color. The lighter the color is, the higher the priority of that ERM partition. It is also possible that two ERMs may share the same priority, as shown in the example. This is due to the fact that the criteria of selecting priority relied on the grid capacity, which resulted in grids that have the same vehicular content. It should also be noted that the ERM area is tightly coupled on the constraints given.

Common to all clustering techniques is the fact that they rely on the initial seed to start the clustering. In ERM, the starting point is located at the most dense grid, but can also be changed according to the application on hand.

4 Conclusion

In this work, we have evaluated the performance of the irregular partitioning presented by the Effective of Regions of Movement (ERMs) concept. ERMs are derived from the spatiotemporal network characteristics of the map under study to consider the effects of the instantaneous vehicular capacity at the current time and space. ERM has been also compared with the k -means and DBSCAN clustering techniques. Common to these three partitioning methods is the selection of initial seeds where clusters will begin grouping. However, unlike these two methods, ERM prioritizes partitions and also encompasses the empty grids which may be useful in some applications, e.g., vacant lots where roadside units can be built.

In the future, we explore the effects of performing node-adjacency in forming ERMs and further evaluate the ERM concept in other empirical mobility datasets, particularly those that do not have as many points as the Beijing dataset.

References

1. Akey Sungheetha RSR (2021) Classification of remote sensing image scenes using double feature extraction hybrid deep learning approach. *J Inf Technol* 3(02):133–149
2. Bhatti F, Shah MA, Maple C, Islam SU (2019) A novel internet of things-enabled accident detection and reporting system for smart city environments. *Sensors* 19(9):2071
3. Cheng H, Fei X, Boukerche A, Almulla M (2015) Geocover: An efficient sparse coverage protocol for rsu deployment over urban vanets. *Ad Hoc Netw* 24:85–102
4. Dhaya R (2021) Hybrid machine learning approach to detect the changes in sar images for salvation of spectral constriction problem. *J Innov Image Process (JIIP)* 3(02):118–130
5. Ester M, Kriegel HP, Sander J, Xu X et al. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: *kdd*, vol 96, pp 226–231
6. Ho IWH, Chau SCK, Magsino ER, Jia K (2019) Efficient 3d road map data exchange for intelligent vehicles in vehicular fog networks. *IEEE Trans Veh Technol* 69(3):3151–3165
7. MacQueen J et al. (1967) Some methods for classification and analysis of multivariate observations. In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol 1. Oakland, pp 281–297
8. Magsino ER, Arada GP, Ramos CML (2020) Investigating data dissemination in urban cities by employing empirical mobility traces. In: *2020 IEEE 12th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM)*. IEEE, pp 1–5
9. Magsino ER, Ho IWH (2022) An enhanced information sharing roadside unit allocation scheme for vehicular networks. *IEEE Trans Intell Transp Syst* 1–14 (2022). <https://doi.org/10.1109/TITS.2022.3140801>
10. Mayuga GPT, Magsino ER (2019) Adaptive cruise control employing taillight tracking for a platoon of autonomous vehicles. *Int J Adv Trends Comput Sci Eng* 8(3):640
11. Nikookaran N, Karakostas G, Todd TD (2017) Combining capital and operating expenditure costs in vehicular roadside unit placement. *IEEE Trans Veh Technol* 66(8):7317–7331
12. Sarubbi JF, Silva TR, Martins FV, Wanner EF, Silva CM (2017) A grasp based heuristic for deployment roadside units in vanets. In: *2017 IFIP/IEEE symposium on integrated network and service management (IM)* IEEE, pp 369–376

13. Schubert E, Sander J, Ester M, Kriegel HP, Xu X (2017) Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Trans Database Syst (TODS)* 42(3):1–21
14. Sheikh MS, Liang J (2019) A comprehensive survey on vanet security services in traffic management system. *Wirel Commun Mob Comput* 2019 (2019)
15. Wang C, Li X, Li F, Lu H (2014) A mobility clustering-based roadside units deployment for vanet. In: *The 16th Asia-Pacific network operations and management symposium, IEEE*, pp 1–6
16. Zhao Q, Shi Y, Liu Q, Fränti P (2015) A grid-growing clustering algorithm for geo-spatial data. *Pattern Recogn Lett* 53:77–84

Demand-Based Deployment of Electric Vehicle Charging Stations Employing Empirical Mobility Dataset



Camille Merlin S. Tan and Elmer R. Magsino 

Abstract In this study, we uniformly partition an empirical vehicular network map and determine the locations where electric vehicle (EV) charging stations will be deployed by implementing the particle swarm optimization (PSO) algorithm. The optimal placements utilize the spatiotemporal taxi movements representing energy demands. From each partition, the taxi GPS coordinates are extracted and used to obtain the global best location of the charging station. Parameters such as the total number of taxis accessible to each deployed EV charging station and the average distance of each taxi from all the charging stations have also been computed to evaluate the deployment method. Results have shown that the total number of taxis within the 1-km radius distance of each EV charging station ranges from 797 to 1218 taxis, while the average distance of each taxi from all of the charging stations ranges from 3.14 to 4.46 km. We also note that there is a small inter-distance separation among all charging stations.

Keywords Electric vehicle charging stations · Global best · Particle swarm optimization algorithm · Optimal deployment · Smart city · Intelligent transportation systems

1 Introduction

In recent years, electric vehicles (EVs) have been able to attract the interest of many people because they are environmental-friendly. Compared to conventional vehicles

C. M. S. Tan · E. R. Magsino (✉)
Department of Electronics and Computer Engineering,
Gokongwei College of Engineering, De La Salle University, Manila, Philippines
e-mail: elmer.magsino@dlsu.edu.ph

C. M. S. Tan
e-mail: camille_merlin_s_tan@dlsu.edu.ph

that use gasoline, the acceleration capability of EVs is also excellent while producing less emissions, requiring lower maintenance, and dissipating less power and noise pollution [8]. However, people often worry about the battery life and the available charging stations in the area of their trajectory. For this reason, it is necessary to find an optimal location for EV charging stations that addresses the driving range of EV and the traffic density in the specific area [4].

Most EV batteries have around three years of warranty and a 60,000-mile mileage, with some well-known EV brands having longer warranty periods, i.e., approximately eight years with mileage of 100,000 miles. Tesla and Hyundai both have eight years of warranty, with Tesla having a mileage limit of 150,000 miles. For Hyundai, Model S, Model X, and regular models have mileage limits of 150,000, 120,000 and 100,000 miles, respectively. Another important thing to take note about the EV battery is that it degrades over time. Moreover, the speed in driving the EV also affects the battery life, i.e., the faster it is driven, the quicker it can deteriorate the battery life because of heat build-up. Therefore, driving at optimal speed is also recommended to ensure that the battery life lasts long [6].

The EV charging stations are classified into three levels, namely, Level 1, Level 2, and Level 3, which are based on their transmitted power and charging speed. These two parameters dictate the amount of waiting time when an EV needs to recharge [16]. Level 1 stations have the same rate as charging on household outlets. It utilizes the 120-volt voltage level and its charging time ranges from 4-36 hours to achieve full charge. On the other hand, Level 2 stations are suitable for public places such as parking lots and commercial areas because they use the 240-volt voltage level. Its charging rate ranges from 1-6 hours to reach full charge. Lastly, Level 3 stations are made to provide commercial vehicles such as taxis and company cars rapid charging. They are usually placed on highways since they require large grid connection capacity due to its large power units. Its charging rate is very fast such that it can achieve full charge within an hour.

In addition, there is an issue between the compatibility of EV chargers and EV, e.g., Tesla supercharger can only be used to charge Tesla vehicles. On the other hand, non-Tesla chargers can be used to charge Tesla vehicles by using an adapter. Also, since level 3 chargers provide a fast charging option, the charging cost is very expensive and varies depending on the station owners. Some charges depend on the duration of the EV being connected to the charger, while other charges are based on the total amount of energy transferred. In most cases, level 3 fast chargers cost around 2-3 times more than normal chargers, which may end up costing the same amount as using gasoline in the long run. However, the carbon emissions would have been very much lessened during this period [15].

At present, the increasing number of EV demands has led to the expansion of deploying charging station networks in a locale. The main target locations are commercial areas such as parks, business, and shopping centers because of vehicular traffic. Placing the optimal number of EV charging stations at the correct place becomes crucial. In planning for the optimum location of the EV charging stations, several factors need to be considered such as user demands in a specific location, installation cost, size, maintenance cost, and type of charging station [7].

In this work, we simplify the EV charging station deployment study by deploying only homogeneous EV charging stations, i.e., only one level. We utilize taxi mobility traces to mimic the movement of electric vehicles roaming around the city and predict the location demand over a certain period of time. By understanding the area demand in terms of vehicular capacity, we employ the particle swarm optimization (PSO) concept to determine where the EV charging station should be located at a certain area. PSO is a good choice in this study since it is based on the dynamic movements of a natural social behavior like flock of birds, fish and insects, wherein in this study, the taxis are considered the flock of particles, [3] trying to find out the charging infrastructure for recharging. In addition, the PSO algorithm uses a simple implementation given a few needed parameters, and an efficient global search algorithm. On the other hand, the main disadvantages include slow convergence in the refined search stage and easy to fall into local optimum in high dimensional space [1].

The rest of the paper is organized as follows. Section 2 provides a brief overview of related literature. Section 3 discusses our proposed methodology. Simulation results and its discussion are provided in Sect. 4. Finally, the conclusion of the work is given in Sect. 5.

2 Review of Related Literature

The research in [10] introduced the methods of using the improved particle swarm optimization (IPSO)-based and K-means algorithms to determine the location of the charging station and the number of charging piles required based on the user's demand. Various datasets were used for different time slots. Results show that IPSO based algorithm is better than K-means clustering technique because it was able to reduce the total cost of deploying charging stations by 9.19%.

The research in [17] presented on implementing an optimal placement of the EV charging stations in IEEE 33 node radial distribution network divided into three areas to have charging stations at different locations of a specific area. Two methods were applied in performing the allocation problem, namely Whale Optimization Algorithm (WOA) and Grey Wolf Optimization (GWO). Results show that both WOA and GWO were able to achieve the same outcome, which confirms that the solution is optimal. In addition, although the deployment of EV charging stations plays an important role in the distribution system, it should not be the only parameter to be considered. Some parameters such as power loss, voltage drop and cost are necessary.

The research in [9] discussed on implementing a method to determine the optimal size and location of EV charging station with maximum capacity in the distribution network while at the same time, stability and safety are also considered. Moreover, the study also evaluated and analyzed other parameters of the charging station such as power flow, reactive and active losses, and voltage deviation to examine their effect on the whole electric distribution network. Results show that the proposed system can only charge up to 200 EVs at a time in the IEEE 3 bus distribution system, while using the existing generators.

The research in [5] studied on implementing a framework which optimizes the EV charging station deployment by considering the budget and EV charging capacity. The study was done by collecting all of the traffic patterns and EV properties such as initial charge and battery capacity in order to make a model for the deployment of the charging stations such that each EV can reach the target destination within the expected time. Computer simulations were performed to evaluate the proposed framework and evaluate its scalability. Results show that the proposed framework was able to perform efficiently.

Finally, the research in [18] implemented a technique to minimize the social cost by reducing the EV travel distance. The number and location of the fast charging stations (FCS) were identified based on the EV demands. The total EV charging demand was computed by estimating the number of households, leading to altering the IEEE123 bus distribution system. A Voronoi diagram was used to show the FCS service area. Moreover, two types of clustering techniques were applied to find the locations of the FCS, namely, fuzzy C-means clustering and K-means clustering. Results showed that both clustering techniques were able to give good results but K-means clustering has a better performance.

Different from these previous studies, we locate optimal EV charging locations by studying the behavior of empirical taxi mobility traces. Understandably, public utility vehicle drivers are more knowledgeable in the ins and outs of an urban city. They understand the road networks and make necessary adjustments to arrive at their destination the soonest possible time. Given this, we utilized the Particle Swarm Optimization technique to aid us in determining the convergence point of these taxis in a region. This location will serve as the point where the EV charging station will be deployed. Since EV charging stations are expensive and cannot be placed side-by-side because of electrical energy demand in urban cities [2, 19] and household energy anomaly [11], we need to place them optimally such that it will not also disrupt the energy supply.

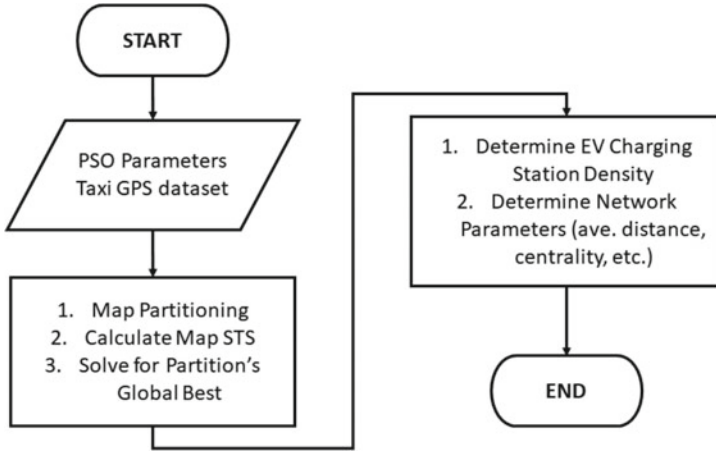


Fig. 1 Flowchart to obtain the EV charging station locations and other network parameters

3 Methodology

Figure 1 shows the flowchart of deriving the locations of EV chargers. Initially, all PSO parameters are initialized along with all of the taxi data roaming on a targeted area.

To implement the PSO algorithm, we extract the necessary empirical mobility traces to flood the desired urban map under study. These mobility traces represent the locations of our electric vehicles in this study. The dataset is acknowledged to have been sampled at different locations and times of the day for a week. To represent this area by using only one snapshot, we characterize it according to its spatiotemporal stable network characteristics [14].

Given the area, we first divide the area under study into partitions based on [12] where we want to place an EV charger. For simplicity, we divide it first into equal partitions. This will allow the PSO algorithm to only obtain one charging station (global best) among all of the taxis in a partition. From these partitions, we then, divide them to N equal sub-regions and we utilize the spatiotemporal stable (STS) network characteristics, $\zeta_{p,q,STS}$, to capture the taxi mobility in a snapshot [14].

In (1), $\zeta_{p,q,STS}$ is equal to the grid's spatiotemporal stable vehicular capacity. This value represents the general weighted average of all vehicles entering the grid at all times. $\alpha(iT_S)$ in (2) denotes the feature scaling of the vehicular capacity of a partition. This represents the temporal importance of a certain map grid. Lastly, $\omega(iT_S)$ in (3) is equal to the correlation weight at iT_S , i.e., temporal value of the map with respect to the other timestamps.

$$\zeta_{p,q,STS} = \sum_{i=0}^I \alpha(iT_S) \omega(iT_S) \quad (1)$$

where

$$\alpha(iT_S) = \frac{\zeta(iT_S) - \min[\zeta(iT_S)]}{\max[\zeta(iT_S)] - \min[\zeta(iT_S)]} \quad (2)$$

$$\omega(iT_S) = \frac{\zeta(iT_S)}{\max[\zeta(i = 0, \dots, IT_S)]} \quad (3)$$

From the derived locations of the EV charging stations, we obtained the number of viable taxis within the a 1-km radius to recharge itself on a certain area. We note that we only set it to one kilometer since practically these EVs are already considered near the charging station. We also derived the average inter-distance of between charging stations to have a picture of how far are the deployments from each other.

4 Simulation Results and Discussion

In performing the simulations in generating the electric vehicle charging station locations, the empirical mobility datasets from the Beijing taxis are utilized [13]. This dataset has 24, 845 taxis plying the roads of Beijing City sampled every 10 s for seven days.

In the simulation, we restricted the area of interest to the range $39.925 \leq lat \leq 39.977$ and $116.376 \leq lon \leq 116.462$, where *lat* and *lon* are the latitude and longitude coordinate in degrees, respectively.

4.1 The Deployment of 30 Electric Vehicle Charging Stations

Figure 2 shows the location of the 30 EV charging stations, represented by the green marker, deployed in the chosen area when using the PSO algorithm [20] running under the Matlab environment. There are a total of 17, 259 unique taxis plying the targeted area throughout the seven-day dataset, which means that the location is well-traversed. It should also be noted that there are taxis roaming the area only once in its trajectory sets. It can be observed that some parts of the area under study have more charging stations than the other sections. This is because the demand of taxis are denser in those certain areas than the others.

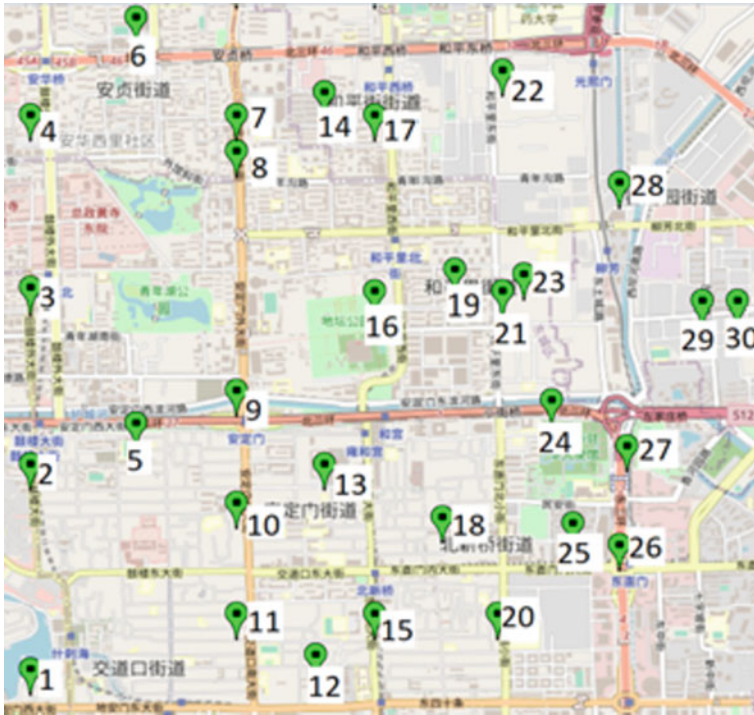


Fig. 2 The 30 Electric Vehicle (EV) charging stations deployed in the area. The map is generated by superimposing the EV charging locations on Beijing City and available for Matlab integration

4.2 Total Number of Taxis Within the 1 km Distance of Each Electric Vehicle Charging Station

We then calculated the amount of taxis within the 1-km radius distance from a charging station. For some EV charging stations which are deployed near the boundary of the whole area of interest and have less than 1 km distance on one side, the maximum boundary of the area at that side is considered instead. The EV charging station vehicular density is shown in Fig. 3. The average taxi density is approximately 1,100 taxis throughout the day and is computed by counting all nearby taxis that are one kilometer from all charging stations.

We note that 1,100 taxis may be too much for a specific charging station specifically if the charging time is 79 s. (Note that 79 s of charging time multiplied by 1,100 equals 86,400 s, i.e., one day.) However, we still have to include in our analysis that out of these 1,100 EVs will really commit itself to charging at that station, which will be covered in future undertakings.

Charging station number 1 is the least accessible among all of the charging stations while charging station number 20 is the most accessible, based on its location and

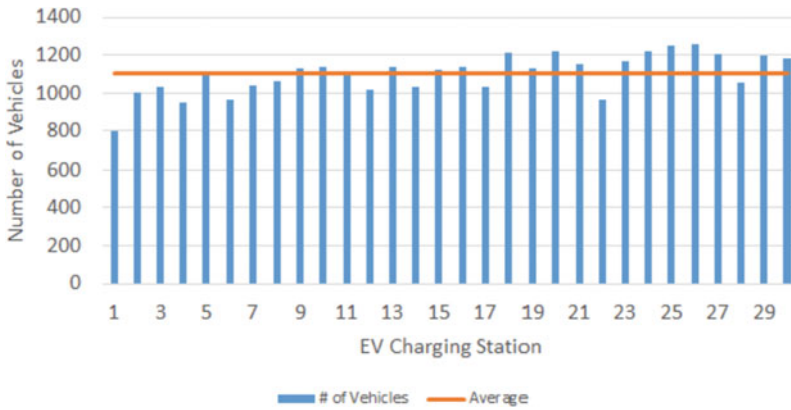


Fig. 3 The total number of taxis found in each region where an EV charging station is deployed

taxi demand. Obviously, EV charging stations found at the area center have high demands. It is only assumed in this study that waiting time is considered minimal, as such, there are no queues formed at each charging station. In the future, we will include this in the study since not all incoming vehicles have the same required energy to fully charged it.

To initially study the fact that there is a queue formation in one charging station that will require longer waiting time, we randomly chose 10 taxis to observe. From an initial charging station, we computed the average distance to the other EV charging stations. We found out that for these random sample, the average distance ranges from 3.14 km to 4.46 km. Therefore, as long as correct and real-time information are provided by the deployed charging station to the nearby vehicles, these taxis can immediately choose to go to the next nearest station or elsewhere.

4.3 Average Distance of Each Taxi from All of the Charging Stations

Figure 4 shows the average inter-distance of each charging station from the others. Given that the average inter-distance between stations is 2.1 km, taxis can go to nearby charging stations if they feel that they will need to wait for a longer time. However, we emphasize that as we make the area under study larger while maintaining the 30 uniform partitions, then, this charging station inter-distance will also increase. If that is the case, real-time infrastructure information can lessen the queue formation.

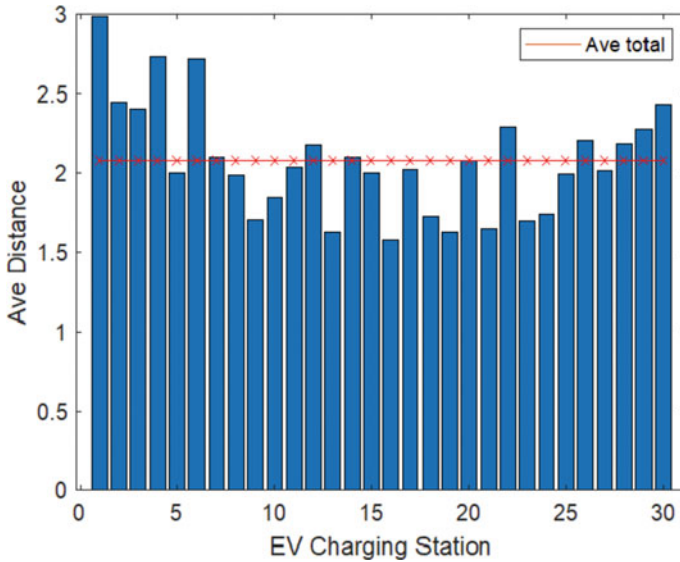


Fig. 4 The average inter-distances of an EV charging stations from the rest of the deployed EV charging stations

5 Conclusion

The optimal deployment of EV charging stations in a given study area remains to be a challenging task since every aspect of vehicle behavior should be considered and is dependent on the total area size and the total number of vehicles around the area at a specific time frame. In this work, we have deployed 30 EV charging stations from a uniformly partitioned vehicular map network area with the main objective of having a homogeneous charging deployment of taxis for all deployed charging stations. To generate only one vehicular capacity snapshot of the taxi mobility, we utilized the spatiotemporal stable network characteristic method to characterize the area under study. Our results have shown that the deployed charging infrastructures are within two kilometers from each other while allowing randomly chosen taxis to be within four kilometers from the other charging stations.

For future improvements, we further study the mobility behavior and use them to classify the area in a more homogeneous partition, which may result to a non-uniform partitioning of the system. We will also incorporate the time-to-full-charge component for each car such that the average waiting time is displayed at each charging station and becomes the main objective function in determining the optimal location of deploying the EV charger.

References

1. Bai Q (2010) Analysis of particle swarm optimization algorithm. *Comput Inf Sci* 3(1):180
2. Bindhu V, Ranganathan G (2021) Energy storage capacity expansion of microgrids for a long-term. *J Electr Eng Autom* 3(1):55–64
3. Cai J, Peng P, Huang X, Xu B (2020) A hybrid multi-phased particle swarm optimization with sub swarms. In: 2020 international conference on artificial intelligence and computer engineering (ICAICE). IEEE, pp 104–108
4. Catalbas MC, Yildirim M, Gulden A, Kurum H (2017) Estimation of optimal locations for electric vehicle charging stations. In: 2017 IEEE international conference on environment and electrical engineering and 2017 IEEE industrial and commercial power systems Europe (EEEIC/ ICPSE Europe), pp 1–4. <https://doi.org/10.1109/EEEIC.2017.7977426>
5. Datta A, Ledbetter BK, Rahman MA (2017) Optimal deployment of charging stations for electric vehicles: a formal approach. In: 2017 IEEE 37th international conference on distributed computing systems workshops (ICDCSW), pp 83–90
6. Disdale J, What is the battery life of an electric car? <https://www.autocar.co.uk/car-news/advice-electric-cars/what-battery-life-electric-car>
7. Furat MF, Dagilgan NSD, Kocaoglu SK (2019) Horizontal rotating charging platform for electric vehicles
8. Ghasri M, Ardeshiri A, Rashidi T (2019) Perception towards electric vehicles and the impact on consumers' preference. *Transp Res Part D: Transp Environ* 77:271–291
9. Kunj T, Pal K (2020) Optimal location planning of ev charging station in existing distribution network with stability condition. In: 2020 7th international conference on signal processing and integrated networks (SPIN), pp 1060–1065
10. Li W, Yang L, Wen Z, Chen J, Wu X (2021) On the optimization strategy of ev charging station localization and charging piles density. *Wirel Commun Mobile Comput* 2021
11. Magsino ER (2019) Energy monitoring system incorporating energy profiling and predictive household movement for energy anomaly detection. *Int J Emer Trends Eng Res* 7(8):151
12. Magsino ER, Abello AJ, Lalusin JM (2021) Taxi hotspots identification through origin and destination analysis of taxi trips using k-means clustering and h-indexing. In: *Journal of physics: conference series*. IOP Publishing, p 012006
13. Magsino ER, Arada GP, Ramos CML (2020) Investigating data dissemination in urban cities by employing empirical mobility traces. In: 2020 IEEE 12th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM). IEEE, pp 1–5
14. Magsino ER, Ho IWH (2022) An enhanced information sharing roadside unit allocation scheme for vehicular networks. *IEEE Trans Intell Transp Syst* 1–14 (2022). <https://doi.org/10.1109/TITS.2022.3140801>
15. Moloughney T, What are the different levels of electric vehicle charging? <https://www.forbes.com/wheels/advice/ev-charging-levels/>
16. Nicholas M (2019) Estimating electric vehicle charging infrastructure costs across major us metropolitan areas. https://theicct.org/sites/default/files/publications/ICCT_EV_Charging_Cost_20190813.pdf
17. Pal A, Bhattacharya A, Chakraborty AK (2019) Allocation of ev fast charging station with v2g facility in distribution network. In: 2019 8th international conference on power systems (ICPS). IEEE, pp 1–6
18. Shukla A, Verma K, Kumar R (2016) Consumer perspective based placement of electric vehicle charging stations by clustering techniques. In: 2016 national power systems conference (NPSC), pp 1–6
19. Wang H (2021) Fault diagnosis in hybrid renewable energy sources with machine learning approach. *J Trends Comput Sci Smart Technol (TCSST)* 3(03):222–237
20. Yarpiz, Particle swarm optimization (pso). <https://www.mathworks.com/matlabcentral/fileexchange/52857-particle-swarm-optimization-pso>

Design and Development of a Digital Preservation Voice Enabled Application for Cultural Heritage Towards Fishing in Vernacular Language



Prasad Vadamodula, R. Cristin, and T. Daniya

Abstract This research is about intangible cultural heritage towards fishing, fisherman and its missing traditional culture. This research makes us to build the scope for understanding the cultivation of selected fishes or small fish farms with confirmed areas through fishing experts and their recorded documentaries. The documented content is translated into local languages of the surrounding states of Andhra Pradesh like Telangana, Chhattisgarh, Maharashtra, Odisha, Karnataka and Tamil Nadu which has water resources Inland for fish farming in a traditional way and make them available in the web portals. Here, the human expert is the fisherman who has well knowledge about the seasons, fishes and its types and the confirmed places, of their availability. This research targets to build and enlighten the life of the fisherman based on social and economic needs in this current society by building an aquatic ecosystem.

Keywords Heritage · Culture · Fish farming · Natural language processing · Deep learning · Synthesizers · Coders and encoders

1 Introduction

The concept of digitizing the things and maintain in the web portal is now a days a normal style, but what about the traditions which are being missing along with the demise of the senior human race (related to fisherman or his ancestor's family). However, the digitization of the data requires a very high level of cleaning to ensure the consistency and availability of the program for making the documentation free

P. Vadamodula (✉) · R. Cristin
Department of Computer Science and Engineering, GMRIT, Rajam, Andhra Pradesh, India
e-mail: prasad.v@gmrit.edu.in

R. Cristin
e-mail: cristin.r@gmrit.edu.in

T. Daniya
Department of Information Technology, GMRIT, Rajam, Andhra Pradesh, India
e-mail: daniya.t@gmrit.edu.in

from bugs, noise and errors. The diagram listed below shows the exact rooting where the Traditional Farming happens [1].

Though digitization is playing a major role in preserving the language technology, the role of NLP is important for the convenient approach in the vernacular process. Hence, the idea of Heritage Culture [2] with NLP produces astonished outcomes relevant to the futuristic society described and mentioned in this project in a detailed way.

2 Literature Survey

Speech recognition is the ability of a machine or a program to respond for spoken commands. It works using algorithms through acoustic and language modelling. It is a technique that enables hands-free control of various devices to interact with machines. These speech recognition techniques had a mixture notation with Machine learning and Deep learning algorithms [3]. Natural Language Processing could act as a primary component in this research for Cultural Heritage Domains.

2.1 Fishing

This research deals with Intangible Cultural Heritage towards Fishing for identifying the.

- (a) Oral and traditional expressions of the elder and expert fisherman in identifying the seasons of fishing at a wide scope.
- (b) Science and Habits related to the fisheries and their scope, to know about the practices of fishing in the coastal areas of Srikakulam and Visakhapatnam.
- (c) Traditional Skills

Classification of fish farming can be done by the confirmed geographical area and the collaborations. Mutations are those where, generally they undergo various fished holds or survives likely on culture and confirmed areas.

Classification of fish based on culture.

- (a) Mono Culture
- (b) Mono-sex Culture
- (c) Poly Culture
- (d) Poly-sex culture.

Classification of fish based on confirmed areas.

- (a) Pond Fish
- (b) Cage Fish
- (c) Pen Fish
- (d) Sea Fish
- (e) Integrated Fish

2.2 Traditional Fishing and Farming

- a. Paddy Cultivation in Rainy Season
- b. Spear Fishing

- c. Angling
- d. Ice Fishing
- e. Trapping
- f. Hand Gathering
- g. Netting
- h. Kite Fishing

Fish culture [1] is an almost ideal method of land use, combining the production of both vegetal and animal proteins. Its development can make an important contribution to the future world food supply.

With the aim of determining and evaluating the condition of freshwater fisheries in Tunisini [4], surveys were undertaken in the most important reservoirs of the country for finding and tracing the traditional culture of getting the rear species of fish. Species size, location including the depth and surface areas that they meant to grow is considered. Fish farming can be implemented in a multitude of environments and with a variety of famous and traditional technique acquired from the fishermen.

Studies to determine the adoption of fisheries co-management [5] impacted on the ownership, empowerment and access of fishers and other stakeholders in the community to resources. Improving the fishing communities and their livelihood for reducing poverty is a strong point in raising the voice for fisherman's.

State of Assam has seen a significant increase in fish productivity [6] which more than that of country's productivity. The author also suggested that the best utilization of ponds and the replacement of traditional fish farming methods with scientific approaches can significantly increase fish production in the village.

Introspects that recording the voice spoken by the user and comparing the level of sound produced by the system and concentrating on the noises are also considered as a part of clearing the noise in the voice data [7]. This study took five different types of voices containing male and female combination. Each one speaks 15 different words and checked whether the target are same or not keeping the training process as a strong input for signal extraction. In this extraction process the source voice is different from the target voice because of the surrounding voices.

Aquaculture of fish species is very important [8] and as a practice that can be used to meet several different objectives such as rearing of ornamental species for home aquaria, educational and recreational purposes, for strengthening stocks of existing fish.

Developed a IOT based system [9] which can help fish farmers intelligently control and manage fishpond water-quality treatment equipment, as well as consumers track and view historical farming process data using a QR code tag on an aquatic product, allowing fish farmers to increase revenue while also ensuring consumer food safety.

Analyzed the data and found out the risk characteristics [10] and calculated accuracy of machine learning algorithms. Admitted that fish farming would become an efficient trade [11], if farmers in Tripura followed scientific methods in fish farming. The authors followed the technique of Random sampling in order to get response from the farmers who mostly use classical methods of fish farming.

Based on Smart Partial Least Squares (PLS) [12] Version 3, the authors confirmed that fish farmers should target high on pricing rather than promotion for achieving more performance in fish trade.

The days of the hunter-gatherer are largely past said [13]. It costs less energy and requires less efforts, to grow crops. Almost anything, that is, except ocean fish. The world's fishing industry is very much alive, but it has approached its limit in harvesting many of the traditional food fishes.

Lake Tanganyika hosts one of the largest inland fisheries in Africa and is a significant source of food and livelihood to millions dwelling inside and outside of its basin said Molsa 2019. The idea behind this is to commercialize the activity as well as a remarkable assemblage of tropical flora and fauna, including highly diverse populations of endemic fish. This idea in the research, shows how geographical locations are responsible in fish cultivation.

In this study [14], presented a review of Automatic Speech Recognition Systems (ASRS) and decided that the most used solution for developing a system is the neural networks using artificial intelligence. This study deals with the two phases of ASRS in detail.

Neural data acquisition with three participants had given three different channel ECoG arrays one and three had left hemisphere coverage and two had right hemisphere coverage. A data acquisition (DAQ) provided to process the local field potentials recorded from these arrays at multiple cortical sites from each participant and similarly the High gamma feature extraction using Real-Time and Network Systems (RTNSR) package implemented a filter chain comprising three processes to measure high gamma activity in real-time [15].

Restrictive Boltzmann Machines (RBM) [24] for Speech Synthesis is used for modelling speech signals, such as speech recognition, spectrogram coding. In these applications, RBM is often used for pre-training of Deep auto-encoders (DAE). Multi-Distribution Deep Belief Networks (DBN) for Speech Synthesis. The Method of modelling the joint distribution of context information models the continuous spectral, discrete voiced and unvoiced (V/UV) parameters and the multi-space simultaneously with three types of RBMs.

Granted out a typical oriented Decision Support System [17] in order to improve the mechanism of fish farming such that production approaches are scientific, natural, cost efficient and organic. The approaches are identified using the algorithm of Particle swarm optimization. The proposed system attains 2 categories of approaches like operational and strategic based on specific circumstances. A smart aquaculture system makes use of available information and resources as well as technology of Artificial Intelligence to boost production efficiency aquaculture systems that require fewer inputs and costs.

Linear Predictive Coding (LPC) for analysis as a method that is based on the source-filter model of the speech signal for signal source modelling in speech signal processing. In linear prediction, the unknown output is represented as a linear sample, and the prediction coefficients [18] are selected to minimize the mean square error. Similarly, the Nearest Neighbor (NN) rule achieves consistently high performance, without a prior assumption about the distributions from which the training

examples are drawn. A new sample is classified by calculating the distance to the nearest training case; the sign of that point then determines the classification of the sample.

Farming applications are being progressed using IOT where the fish farming and production have some certain parameter [19]. The accuracy of the proposed model is found to be quite excellent, which proposes a machine learning-based automated approach for early diagnosis of fish infections [20].

Reservoir authorities [21] should provide proper training in current fishing methods to the fishing community in order to increase the commercial output of fish production.

Fisher has suffered significant losses in assets and other areas due to cyclone. It is necessary to give specialized training programmed to raise knowledge about natural disasters.

The study [22] indicated that there are specific mobile apps in this subject of aqua sector, some of which have high ratings and downloads.

An overview of the marine fisheries [23] sector in Andhra Pradesh. The information in this study is derived from secondary sources. Simple statistical tools, such as percentages and growth rates have been provided.

This study focuses on the basic ideology of Linear Predictive Coding (LPC) [24] with the usage of Mel-Frequency Cepstrum Coefficients (MFCC). In LPC, The basic idea of LPC is letting know the basic precise boundaries of a speech given through audio or text files [25] known as discourse and similarly MFCC is the illustration of LPC about the capabilities used in discourse acknowledgment.

The study suggests that fish farmers construct location-specific mobile apps [24] in a participative manner, which will aid in the dissemination of information in their native languages. Revealed that in the majority [26] of the markets investigated, a lack of proper infrastructure and services, as well as a lack of product diversity, are limiting factors in the fish trade.

Developed automated fish [27] farming monitoring system that saves the farmer time, money, and energy. A model is developed as an energy saving model using ZigBee wireless transmission and MSP430 chip and raspberry pi.

2.3 Major Findings

- Fisherman welfare and holding their traditional skills.
- Fish welfare and nurture the water and land resources.
- The role of geographical location, climate, and heights for fish cultivation.
- Harvesting many of the traditional food fishes.
- Adoption of fisherman's and fish's cultivation by co-management which ends traditional fish farming.
- Speech recognition is done in two phases to extract noiseless data

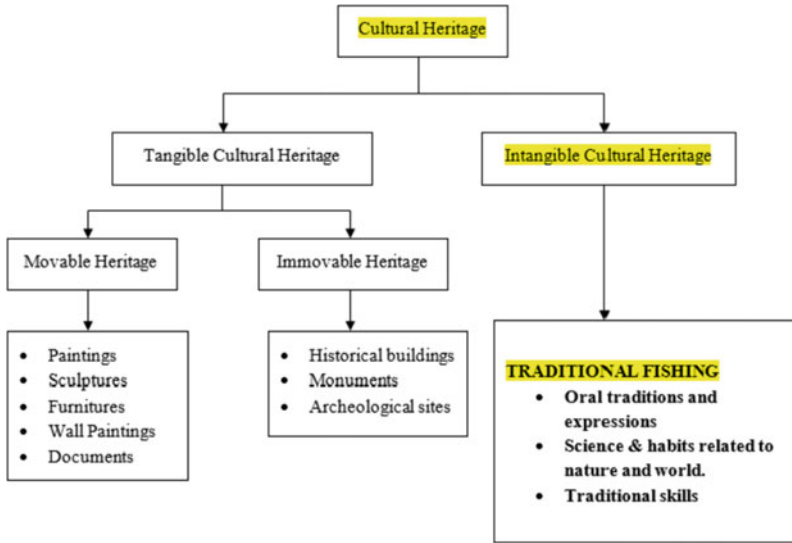


Fig. 1 Cultural heritage and spotting traditional fishing and fish farming

2.4 Relevance for Heritage Science

This research deals with Technologies for IHC. Conservation and Preservation of IHC of fisherman and their fisheries at local fish farms and local wetlands are available. This deals with the study of traditional fish farming skills like Hand gathering, Spearfishing, Netting, Angling and Trapping at various seasons and various geographical locations based on the characteristics of water resources available. Though, these techniques sound to be simple, but these techniques are rich enough in protecting and make the rare species of fish to survive in this aquatic world. This culture and heritage come up with the study of local water bodies and fisheries and small fish farms located in the coastal areas of Srikakulam and Visakhapatnam. Identify, the missing traditional culture as shown in Fig. 2.



Fig. 2 Missing fishing traditional culture

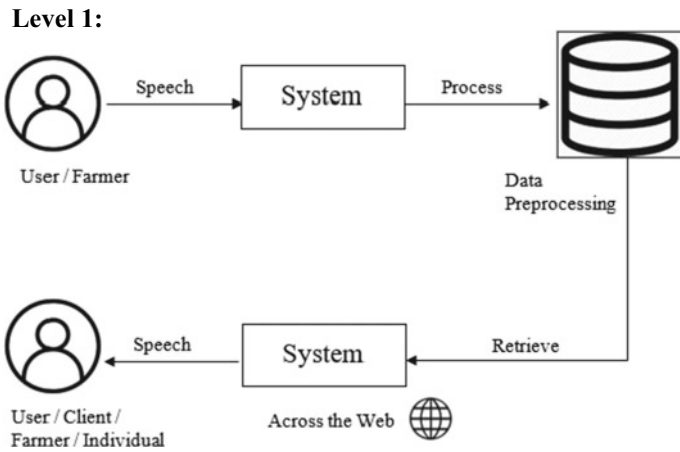


Fig. 3 Access for speech to speech communication and interaction based on the information stored across the web

Level 2:

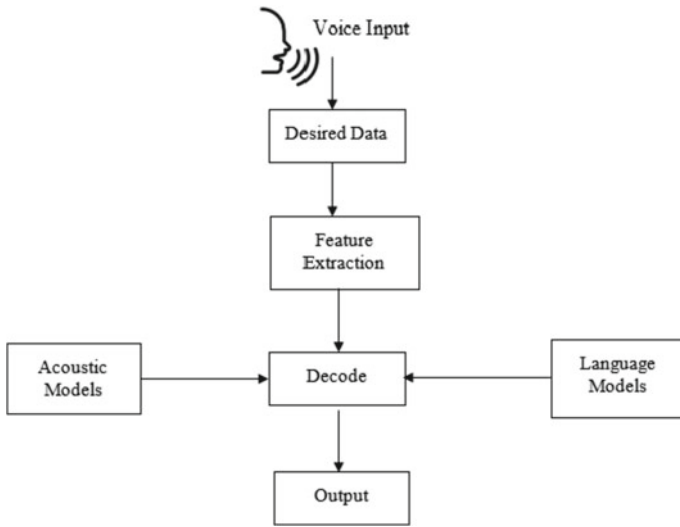


Fig. 4 Detailed procedure for converting the video and audio frames to the database required

Level 3:

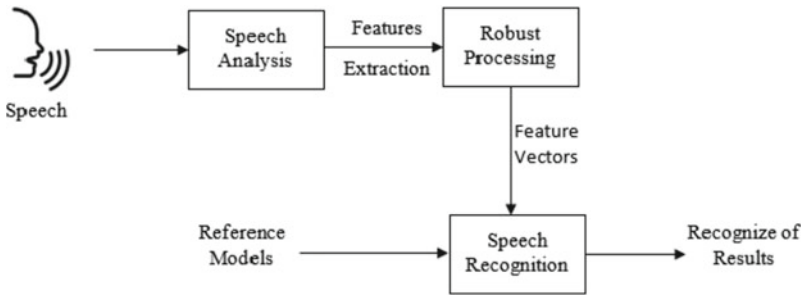


Fig. 5 Detailed procedure for converting the speech to speech on various speech models

Level 4:



Fig. 6 Edge component NLP as a major role in conversion from speech to speech models

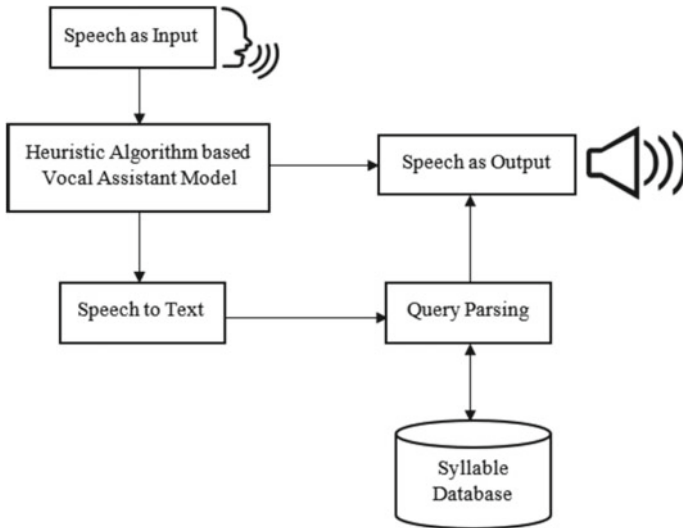


Fig. 7 Detailed step by step representation of voice modules

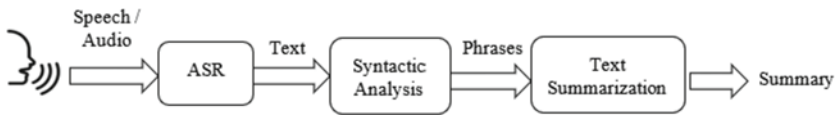


Fig. 8 Brief process of accessing the sound input and convert it into textual information and store in the database

This traditional fishing methodologies are missing because of the Water Pollution, overfishing by third parties or co-management, decrease in the bio mass and marine debris.

3 Scope and Objectives of the Project

- (a) To identify, interrogate and document fishermen’s experience in coastal areas to classify fish farms.
- (b) To extract audio from the document and convert into vernacular languages using Natural Language Processing Techniques.
- (c) To develop an Interactive voice application with a speech corpus for sensitizing the targeted fisherman community.
- (d) Encouraging the identified fisherman community to establish fishing farms as per traditional skills with least financial aid.

3.1 Use Case Scenario

Science and Technology Component is recorded in levels to understand the scope of the project in a technical way.

4 Methodology

4.1 Methodology Detailing Stepwise Activities and Sub-Activities

- **To identify, interrogate and document fishermen's experience in coastal areas to classify fish farms.**

Experienced farmers are investigated, and their information is captured in both audio and video forms by using highly equipped essentials.

- **To extract audio from the document and convert into vernacular languages using Natural Language Processing Techniques.**

NLP is applied on the audio frames by extracting them from the videos using suitable encoders/ decoders for translation of the voice note into vernacular database and store the same in the syllable database.

- **To develop an Interactive voice application with a speech corpus for sensitizing the targeted fisherman community.**

An application is developed using language models and speech synthesizers and convert them into languages related to other residing states of Andhra Pradesh

- **Encouraging the identified fisherman community to establish fishing farms as per traditional skills with least financial aid.**

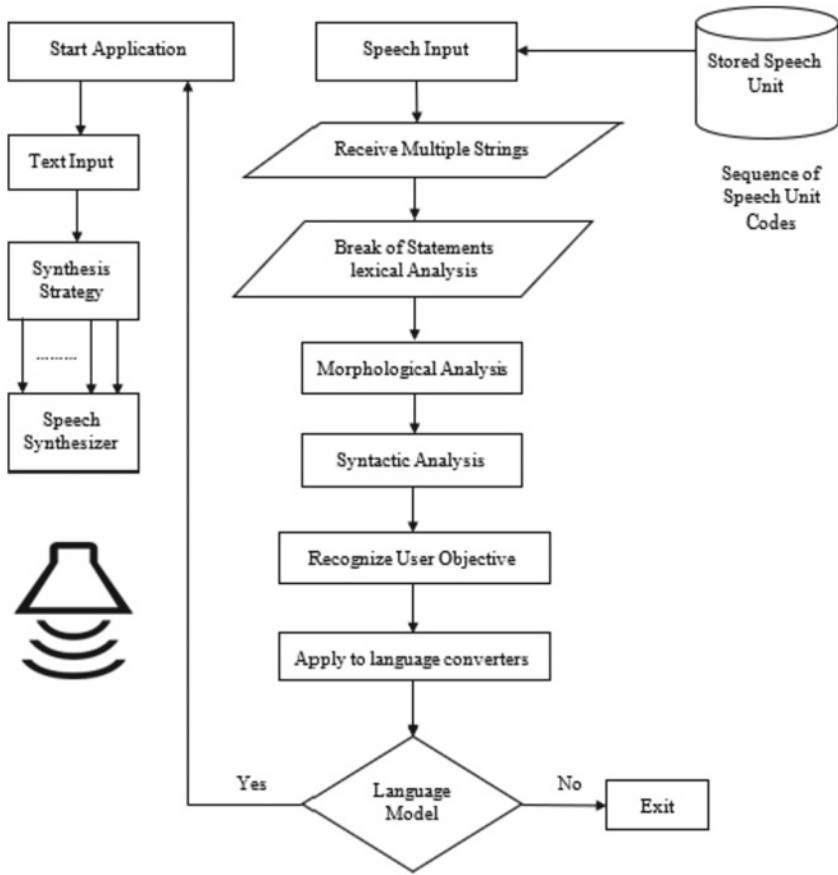


Fig. 9 Step by step detailed process of accessing the sound input and convert it into other language models

The traditional fisherman communities are identified from the region of Kalingapatnam, Magadhalapatnam and Bheemunipatnam and several workshops and events are conducted displaying the expertise videos and enlighten them to continue the traditional fish farming techniques (Figs. 9, 10, 11).

The detailed description of the methodology is represented in the form of a pictorial representation in Fig. 9.

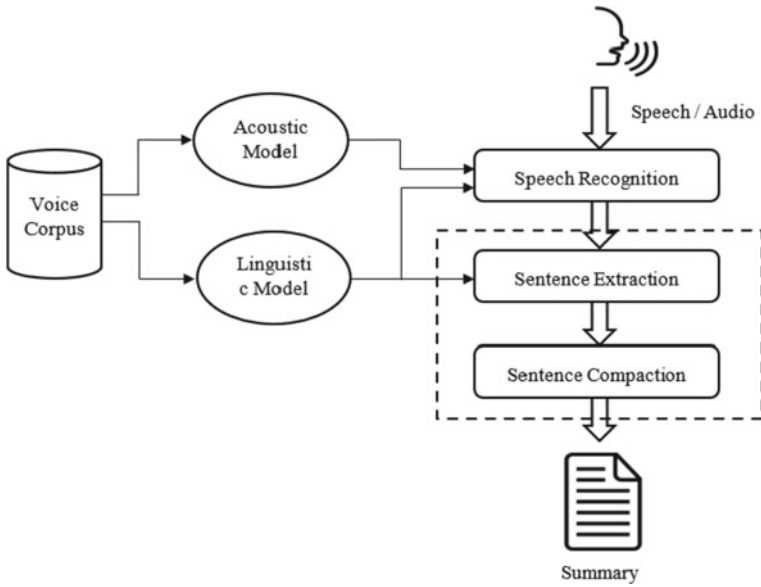


Fig. 10 Extraction module from voice to text

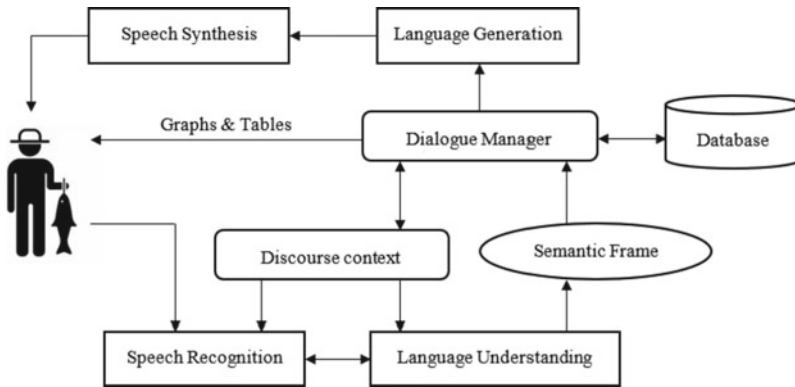


Fig. 11 Framework of the methodology

5 Conclusion

An Interactive voice enabled application to enlighten the community of the farmers to enrich in fish farming using traditional approach of the preserved content.

References

1. Coche AG (1967) Fish culture in rice fields a world-wide synthesis and Mustafa, Faizan. (2016). A review of smart fish farming systems. *J Aquacult Eng Fisher Res* 193–200
2. Aejas B, Bouras A, Belhi A, Gasmi H (2021) Named entity recognition for cultural heritage preservation, data analytics for cultural heritage, pp 249–270
3. Bashar A (2019) Survey on evolving deep learning neural network architectures. *J Artif Intell* 1(02):73–82
4. Mili S, Ennouri R, Laouar H (2021) Freshwater fish farming and fishery management in tunisian reservoirs: limitations and opportunities
5. Satia BP, Njifonjou O (2021) Fisheries co-management and poverty alleviation in the context of the sustainable livelihood approach: a case study in the fishing communities of Aby Lagoon in Côte d’Ivoire
6. Barasha Rani Das (2020) Trend of fish production in a flood plain village of Kamrup(M) district, Assam, India. *Int J Adv Res Eng Technol (IJARET)* 11(11):820–822
7. Mukhneri FM, Wijayanto I, Hadiyoso S (2020) Voice conversion for dubbing using linear predictive coding and hidden markov model. *J Southwest Jiaotong University*, 55(5)
8. Rito J, Viegas I (2020) Fish farming. In: Leal Filho W, Azul AM, Brandli L, Lange Salvia A, Wall T (eds) *Life below water. Encyclopedia of the UN sustainable development goals*. Springer, Cham.
9. Gao G, Xiao K, Chen M (2019) An intelligent IoT-based control and traceability system to forecast and maintain water quality in freshwater fish farms. *Comput Electron Agricult* 166:105013. ISSN 0168-1699
10. Ahmed M, Rahaman MO, Rahman M, Abul Kashem M (2019) Analyzing the quality of water and predicting the suitability for fish farming based on IoT in the context of Bangladesh. In: 2019 international conference on sustainable technologies for industry 4.0 (STI), 2019, pp 1–5
11. Borah YJ, Singh A, Sarkar SK, Singh P, Pal ON, Khuman CP (2019) Adoption of scientific fish farming practices in West Tripura district of Tripura, India. *Patnagar J Res, ResearchGate* 17(2)
12. Fitriah A, Rosdi S, Mohamad M, Mustapha NA, Abdul Aziz Z, Ibrahim W, Md Radzi MS, Aniesha N, Yaacob A (2019) The effects of marketing mix on small fish farming business performance 6:1–16
13. Paulin NS, Anupriya N, Prasanthi S (2017) Pisciculture environment control using automated monitoring system. *Asian J Appl Sci Technol (AJAST)* 1(2):60–65
14. Benk S, Elmir Y, Dennai A (2019) A study on automatic speech recognition. *J Inf Technol* 10(3):77–85
15. Moses DA, Leonard MK, Makin JG, Chang EF (2019) Real-time decoding of question-and-answer speech dialogue using human cortical activity. *J Nat Commun*
16. Ning Y, He S, Wu Z, Xing C, Zhang L-J (2019) A review of deep learning based speech synthesis. *J Appl Sci*
17. Cobo A, Llorente I, Luna L, Luna M (2018) A decision support system for fish farming using particle swarm optimization. Elsevier
18. Raj JS, Vijitha Ananthi J (2019) Recurrent neural networks and nonlinear prediction in support vector machines. *J Soft Comput Paradigm (JSCP)* 1(01):33–40
19. Shinde S, More S, Paithane P Rane S (2021) IOT based fish farming and its applications, May 7 2021
20. Nayan A-A, Mozumder AN, Saha J, Mahmud KR, Al Azad AK (2020) Early detection of fish diseases by analyzing water quality using machine learning algorithm. *Int J Adv Sci Technol* 29(5):14346–14358
21. Sridhar D, Ramaneswari K (2015) An introductory study on fishing craft and gears in Thotapalli and gotta reservoirs, of Vizianagaram and Srikakulam districts of Andhrapradesh. India, *Int J Adv Res* 3(8):1980–1984

22. Dhenuvakonda K, Sharma A, Pani Prasad K, Sharma R (2019) Socioeconomic profile of fish farmers of Telangana and usage of mobile apps. *Asian J Agricult Extens, Econ Sociol* 37(3):1–9, 2019; Article no. AJAEES.52753 ISSN: 2320-7027
23. Babu KM, Sudhakara Rao B (2020) Growth and performance of marine fisheries sector in Andhra Pradesh. *Int J Southern Econ Light (JSEL)*| SJIF Impact Factor: 6.244, 7
24. Shanthi H, Pasumarthi RG, Suneel Kumar P (2020) Estimation of speech parameters using linear predictive coding (LPC). *J Compos Theory* 8(7)
25. Dhawale AD, Kulkarni SB, Kumbhakarna V (2019) Survey of progressive era of text summarization for indian and foreign languages using natural language processing. In: *International conference on innovative data communication technologies and application*. Springer, Cham, pp 654–662
26. Raju SS, Pattnaik P, Shyam SS, Narayanakumar R (2021) Impact of Titili cyclone on marine fisheries of Srikakulam district of Andhra Pradesh. *J Marine Biol Res* 1(2):1–10
27. Usha Kiruthika S, Kanaga Suba Raja S, Jaichandran R (2017) IOT based automation of fish farming. *J Adv Res Dyn Control Syst* 9(1)

Krishi Mitra: Crop and Fertilizer Recommendations System Using Machine Learning Algorithm



Vijay Mane, Akash Gajbhiye, Amisha, Chinmay Deshmukh, and Kunal Gaikwad

Abstract Farming is a livelihood for many people in India, especially in rural parts. Many people in this field are unaware of technological innovations and keep using old farming techniques and the same cropping pattern. But in today's highly technological world, we can use various technologies to help them to select an appropriate crop according to soil condition, good fertilization for farms etc. We can use the data such as the amount of Nitrogen, Phosphorus, Potassium, rainfall and other environmental parameters to select the most suitable crops. In this paper we are training a machine learning algorithm to predict the crop, also we suggest fertilizers based on N, P, K and crop values. The model trained using various machine learning algorithms. Then we chose a model trained by random forest because of its high accuracy. Lastly, integrated the model with a website using Flask API.

Keywords Random forest · Machine learning · Flask · API

1 Introduction

India is mostly a farming country. Agriculture employs around 70% of our people. One-third of our national GDP comes from agriculture. The progress of agriculture development is critical to the economic welfare of our country. Our country is now food grain self-sufficient. Agriculture will transform drastically in the next few years.

V. Mane (✉) · A. Gajbhiye · Amisha · C. Deshmukh · K. Gaikwad
Department of Electronics Engineering, VIT, Pune, India
e-mail: vijay.mane@vit.edu

A. Gajbhiye
e-mail: akash.gajbhiye18@vit.edu

Amisha
e-mail: amisha.singh18@vit.edu

C. Deshmukh
e-mail: chinmay.deshmukh18@vit.edu

K. Gaikwad
e-mail: kunal.gaikwad18@vit.edu

There are many farmers Even small-scale farmers in India lack access to knowledge and technology resources that could help them promote growth and achieve higher pricing for their commodities. Data on farming can be found in a variety of formats and structures, including written media, Radio, newspapers, television, the web, cell devices, and so on. While processing data, various manual procedures may be required to convert data from one format to another [1].

Many farmers in India are unaware of the outside world or agricultural advancements in technology. The majority of farmers have no notion how much their crops and products are worth, therefore they sell them at any price. So, we have come up with a solution that will help and guide the farmers about the growth of their crops. We have created a web platform that will take different parameters that are important for farming as input and by using the ML (machine learning) we will suggest the farmers about the best crop they can grow in given conditions which will increase their yield indirectly resulting in more income. We will also suggest different methods which will help them in more growth of their crops based on the different parameters of soil [2].

2 Related Works

The authors have implemented a model which will help farmers in crop selection based on environmental and geographical factors [3]. In their approach they have used different machine learning algorithms such as KNN, Random Forest, Decision tree and neural networks. Their system architecture has two subsystems 1. Crop suitability predictor and 2. Rainfall predictor.

A comparative study has been carried out between 4 commonly used algorithms such as logistic regression, Naive bayes, decision table and random forest [4]. They did analysis of F1 score, accuracies, recall and confusion matrix. Using the results, they got they classify crops into 4 seasons i.e., kharif, Rabi, summer, and whole year. They tested the proposed model on a dataset for Uttar Pradesh and Karnataka mainly using different statistical methods.

An android application developed with integrated a machine learning model with the application by [5]. Their application has two modules: prediction and fertilizer. They also predicted the amount of yield per square hector. They have used ANN, SVM and RF. Also, they got the highest accuracy for RF.

The authors of [4] did not implement their model on any software platform while authors of [3] and [5] did the experimental implementation on android and Web. In [1] they have implemented Maps feature in their system which can display which crops are being cultivated by the neighbouring farmers in order to help the farmer (user) take better decision for their crops.

3 Methodology

3.1 Gathering Data

In this project our aim is to build a machine learning algorithm which can recommend crops based on inputs about soil composition and climatic conditions and also to recommend fertilizer on the basis of N, P, K values and crop. We need a dataset with a high number of values for N, P, K, rainfall, and other parameters in order to develop a viable model. We tried to find such a dataset on the internet, and we found one from one of the open competitions on KAGGLE [6].

3.2 Data Preprocessing

We worked on the pre-processing of the dataset which included data cleaning, filling up the missing values and so on. Final dataset of crop recommendation contains N, P, K, Humidity, Temperature, Rainfall, Crop name with total entries of more than 2000. While the second dataset of fertilizers contains Crop and N, P, K values for 22 labels of different crops [7].

3.3 How Machine Learning is Used to Predict Crop?

We have used a supervised machine learning to train our model where we splitted our data in 80% for training purposes and 20% for testing. So, 1600 entries were used for training purposes and remaining for testing with the same 22 labels of the fertilizer dataset. In this case, the outcome variable is the crop name which is predicted from independent variables such as values of Nitrogen, Phosphorus, Potassium, ph level, temperature, rainfall and humidity. We have tested 5 different algorithms namely Decision Tree, SVM, Naïve bayes, Logistic regression and random forest. In the dataset used for each crop there are multiple data entries using which this model is trained [1, 8]. The proposed system uses Random Forest algorithm amongst the rest to predict based on the values inputted by the user.

3.4 Classification by Random Forest

Random forest can be used as a classification algorithm. It consists of many decision trees. Using RF decision values can be predicted from the class variable for a qualitative or a categorical type of variable. This is done by the calculation with one or more independent or predictor variables. Application of RF can be seen in

fields like computer science (Data Structure), medicine (diagnosis), botany (classification), and psychology (decision theory), etc. It helps to interpret data easily by visualizing graphically [1, 2]. To predict a qualitative response, a variety of relevant categorization algorithms, or classifiers, can be used. For e.g., Logistic regression, k-nearest neighbours, Bayes’ theorem, etc. Two of the most significant characteristics of classification trees are hierarchical property and flexibility.

3.5 Choosing the Right Model

Looking at the accuracy of different classifiers, Naive Bayes and Random Forest had the highest accuracy of 99.09% as shown in Figs. 1 and 2. While the F1 score of Naïve Bayes and Random Forest were highest among all with score of 0.99. For our project we used a model which is developed using random forest because of following reasons.

- Random Forest is well suited for working with Numerical and Categorical data.
- Among all algorithms, random forest was considered to give the outstanding results in terms of accuracy.

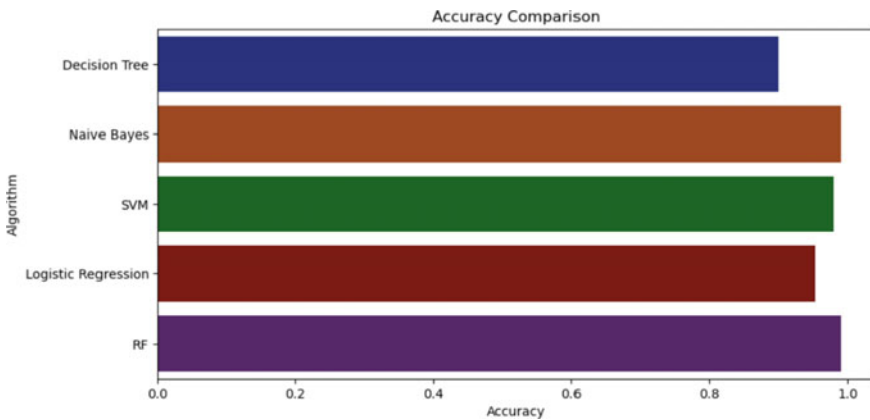


Fig. 1 Graphical format of accuracies

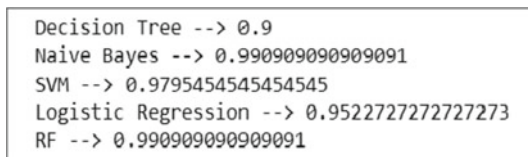


Fig. 2 Numerical format of Accuracies

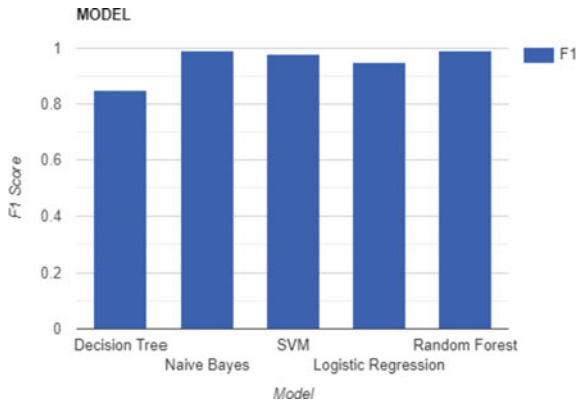


Fig. 3 F1 scores versus models

- It is the most well organised algorithm for working on huge data. Even if it accepts hundreds or thousands of data then also, we don't have to trim or cut any input for getting the desired result.
- During the forest creation phase, the random forest algorithm generates internal reliable estimates of the generalisation error. These estimates are used to showcase the response to increasing the number of features used in splitting and are also used to measure variable importance.
- The produced forest has the capacity to perform well enough in the future as data is added.

Figures 1 and 2 represents the comparison of accuracies from different algorithms in two different forms.

As shown in Fig. 3, we have calculated the F1 score which is generally calculated using the following equation:

$$F1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{1}$$

The F1 score is a measure of model's accuracy on a dataset or is harmonic mean of precision and recall. F1 score ranges from 0 to 1 where 1 being the best score.

3.6 Generating a Model File

As our aim is to integrate trained model with a website, we have to generate a model file. For that we used a pickle module from Python. Pickling is beneficial for applications that require some level of data permanence. The data is stored in a file, so that we can use it later whenever we need. It is also essential to save python objects

in a database or transfer files over a Transmission Control Protocol (TCP) or socket connection. Pickle comes in handy when dealing with machine learning algorithms and you need to save them so you may make new predictions later without having to reconstruct the model or train the model from scratch. As we have implemented random forest, pickle file of random forest is generated.

3.7 Fertilizers Suggestion

To suggest fertilizer based on input values of N, P, K and Crop name given by user, we have created a python dictionary. In this dictionary there are 9 cases mentioned which includes all the possible real situation regarding the level of N, P, K content of soil and this is done by studying required fertilization for soil to make it fertile and balanced for crop production. The suggestion is given by checking the difference of input value from required value. **For e.g.**, if user input shows that soil have low content of nitrogen and high content of Phosphorous and Potassium, then user gets suggestion related to N_Low, P_High and K_High. N_Low, P_High and K_High are the cases from dictionary.

3.8 Integration of Model with Website

3.8.1 Role of Flask

We have used flask API to build our website. Flask is an open-source micro web application framework written in Python.

- Flask reduces development time and allows programmers to build faster and smarter.
- Flask has a lightweight and versatile design, making it simple to change it into the web framework developers need with just a few extensions.
- Extensible framework
- Fast debugging.
- Highly secure
- In a production environment deployment of flask is comparatively easy.

The website is mainly made with HTML, CSS and JavaScript. HTML is an abbreviation of Hyper Text Markup Language. It is used for creating web pages. CSS is used for designing of this webpages and JavaScript helps us to make webpages interactive (Fig. 4).

As shown in Figs. 5 and 6, these pages have input forms written in HTML which posts all the input data to a model file and python dictionary respectively.

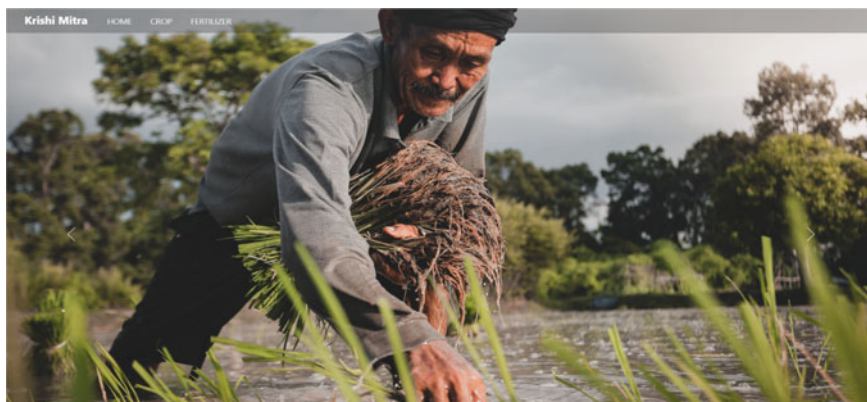


Fig. 4 Homepage of website

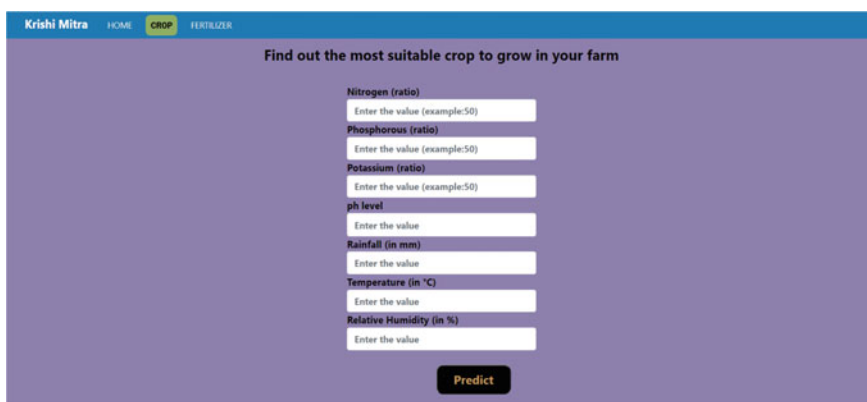


Fig. 5 Crop recommendation page

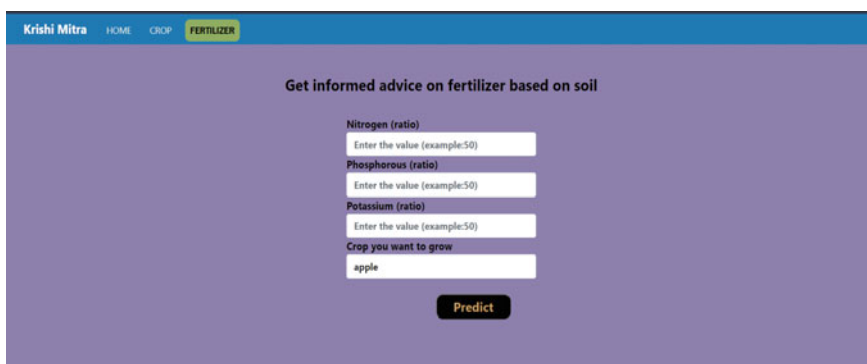


Fig. 6 Fertilizer suggestions page

4 Results

We have made a fully functional web platform for farmers where different parameters such as values of Nitrogen (N), phosphorus(P), potassium(K), ph value(0–14) of soil, humidity of soil and rainfall will be given as input to the site and after the processing it will suggest the best crop to grow under the given conditions. The platform will also give different suggestions/methods to improve the crop growth (Figs. 7 and 8).

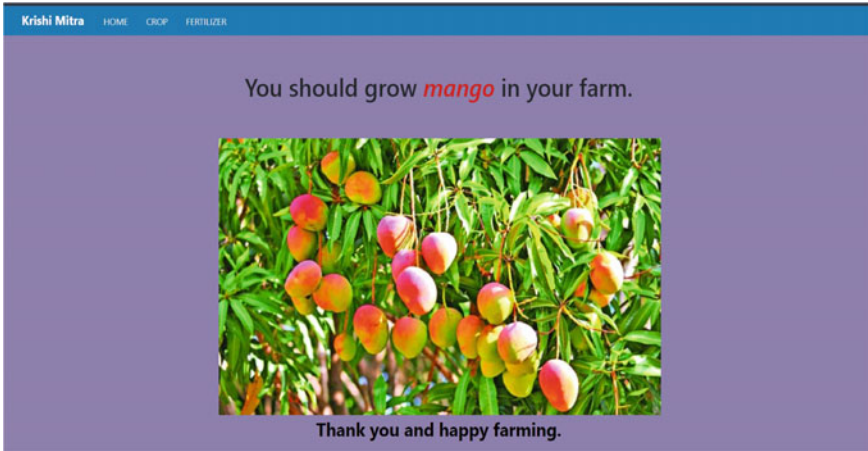


Fig. 7 Crop recommendation result page

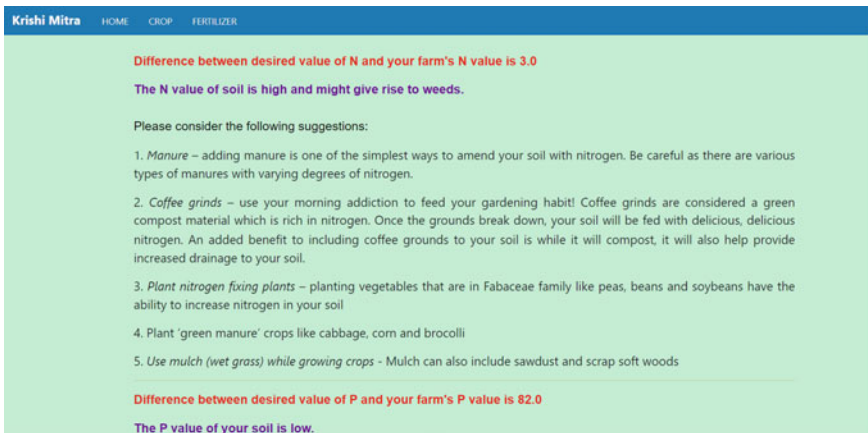


Fig. 8 Fertilizer suggestion result page

5 Conclusion

We have successfully implemented a Random Forest classifier for Crop recommendation and recommend fertilizer based on N, P, K values and crop with high accuracy. To make our work reachable to users we have implemented it using website and website is hosted so anyone can access it. As we know ground level reality that many farmers are not well acquainted with technology. So looking at this we can integrate our software with hardware sensors and take reading directly into the software algorithm with sensors rather than having the farmers put it manually and thus making it an almost fully automated system.

References

1. Manoharan JS (2021) Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm. *J Soft Comput Parad (JSCP)* 3(02):83–95
2. Bashar A (2019) Survey on evolving deep learning neural network architectures. *J Artif Intell* 1(02):73–82
3. Doshi Z, Nadkarni S, Agrawal R, Shah N (2018) AgroConsultant: intelligent crop recommendation system using machine learning algorithms. Fourth international conference on computing communication control and automation (ICCUBEA) 2018:1–6. <https://doi.org/10.1109/ICCUBEA.2018.8697349>
4. Katarya R, Raturi A, Mehndiratta A, Thapper A (2020) Impact of machine learning techniques in precision agriculture. In: Proceeding 3rd international conference on emerging technologies in computer engineering: machine learning and internet of things (ICETCE), pp 1–6. <https://doi.org/10.1109/ICETCE48199.2020.9091741>
5. Pande SM, Ramesh PK, Anmol A, Aishwarya BR, Rohilla K, Shaurya K (2021) Crop recommender system using machine learning approach. In: Proceeding 5th international conference on computing methodologies and communication (ICCMC), pp 1066–1071. <https://doi.org/10.1109/ICCMC51019.2021.9418351>
6. Kumar R, Singh MP, Kumar P, Singh JP (2015) Crop selection method to maximize crop yield rate using a machine learning technique. In: Proceedings of international conference on smart technologies and management for computing, communication, controls, energy, and materials (ICSTM). IEEE, pp 138–145
7. Kale SS, Patil PS (2019) A machine learning approach to predict crop yield and success rate. In: 2019 IEEE pune section international conference (PuneCon). IEEE, pp 1–5
8. Nishant PS, Venkat PS, Avinash BL, Jabber B (2020) Crop yield prediction based on Indian agriculture using machine learning. In: Proceedings of International conference for emerging technology (INCET). IEEE, pp 1–4
9. Bashar A (2019) Survey on evolving deep learning neural network architectures. *J Artif Intell* 1(02):73–82
10. Jaiswal JK, Samikannu R (2017) application of random forest algorithm on feature subset selection and classification and regression. In: Proceedings of world congress on computing and communication technologies (WCCCT)

Development of DDoS Attack Detection Approach in Software Defined Network Using Support Vector Machine Classifier



Oluwashola David Adeniji , Deji Babatunde Adekeye,
Sunday Adeola Ajagbe , Ademola Olusola Adesina ,
Yetunde Josephine Oguns , and Matthew Abiola Oladipupo 

Abstract Software defined networking (SDN) is an approach to network management that enables the network architecture to be centrally and intelligently controlled with the use of software applications. It is an innovative way in which networks are built by separating the control plane of network devices and data plan and thereby making their management easier through centralized control. The design of SDN allows a secured flexible network configurations. The major drawback in SDN is the formation of Distributed Denial of Service (DDoS) attacks. The DDOS introduces flood with large numbers of malicious packets from malicious hosts that consume network resources. However, the trust between the control planes and forwarding planes is the vulnerability that can be exploited by attackers to perform distributed denial of service attacks in an SDN. Also, there is no form of authentication or authorization between the controllers and switches before a flow is passed between the layers. This problem ultimately causes a denial of access to legitimate network users or degrade the performance of the network. The objective of this paper is to develop an approach to detect DDos Attacks in IPv6 Enabled SDNs using the Support Vector Machine (SVM) algorithm. The normal and attack traffic packets generated was 500,000 packets during the test for 20 min. The packets were pre-processed and the

O. D. Adeniji · D. B. Adekeye

Department of Computer Science, University of Ibadan, Ibadan, Nigeria

e-mail: od.adeniji@ui.edu.ng

S. A. Ajagbe (✉)

Department of Computer Engineering, Ladoke Akintola University of Technology LAUTECH, Ogbomoso, Nigeria

e-mail: saajagbe@pgschool.lautech.edu.ng

A. O. Adesina

Department of Mathematical Sciences, Olabisi Onabanjo University, OOU, Ago-Iwoye, Nigeria

e-mail: ademola.adesina@oouagoiwoye.edu.ng

Y. J. Oguns

Department of Computer Studies, The Polytechnic Ibadan, Ibadan, Nigeria

e-mail: oguns.yetunde@polyibadan.edu.ng

M. A. Oladipupo

University of Salford, Greater Manchester, Salford M5 4WT, UK

e-mail: m.a.oladipupo@edu.salford.ac.uk

SVM was trained with 25% of the normal and the attack data. The SVM achieved an accuracy of 99.69% and a DDoS attacks detection rate of 100%.

Keywords Distributed denial of service (DDoS) attacks · Software defined networking (SDN) · Support vector machine (SVM) · Open flow · Internet protocol version 6 (IPv6) · Exploit

1 Introduction

Software defined networking (SDN) concept provides a granular approach to security which is determined by centralizing the security control into one entity like the SDN controller. With the standardization of Open Flow protocol defined by the open networking foundation, networking equipment's from different vendors can work within an SDN architecture. The new paradigm for applications to interact with the network call for a declarative abstract within the SDN [1]. The APIs can be configured and operate within the network. In order to plan and optimize operations, the APIs query the network for data. SDN computes and abstracts heterogeneous computing resources in terms of capacity and capability. The SDN can provide network virtualization, define network policies and have complete control over the entities within a network [2, 3]. It is also worthy of note that, DDoS assaults post major concern in the current SDN because SDN controllers are vulnerable sites of failure throughout the SDN architecture. Recently, research on DDoS attack detection in SDN has concentrated on determining how to exploit data plane programmability, enabled by the P4 language, to detect attacks directly in network switches, with just a minor role for SDN controllers [18].

Centralized control and network programming implementation tackle the interdependence problem in three planes: control plane, data plane, and application plane. Northbound API connects the control plane to the application plane. The southbound API primarily consists of the Open Flow protocol, which allows the control plane to interface with the data plane [4]. SDN liberates network architecture from the constraints of traditional network complexity and coupling, allowing it to meet flexibility, reliability, and security all at the same time. It separates the control plane from the data plane, as well as the network's control function from the data forwarding function [5]. The open networking foundation (ONF) defines SDN as the physical separation of the network control plane from the forwarding plane, with a control plane in charge of many devices. Arising from the problem that will ultimately cause a denial of access to legitimate network users or degrade the performance of the network. Hence, the research. The objective of this study, therefore is to develop an approach to detect DDoS Attacks in IPv6 Enabled SDNs using a machine learning (ML) algorithm, support vector machine (SVM) algorithm. Other parts of this study are organized in the following ways, Sect. 2 literature review while Sect. 3, reported the methodology and testbed design specification, Sect. 4 contains the implementation and results. Finally, in Sect. 5 conclusion of the study is presented.

2 Literature Review

The flooding of UDP, ICMP, SYN attack, HTP, ping of death attack, smurf attack, and Slowloris attack are the most common types of DDoS attacks. The flooding of UDP is a type of flooding attack in which a malicious system generates multiple UDP datagrams [6]. The flooding of ICMP attacks can exploit by consuming a large number of ICMP pings [7]. The flooding of HTTP attack can be exploited by using legitimate GET [8]. The Smurf assault, also known as the ICMP amplification attack, is one of the most popular DDOS attacks in IPv4 networks and still occurs in IPv6 networks [9]. Slow HTTP DoS attacks, such as Slowloris and Slow POST, target the application layer, in this example the HTTP protocol, and can be implemented in a number of ways [10]. The flooding of the SYN attack makes use of a typical TCP implementation practice of allocating a connection state structure when a TCP SYN segment is received, while also restricting the number of connections in a certain subset of states for a given application [11].

The intrusion detection systems, such as host-based intrusion detection systems (HBIDS), are a type of technique that is installed on and monitors a single host machine. Analyzes system activity for events or groupings of events that fit a specified pattern of events describing a known attack. The core concept is to employ knowledge of known attack patterns to discover attacks in various sources of data that are being watched. The tradeoff between the two protocols has the potential to have a big influence on networks. Furthermore, one potential choice of selecting any of the protocols that can increase or decrease the degree of application is used as adapted in [12]. However, SDN implements network protocols that take years of testing, standardization and interoperability in [13]. The present revolution and innovation in networking are the SDN. The design in SDNs is focused on the required topology that performs the programming job. With a better control plane, the packet handling rate should be scalable with the number of CPUs. It is preferable for the controllers to access the network status at the packet level at all times [3].

In the research work done by [14] they discussed about an Advance SVM approach that detects DDOs attacks but the implementation procedure was not discussed. The ASVM technique is a multiclass classification method consisting of 3 classes. The results were evaluated by measuring false alarm rate, accuracy and detection rate. The results show their detection technique was 97% accurate. The only drawback to their research work was that it was conducted on IPv4, SDN Network [15]. A similar research was done by [15]. They analyzed the SVM classifier and compared it to Radial Basis Function Network, Naives Bayes, Bagging, J48 Decision Tree, and Random Forest. The research showed SVM had the highest percentage of accuracy. The project is well structured and organized but was based on DARPA 2000 dataset which contains outdated attack methods. Another worthy mention to this research is [12], which was focused on using SVM for the Detection of DDoS attacks in SDN on an IPv4 network. Their classification generated a 95.24 detection accuracy [16]. Also, in the research work [13, 14] they discussed about using the Random Tree

Forest Algorithm for classifying DDoS attacks in an SDN networks. Their research work got a 95.24 detection rate in the classification of normal and abnormal traffic.

Kokila et al. [19] discovered that it was vital to detect the attack in the controller at an earlier stage because the controller of SDN is vulnerable to DDoS attacks, which lead to resource exhaustion, resulting in the inability to reach the controller's services. The DDoS detection necessitates a flexible and accurate classifier that makes decisions based on unclear input. SVM is a popular classifier that has high accuracy and a low false-positive rate. The study examined the SVM classifier and compare it to existing DDoS detection classifiers. Experiments demonstrate that SVM outperforms others in classification accuracy.

The DDoS assaults post major security concerns in the current SDN because SDN controllers are vulnerable sites of failure throughout the SDN architecture. Recently, research on DDoS attack detection in SDN has concentrated on determining how to exploit data plane programmability, enabled by the P4 language, to detect attacks directly in network switches, with just a minor role for SDN controllers. To address the issue, this study explored an ML algorithm, SVM as a detector/classifier. The success of SVM relating to classifications in previous studies such as [19] informed its choice in this experiment.

3 Methodology

The SVM approach was used in this study to learn how to classify the data due to its success in many pieces of research including mortality prediction research by [17]. SVM classifier is a supervised ML method that separates classes by a hyperplane. The SVM learns from the data that is provided to and trained with, and it compares the data assigned to the algorithm with the data used to train it. It generates a pattern map that distinguishes between regular traffic and attack traffic aspects. In this research work, the algorithm was implemented virtually using the Oracle Virtual Box. The Ubuntu 18.04 was installed in Virtual Machine to create the simulation's operating environment. Other technologies and tools were employed to put the methodology into action.

- The OpenFlow is a programmable network protocol that uses standardization to govern and guide traffic between routers and switches from diverse vendors. It abstracts the router and switches forwarding from the underlying hardware. Mininet is a network emulator that creates a virtual network of hosts, controllers, and links. Mininet switches support OpenFlow for highly flexible custom routing and SND, and Mininet hosts run standard Linux network software.
- The Mininet facilitates research, development, prototyping, learning, debugging, testing and any other tasks that would benefit from having a fully functional experimental work on a laptop machine to put SDN into action. It also provides a low-cost platform for developing, testing, and building unique network topologies.

- The Ryu Controller is an open SDN controller that simplifies traffic management and adaptation to increase network agility. The SDN Controller is the system’s brain, sending data down to switches and routers via southbound APIs and up to applications and business logic via northbound APIs. The RYU controller was installed and placed in NTT cloud data centers. The network can adapt to the demands of its applications as they change.
- The Wireshark is an open-source program for capturing packs and logging traffic over a computer network. It is used for troubleshooting, analyzing, protocol development and research. Wireshark is cross-platform and it has a GUI interface and can also be run with a terminal-based version called T Shark.
- The Iperf is a network performance monitoring and optimization tool. It’s a cross-platform utility that can generate standardized network performance measurements. Iperf can create data streams to measure the throughput between the two endpoints in one or both directions, and it provides client and server features. A time-stamped report of the amount of data transported and the throughput measured is typical of Iperf output. Both IPv4 and IPv6 are supported.

For evaluating and identifying DDoS assaults, the following traits and parameters are watched and gathered. As shown in Fig. 1, the incoming traffic goes into the system and the traffic collector takes its input from incoming traffic and sends its output to statical monitoring, which in turn supplies and combines with data collection. The storage data extraction was the next stage and relevant feature was extracted from the data. The SVM classifier was then trained on the machine and prediction was done based on the default parameters and the prediction results were obtained.

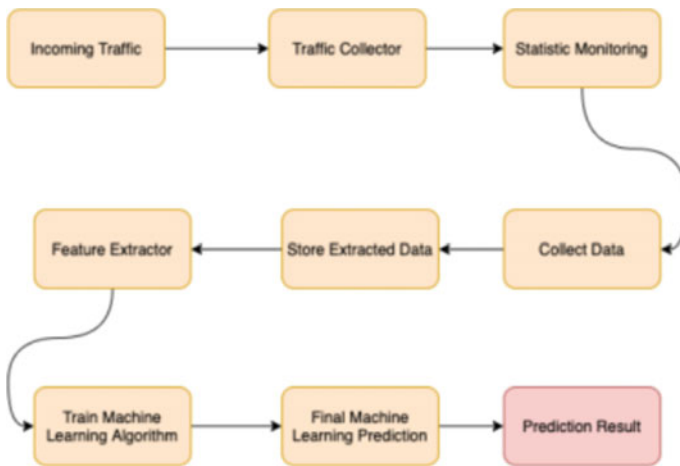


Fig. 1 Study approach

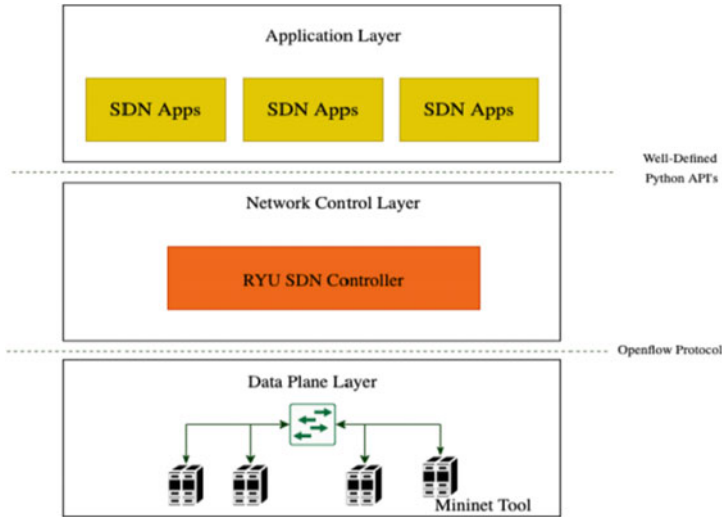


Fig. 2 SDN network architecture

3.1 Testbed Design Specification

The architecture of a standard SDN is shown in Fig. 2, it includes the Application, Network control, Data plane layers. The data plane in the proposed SDN framework has multiple hosts that are virtually created using mininet and are all connected to the Open Flow switch, which defines the SDN protocols and the Open Flow protocol that communicates with the framework’s control plane. The data plane and switches are controlled by the control plane, which also defines rules and monitors network traffic flow. The RYU controller is a simulation-based controller that provides configuration capabilities and allows us to control network routing operations. The control plane is written in Python since the RYU is a Python-based controller that communicates with the application layer via a Python-based API, which in this case is network traffic apps. The choice of this implementation environment was informed by the easy simulation attributes of this environment in an experiment of such nature.

Figure 3 shows the network topology constructed with the Mininet network simulation the network has ten nodes or hosts and one single Open Flow switch coupled to one RYU controller. The switch is connected to the controller and the switch is connected to all of the hosts. The RYU controller is in charge of all of these hosts and switches, and it collects all of the network’s flow data.

The RYU Controller started at port 6653 and Mininet creates one virtual switch and 10 hosts, a ping is run to generate normal traffic for 20 min as Figs. 4 and 5 shows the creation of topology and generation of normal traffic and attack traffic Log by the RYU controller in the research respectively.

The SDN centralized management mechanism was introduced to prevent the network from unauthorized access. In a standard network environment, every

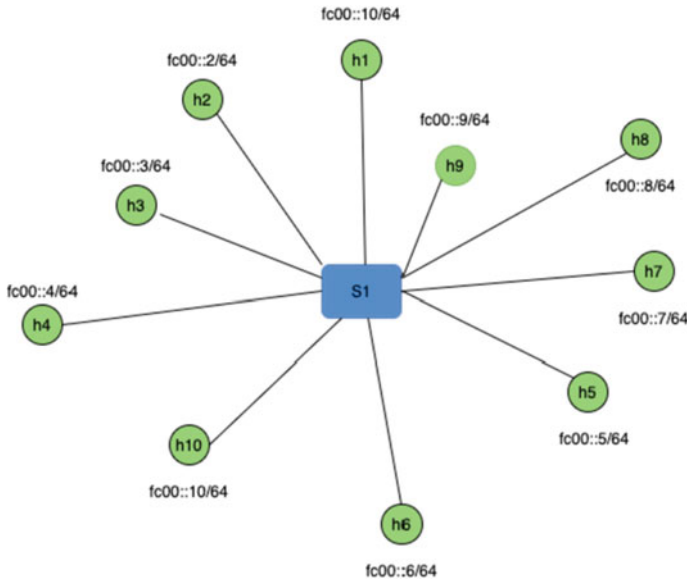


Fig. 3 Network topology showing the controller and hosts

```
(base) spazwtk@spazwtk-VirtualBox:~/ddostest$ conda deactivate
spazwtk@spazwtk-VirtualBox:~/ddostest$ sudo python topo_normal.py
[sudo] password for spazwtk:
connecting to remote controller at 127.0.0.1:6653
*** Creating network
*** Adding controller
*** Adding hosts:
h1 h2 h3 h4 h5 h6 h7 h8 h9 h10
*** Adding switches:
s1
*** Adding links:
(5.00Mbit) (5.00Mbit) (h1, s1) (5.00Mbit) (5.00Mbit) (h2, s1) (5.00Mbit) (5.00Mbit) (h3, s1) (5.00Mbit) (5.00
Mbit) (h4, s1) (5.00Mbit) (5.00Mbit) (h5, s1) (5.00Mbit) (5.00Mbit) (h6, s1) (5.00Mbit) (5.00Mbit) (h7, s1) (
5.00Mbit) (5.00Mbit) (h8, s1) (5.00Mbit) (5.00Mbit) (h9, s1) (5.00Mbit) (5.00Mbit) (h10, s1)
*** Configuring hosts
h1 h2 h3 h4 h5 h6 h7 h8 h9 h10
*** Starting controller
c1
*** Starting 1 switches
s1 ... (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.00Mbit) (5.0
0Mbit)
*** Enabling sFlow:
s1
*** Sending topology
Generating NORMAL Traffic.....
```

Fig. 4 Creation of topology and generation of normal traffic

incoming traffic has certain parameters defined for every network packet flow. These parameters are gathered as training and testing features for the developed method for preventing DDoS attacks on a network built with SDN Architecture.

Figure 6 is the procedure that was followed for the developed testbed in the study. It comprises three phases. The data collection, feature selection and extraction, and the ML classification phase. The operations are described as follows:

```
[ '01/21/2021, 16:00:49', '77', '77', '0.0005451076587626056' ]  
[ '01/21/2021, 16:00:51', '78', '78', '0.0005337603416066186' ]  
[ '01/21/2021, 16:00:53', '77', '77', '0.00052301255523012552' ]  
[ '01/21/2021, 16:00:55', '76', '76', '0.0005128205128205128' ]  
[ '01/21/2021, 16:00:57', '79', '79', '0.0005026380539834129' ]  
[ '01/21/2021, 16:00:59', '36', '36', '0.00049813200498132' ]  
[ '01/21/2021, 16:01:01', '65', '65', '0.0004901960784313725' ]  
[ '01/21/2021, 16:01:03', '65', '65', '0.0004825090470446321' ]  
[ '01/21/2021, 16:01:05', '63', '63', '0.0004752851711026616' ]  
[ '01/21/2021, 16:01:07', '64', '64', '0.00046816479400749064' ]  
[ '01/21/2021, 16:01:09', '64', '64', '0.00046125461254612545' ]  
[ '01/21/2021, 16:01:11', '63', '63', '0.0004546487838145033' ]  
[ '01/21/2021, 16:01:13', '64', '64', '0.00044812906116961686' ]  
[ '01/21/2021, 16:01:15', '63', '63', '0.0004418912947414936' ]  
[ '01/21/2021, 16:01:17', '64', '64', '0.00043572984749455336' ]  
[ '01/21/2021, 16:01:19', '63', '63', '0.0004298302170642596' ]  
[ '01/21/2021, 16:01:21', '64', '64', '0.00042399830400678397' ]  
[ '01/21/2021, 16:01:23', '63', '63', '0.00041841004184100416' ]  
[ '01/21/2021, 16:01:25', '64', '64', '0.00041288191577208916' ]  
[ '01/21/2021, 16:01:27', '64', '64', '0.00040749796251018743' ]  
[ '01/21/2021, 16:01:29', '64', '64', '0.00040225261464199515' ]  
[ '01/21/2021, 16:01:31', '33', '33', '0.0003996003996003996' ]  
[ '01/21/2021, 16:01:34', '55', '55', '0.0003952569169960474' ]  
[ '01/21/2021, 16:01:35', '55', '55', '0.000391006684261974585' ]  
[ '01/21/2021, 16:01:37', '55', '55', '0.00038684719535783365' ]  
[ '01/21/2021, 16:01:39', '55', '55', '0.0003827751196172249' ]  
[ '01/21/2021, 16:01:41', '55', '55', '0.0003787878787878788' ]  
[ '01/21/2021, 16:01:43', '55', '55', '0.00037488284910965324' ]  
[ '01/21/2021, 16:01:45', '54', '54', '0.00037112636852848393' ]  
[ '01/21/2021, 16:01:47', '55', '55', '0.0003673769287288758' ]  
[ '01/21/2021, 16:01:49', '55', '55', '0.0003637024913620658' ]  
[ '01/21/2021, 16:01:51', '55', '55', '0.00036010082823190496' ]  
[ '01/21/2021, 16:01:53', '55', '55', '0.00035656979853806385' ]
```

Fig. 5 Attack traffic log by the RYU controller



Fig. 6 Procedure for the developed testbed

1. **Data Collection Phase:** This phase involves the collection of normal and abnormal traffic for feature extraction and SVM classification. In this phase, normal and DDoS attack traffic will be collected separately.
2. **Feature Extraction Phase:** At this phase, features that can indicate DDoS attacks were selected and extracted from the captured datasets, outliers were eliminated at this stage and indeed final dataset for the study and transported to the ML classifier.
3. **Classification Phase:** At this phase, the features extracted will be preprocessed to an acceptable format for the training and classification, then the classifiers performance will be evaluated. According to the set parameters, 0 was normal traffic and 1 was the traffic attack generated through machine classifiers.

4 System Implementation and Results

The developed method uses ML methods to detect DDoS attacks in an IPv6 enabled SDN. Table 1 contains the implementation environment and specifications. Basically, it contains the software and hardware requirements, programming language, types of software requirements, and the operating system.

The adopted method requires training of the SVM technique for the attack detection with a percentage of the generated data from the network. The collected data on the module was configured on the controller, the collected data comprises both normal traffic data and attacked traffic data which were stored. These data are stored as CSV files for the model to analyze. Normal traffic data were collected first followed by the attack traffic data. However, for higher accuracy, regular traffic data should be collected again following attack traffic data. The system implementation presented approach using the SVM approach to detect and DDoS attacks in an IPv6 enabled SDN. The implemented method requires training of the SVM approach for detecting the attack with a percentage of the generated data from the network, a flow of this process is shown in Fig. 7. Figure 8 is the SFlow RT packet monitoring during normal traffic generation, the blue line demonstrates the simulation of inflow traffic and the red line demonstrates the outflow of the traffic in Bit Per Second. The simulation results show an improvement compared to available ones in the literature.

On the Controller, the data collection module is configured to gather data from both regular and attack traffic and store it in a CSV file for analysis by the ML method. Normal traffic data must be gathered first, followed by attack traffic data. For improved accuracy, normal traffic data should be collected again following attack traffic data. The data is gathered using all of the feature extraction criteria described in the approach, such as source IP speed, flow entry speed, and flow pair entry ratio in distinct columns. Detection of the attack takes place after the data has been collected and processed, then the data is used to train and effectively classify the traffic data into normal and abnormal traffic. The SDN is made up of several protocols and controllers, each of which is meant to execute a specific purpose in order to increase network efficiency and flexibility. The presented methodology is implemented using

Table 1 Contains the implementation environment and specifications

| Software | Types | Programming language | Hardware requirements | Operating system |
|---------------------|------------------------------|--|--|---------------------------------|
| Ubuntu 18.04 | Open-Source Operating System | C, C++ and Python | Free disk space of 20 GB required minimum of 1 GB RAM required | Linux |
| Virtual box 6.1 | Open-Source | Python, C, C++ , × 86 programming language | Free disk space of 10 GB required, minimum of 1 GB RAM required | Linux, Mac OS and Windows |
| Ryu controller 4.34 | Open-Source | Python | Free disk space of 1 GB required, minimum of 256 MB RAM required | Linux, Mac OS and Windows and X |
| Mininet 2.2.2 | Open-Source | Python | Free disk space of 5 GB required, minimum of 1 GB RAM required | Linux |
| Sflow-RT | Open-Source | HTML and JS | Free disk space of 5 GB required, minimum of 1 GB RAM required | Linux, Mac OS and Windows |
| Wireshark 3.4.2 | Open-Source | C, and C++ | Free disk space of 5 GB required, minimum of 1 GB RAM required | Linux, Mac OS and Windows |

```

In [1]: from sklearn import svm, datasets
import matplotlib.pyplot as plt
import numpy as np
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
from sklearn import tree

In [2]: from sklearn.model_selection import cross_val_score

In [3]: data = np.loadtxt(open('result.csv', 'rb'), delimiter=',')
X = data[:, 0:2]
Y = data[:, 2]

In [4]: x_train, x_test, y_train, y_test = train_test_split(X, y, random_state = 0, test_size = 0.30)

In [5]: clf = svm.SVC(kernel='linear',C=0.030)

In [6]: clf.fit(x_train, y_train)
Out[6]: SVC(C=0.03, kernel='linear')

In [7]: classifier_predictions = clf.predict(x_test)

In [8]: print("Accuracy is ", accuracy_score(y_test, classifier_predictions)*100)
Accuracy is 99.8894099378882

In [9]: scores = cross_val_score(clf, x_train, y_train, cv=5)
print("cross-validation score - scores.mean()")
cross-validation score 0.9973333333333333
    
```

Fig. 7 Accuracy test on SVM DDoS

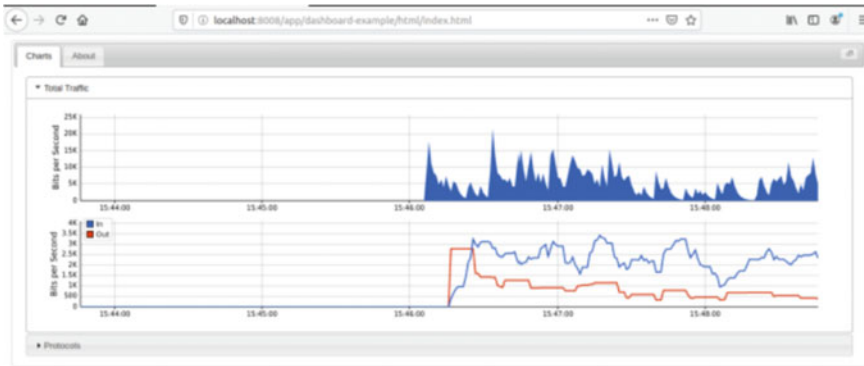


Fig.8 SFlow RT packet monitoring during normal traffic generation

the most popular of all the protocols for DDoS attacks in an IPv6 Enabled SDN. Since the Open Flow Protocol is the most widely used protocol for SDN, open switch is utilized for this project.

The rationale and approaches of the presented method are programmed in Python because it employs ML methodologies. Data analysis includes factors like source IP speed, flow entry speed, and flow pair entry ratio, which are all defined in the controller. The normal and attack traffic packets generated was 500,000 packets during the test for 20 min. The packets were pre-processed and the SVM was trained with 25% of the normal and the attack data. The SVM achieved an accuracy of 99.69% and a DDoS attacks detection rate of 100%.

5 Conclusion and Future Consideration

The previous session described the implementation of an SDN Topology, generation of normal and abnormal traffic packets, preprocessing of the traffic packets and the classification of the data by SVM algorithm. The SVM algorithm classified the traffic packets with 99.69% accuracy and 0 false positives, this result shows the SVM algorithm is applicable in a live SDN network environment, by configuring the controller to block IP addresses and ports that have been classified as DDoS attack hosts for time T in second to guarantee the perfect quality of service for other users on the network. The results assert the need for ML algorithms in IPv6 enabled SDNs. We have also successfully answered the research question stated and we are a step closer to answering the second question, this can be achieved by implementing a mitigation system within the network. ML algorithms will improve the quality of service of SDN if they are implemented as a second layer of security. The study recommends large scale implementations of classifying algorithms like the SVM should be adopted more in battling DDoS attacks. Also, more research should be done into DDoS attacks for SDNs before it becomes a major issue when its fully

deployed on networks. Other ML classifier algorithms should be implemented and compare the results.

References

1. Acharya S, Tiwari N (2016) Survey of DDoS attacks based on TCP/IP protocol vulnerabilities. *IOSR J Comput Eng* 18(3):68–76
2. Adeniyi DA, Wei Z, Yongquan Y (2016) Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. *Appl Comput Inform* 12(1):90–108
3. Bamimore I, Ajagbe SA (2020) Design and implementation of smart home for security using radio frequency modules. *Int J Digit Signals Smart Syst* 4(4):286–303. <https://doi.org/10.1504/IJDSST.2020.111009>
4. Fan C, Kaliyamurthy NM, Chen S, Jiang H, Zhou Y, Campbell C (2022) Detection of DDoS networking using entropy. *Appl Sci* 12(1):370
5. Al-Adaileh MA, Anbar M, Chong YW, Al-Ani, A (2018) Proposed statistical-based approach for detecting distribute denial of service against the controller of software defined network (SADDCS). In: *Proceedings of MATEC web of conferences EDP sciences*. Les Ulis, France, p 02012
6. Agne A, Platzner M, Lubbers E (2011) Memory virtualization for multithreaded reconfigurable hardware. In: *Proceedings of 21st international conference on field programmable logic and applications*. IEEE, pp 185–188
7. Akamai A (2018) *Memcached reflection attacks: A NEW era for DDoS*. Akamai technologies, Cambridge
8. Amaral P, Dinis J, Pinto P, Bernardo L, Tavares J, Mamede HS (2016) Machine learning in software defined networks: Data collection and traffic classification. In: *Proceedings of IEEE 24th international conference on network protocols (ICNP)*. IEEE, pp 1–5
9. Bakker J (2017) *Intelligent traffic classification for detecting DDoS attacks using SDN/Open flow*. Computer science. Victoria University Wellington. Accessed from <http://hdl.handle.net/10063/6645>
10. Banerjee U, Vashishtha A, Saxena M (2010) Evaluation of the capabilities of wire shark as a tool for intrusion detection. *Int J Comput Appl* 6(7):1–5
11. Bawany NZ, Shamsi JA, Salah K (2017) DDoS attack detection and mitigation using SDN: methods, practices, and solutions. *Arab J Sci Eng* 42(2):425–441
12. Bediako PK (2017) Long short-term memory recurrent neural network for detecting DDoS flooding attacks within tensorflow implementation framework. Lulea University of Technology, Sweden
13. Benzekki K, Fergougui A, Elbelrhiti Elalaoui A (2016) Software-defined networking (SDN): a survey. *Secur Commun Netw* 9(18):5803–5833
14. Bhuyan MH, Bhattacharyya DK, Kalita JK (2015) An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection. *Pattern Recogn Lett* 51:1–7. <https://doi.org/10.1016/j.patrec.2014.07.019>
15. Olabisi AA, Adeniji OD, Enangha A (2019) A comparative analysis of latency, jitter and bandwidth of ipv6 packets using flow labels in open flow switch in software defined network. *Afr J Manag Inf Syst (Afr. J. MIS)* 1(3):30–36

16. Adeniji OD, Khatun S, Borhan MS, Raja RS (2008) A design proposer on policy framework in IPv6 network. *Int Symp Inf Technol* 4:1–6
17. Ajagbe SA, Idowu IR, Oladosu JB, Adesina AO (2020) Accuracy of machine learning models for mortality rate prediction in a crime dataset. *Int J Inf Process Commun* 10(1 & 2):150–160
18. Musumeci F, Fidanci AC, Paolucci F, Cugini F, Tornatore, M (November 2021) Machine-learning-enabled ddos attacks detection in p4 programmable networks. *J Netw Syst Manage (JONS)* 30(21):1–27. Springer
19. Kokila R, Thamarai Selvi S, Govindarajan K (2014) DDoS detection and analysis in SDN-based environment using support vector machine classifier. In: 2014 sixth international conference on advanced computing (ICoAC), pp 205–210

Sensor Data Fusion Methods for Driverless Vehicle System: A Review



Nitheesh Kurian and K. Vadivukkarasi

Abstract According to the research studies in recent years, it is predicted that level 5 autonomous vehicles will take the place of conventional vehicles. These driverless vehicles can make decisions according to the environmental data and accomplish the driving task accordingly. In order to accomplish this, autonomous vehicles use different types of sensors in order to detect and perceive their local environment. But sensors can still malfunction due to environmental conditions, manufacturing defects or noise; so the information obtained from one sensor would not be reliable for the tasks associated with driverless vehicles. A feasible solution for this issue is to collaborate multiple sensors and fuse their data to accomplish more accurate driving tasks in Autonomous Driving systems. The various methods used to improve the navigation system of driverless vehicles are being reviewed in this survey.

Keywords Deep learning · Navigation · Localisation · Autonomous driving

1 Introduction

According to the latest studies there has been a great advancement in autonomous driving vehicles over the past few years. The National Motor Vehicle Crash Causation Survey (NMVCCS) says that 94% of conventional vehicle accidents occur due to the driver [1]. When compared with conventional vehicles, probability of accidents are very less in the case of driverless vehicles. The Society of Automobile Engineers (SAE) provides a taxonomy with different levels of autonomous driving, starting from no automation to full automation [2]. Honda is the first to introduce level 3 autonomous driving technology in a car, which will definitely inspire the industry to

N. Kurian (✉)

Electronics and Communication Engineering, Rajagiri School of Engineering and Technology, Kochi, India

e-mail: nitheeshk@rajagiritech.edu.in

K. Vadivukkarasi

Electronics and Communication Engineering, SRM Institute of Science and Technology, Chennai, India

e-mail: vadivukk@srmist.edu.in

commercialize this type of automation for vehicles [3]. BMW's system is coming with an all new electric vehicle which has also adopted level 3 driving automation [4]. Autonomous vehicles are using different types of sensors like LiDAR, RADAR, GPS, IMU etc. to perform different tasks. But one of the most challenging problems is Autonomous Navigation. Navigation is normally classified into four categories such as perception, localisation and mapping, path planning and control. The perception method will be using a set of sensors in the vehicle to detect and understand its surrounding; such as various obstacles, humans, traffic signs, signals etc.

Vehicle environment and surroundings will be transferred to a map by the method localisation and mapping and current location of the vehicle will be continuously monitored with respect to the reference coordinates in the map. Path planning will utilize the output of the last two tasks and recognises an easiest route for the vehicle to reach the end point. Control element will give the appropriate values of acceleration, steering angle and torque for the vehicle to travel along the feasible route [5].

Sensors are the most important components used in AV systems in order to sense and create a picture of vehicle surroundings and whose combined performance directly defines the safety of AV systems [6]. Multiple sensors can work together and fuse their outputs in an AV to enhance the overall quality of the system. So, sensor fusion is a major task for reducing the limitations associated with individual sensors. Since the environment is highly complex and dynamic, vehicles should make very accurate and fast decisions. Deep Learning algorithms can be used to implement perception, mapping and decision making which will help to improve the accuracy and processing speed.

2 Sensor Fusion

Sensor fusion originated in the 1970s and it was studied greatly for robotics and defense research. Joint Directors of Laboratories (JDL) describes that information fusion will detect and combine different types of data from single and multiple sources [7]. A group of sensors are used in autonomous vehicle systems to perform perception tasks. Basically there are two types of perception tasks such as environmental perception and localisation. RGB Cameras, LiDAR, RADAR etc. are used for environmental perception. Global positioning System (GPS), LiDAR, Inertial Measurement Unit (IMU) etc. are used for localisation [5].

In order to get accurate data AV systems use different combinations of sensors and fuse their outputs thereby reducing the issues with the individual sensors. RGB Cameras will capture the environmental information at relatively low cost. But it will provide only 2-D information about the surroundings. To model a 3-D environment Light Detection And Ranging (LiDAR) sensors can be used. It will provide in-depth information about the environment. But LiDAR sensor outputs are usually affected by weather conditions like rain, fog etc. [8]. So to compensate for the problems of individual sensors AV systems will employ heterogeneous sensor fusion.

Infrared and thermal imaging can be very easily used during night time for environmental perception tasks, because these two sensors can detect the presence of objects even in the absence of light [9]. Localisation and mapping are also very important issues in current AV systems and multiple sensor data binding is employed to enhance the efficiency of localisation and mapping in vehicles. Localisation and mapping generally integrate different types of sensors, like GPS, IMU, LiDAR and camera to get reliable and feasible outputs.

There are two types of localisation process if no previous map data is available. Global Positioning System/Inertial Navigation System (GPS/INS) integration method [10] and Simultaneous Localisation and Mapping (SLAM) methods. Due to outages, one of the main limitations of GPS is that it may lose its signal. By combining GPS and INS, INS is able to determine location and orientation during periods when the GPS signal is missing. Signals from satellites are affected by atmospheric conditions. Also received signals will get blocked by the physical infrastructure. This will lead to signal degeneration and thereby reducing the accuracy of GPS/INS systems. Simultaneous Localisation and Mapping (SLAM) [11] using a group of sensors like LiDAR and Camera to build the map and obtain the position of the vehicle and surrounding vehicle with respect to the pre-built map [12]. Table 1 describes various sensor fusion methods and their benefits over single sensor approaches employed in autonomous driving systems.

Table 1 Various sensor fusion methods and benefits in AV systems

| Literature | Sensors used | Fusion benefits | Application of fusion | Limitation without fusion |
|--|------------------|---|-----------------------|---|
| Melotti et al. [13] | LiDAR and camera | Better accuracy Works in bad weather conditions also | Pedestrian detection | Poor night vision in camera and resolution will be very low when LiDAR alone used |
| Kato et al. [14] | RADAR and camera | Distance measured accurately | Vehicle detection | Camera requires additional lenses and RADAR results low resolution images |
| Du et al. [15] | Lidar and camera | Final detection performance improved | Road detection | Unorganized point cloud LiDAR data and illumination problem with camera |
| Bresson et al. [11] and Alatisse et al. [16] | Camera and IMU | Improved accuracy and less computation | SLAM | Drifting error in IMU and lighting problem with camera |
| Xiong et al. [17] | GPS and IMU | Seamless and reliable navigation | Navigation | GPS signal loss in hill areas |

Sensor fusion is a process of combining individual sensor outputs to produce a combined outcome that is enhanced, more accurate and productive than individual sensor outputs.

In literature sensor fusion models are characterized in several ways. Which could be in terms of data level, objectives, application, sorts of sensors employed etc. But one among the major categorisations is predicated on level of detail within the data: data level, feature level and decision level [5]. Data level (Early fusion) fusion deals with the raw data coming out of the sensors. Feature level (Halfway fusion) extract the features of sensor output and fuse them together for better results. In Decision level (Late level fusion) multiple classifiers will take the decisions and finally combine them together to final result. According to Dasarathy [18] five different classification are there in sensor fusion, they are Data In-Data Out (DAI-DAO) Fusion, Data In-Feature Out (DAI-FEO) Fusion, Feature In-Feature Out (FEI-FEO) Fusion, Feature In-Decision Out (FEI-DEO) Fusion and Decision In-Decision Out (DEI-DEO) Fusion.

This classification is depicted in Fig. 1. DAI-DAO is the most basic form of fusion in which raw data from the sensor is given as input and the result will be also raw data. DAI-FEO method will generate a set of features as output which will use fusion of raw sensor data from different sources. FEI-FEO derives features from sensor output and combines them to generate a set of features. This is frequently called feature fusion. In FEI-DEO, input will feature vectors from different sensors and output will be a decision. DEI-DEO is commonly known as decision fusion and both input and output are decisions.

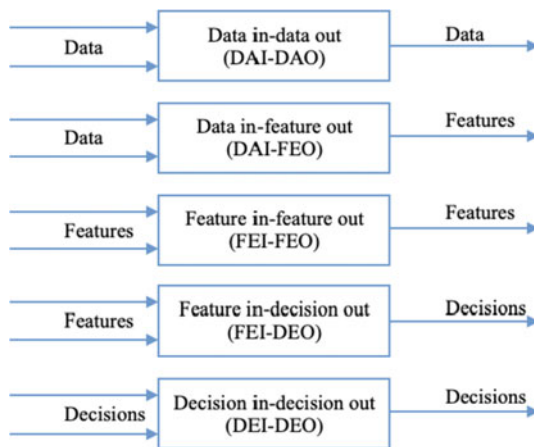


Fig. 1 Dasarathy's classification [18]

2.1 Sensor Fusion: Conventional Algorithms

The fusion algorithms must eventually fuse the data from input sources, regardless of the fusion architecture or sensor components. However, in the real world, sensor fusion approaches must address issues such as data imperfection, data similarity, data uncertainty, and data disparity [19]. Sensor fusion algorithms are categorized based on the data fusion problems they face. Data imperfection is one of the most challenging issues of data fusion systems. To reflect data imperfection, various mathematical theories are used, such as probability theory [20], fuzzy set theory [21], possibility theory [22], and so on. A specific aspect of imperfect data can be modeled by using these theories. For example, uncertainty in data can be represented by a probabilistic distribution, fuzzy set theory can represent data vagueness, and ambiguous and uncertain data can be modeled using evidential belief theory.

Sensor fusion can be performed at various levels, like low, medium and high level fusion [23]. This classification is predicated on the sort of the processed input file and therefore the resultant information providing for the system. Low level deals with the signal coming out of the sensor. Medium level is using features extracted from the sensors for the fusion purpose and decision making fusion is treated as high level. Table 2 explains the conventional fusion algorithms and their benefits.

3 Deep Learning Algorithms: Sensor Fusion

Deep learning is a subset of Machine Learning and Artificial Intelligence which imitates the functionalities of the human brain. Artificial Neural Network (ANN) is considered as the heart of Deep Learning. ANN is a Machine Learning process which is similar to the human nervous system and brain [26]. Deep learning algorithms are better than machine learning methods when we are dealing with very large data sets because, learning and modeling of large data sets can be easily done with deep

Table 2 Various conventional sensor fusion algorithms

| Literature | Algorithm | Fusion benefits | Application of fusion | Level of estimation |
|-----------------------|----------------------|--|-----------------------|---------------------|
| Castanedo et al. [24] | Probabilistic method | Deals nonlinear systems | Estimation | Low [23] |
| Pires et al. [25] | Statistical method | Able to deal with unknown correlations | Estimation | Low [23] |
| Bahador et al. [19] | Evidential | Enables fusion of uncertain and ambiguous data | Decision | High [23] |
| Bahador et al. [19] | Knowledge based | Handles uncertainty and imprecision | Classification | Medium [23] |

learning algorithms, so it can be employed in various applications like autonomous vehicles [27], Robotics [28], object detection [29], health care [30], advertisement [31] etc.

Deep learning algorithms can be used for pedestrian detection, environmental perception, traffic sign detection etc. One of the best applications is environmental perception in which an autonomous vehicle can detect and build a complete idea about the surroundings and objects around the vehicle. Different sensors like Camera, LiDAR, RADAR etc. are mainly used to build environmental perception maps. Since convolutional neural networks are very good with visual data, they are normally used as detection algorithms in environmental perception.

Most commonly used deep learning sensor fusion algorithms in autonomous vehicles are Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). CNN has two types of algorithms such as Single stage detectors and Two stage detectors. Single stage detectors are again classified into You Only Look Once (YOLO), Single-Shot Multibox Detector (SSD) and Deconvolutional Single-Shot Detector (DSSD). Two stage detectors are Region-Based CNN (R-CNN), Spatial Pyramid Pooling Network (SPPN), Fast R-CNN and Faster R-CNN. RNN has two different variations like Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU).

3.1 Convolutional Neural Network (CNN)

CNN belongs to the category of feed-forward neural networks. They are mainly used for image data processing by making use of a grid topology. A CNN is employed to identify and classify different objects in a picture. CNN is able to identify and group specific features from pictures [32]. Convolution in CNN is a linear operation in mathematics in which multiplication between two functions will generate a resultant function which explains about the modification of the first function by the second. Basic architecture of CNN is shown in Fig. 2. Basic architecture of CNN.

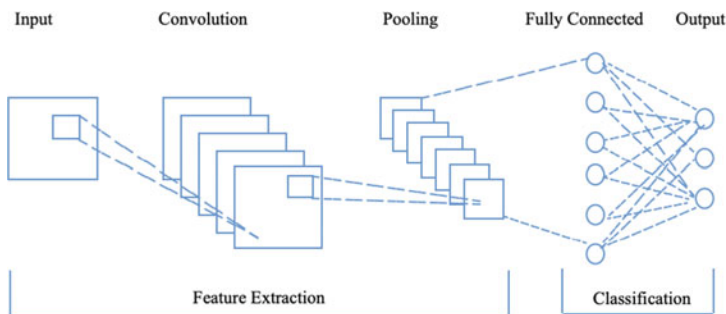


Fig. 2 Basic architecture of CNN

CNN architecture can be classified into two main parts, including.

1. The Feature Extraction process will set apart and detect various features of the image.
2. Convolution output will be given to the fully connected layer and it will predict the class of image by using the features extracted in the last stage.

Three layers are used to construct CNN architecture such as, Convolution layers, Pooling layers and Fully Connected (FC) layers. Apart from these layers, another two important parameters of CNN are activation function and dropout layer.

- Convolution layer: First layer of CNN architecture. The main function of this layer is extracting different features of the input images. Convolution layer consists of a filter with size $M \times M$ and convolution operation will be carried out between this filter and input image. The given filter will glide over the input image and a part of the image having the same size of the filter will be multiplied with the filter using a dot product operation. The output of convolution layers will give the details about the corners and edges and they are called feature maps.
- Pooling layer: Pooling layer will reduce the connections between layers and independently operates on every feature map and hence it can reduce the cost. Pooling operations are classified according to the methods used, such as Max Pooling, Average Pooling etc.
- Fully Connected layer: It is acting as a classifier and basically establishes a connection between the neurons of two different layers. Output layer will be coming after this fully connected layer.

CNN image detectors are classified into two: (1) single stage detectors, during which predictions are going to be administered without an intermediate stage; and (2) multi stage detectors, in which prediction will take place only after region proposal.

3.2 *You Only Look Once (YOLO)*

Redmon et al. [34] proposed YOLO in 2016—You Only Look Once, one of the most popular algorithms for object detection. It is an example of a single stage detector. YOLO algorithm is capable of predicting various bounding boxes using a single neural network. It also determines the class probabilities for the bounding boxes. Figure 3 [34] explains the architecture of YOLO.

Twenty four different convolution layers and two fully connected layers are used to construct the YOLO detection network. 1×1 convolutional layers present in the network helps to reduce the features space from preceding layers. According to Redmon et al. [34] YOLO is having fast detection at a rate of 45 frames per second. But it is having spatial constraints on bounding box predictions and also its very difficult to detect small objects appearing in groups. These challenges were addressed by the algorithms YOLOv2 [35], YOLOv3 [36] and YOLOv4 [37].

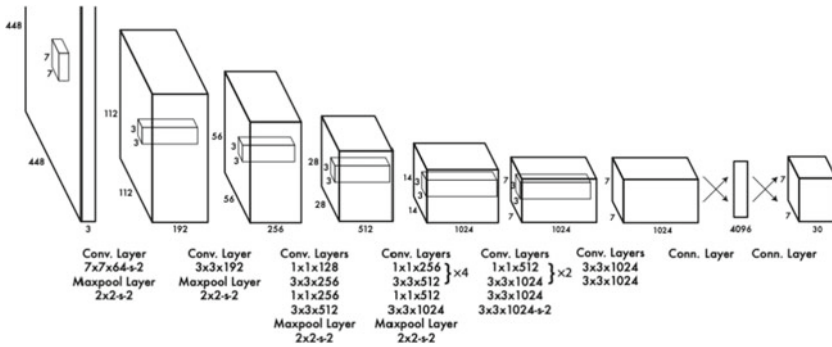


Fig. 3 Basic architecture of YOLO

Asvadi et al. [38] proposed a multi modal and multi sensor vehicle detection system, in which colour image, dense Depth Map (DM) and dense Reflectance Map (RM) from 3 D LiDAR are used as the inputs. They are fused together to detect objects on the road. These inputs are given to the YOLO object detection framework consisting of three different YOLO networks such as YOLO-C (Responsible for colour image), YOLO-D (Responsible for DM) and YOLO-R (Responsible for RM). Each network will generate bounding boxes separately and decision level fusion is employed for object detection.

3.3 Single-Shot Multibox Detector (SSD)

According to Dou et al. [] the efficiency of the YOLO algorithm will be reduced when it deals with small objects. Also YOLO is experiencing spatial constraints on the bounding boxes [40]. Single-Shot Multibox Detector (SSD) [41] can be used for eliminating the challenges faced by the YOLO algorithm. Size and aspect ratios of the bounding boxes used in SSD will be different for different objects, so that it can be employed for detecting different objects with different sizes. When compared with YOLO in terms of accuracy and speed SSD offers faster speed with 59 frames per second and better accuracy [41]. Kim et al. [42] proposed a deep learning fusion algorithm using SSD for object detection. In this method they have fused camera images and LiDAR 3D point cloud data with the help of two separate CNN networks which is similar to SSD.

3.4 Region-Based Convolutional Neural Network (R-CNN)

Girshick et al. suggested Region-based CNN (R-CNN), which belongs to the category of two stage detectors [33]. Major advantage of R-CNN is it can avoid the problem of

selecting a huge number of regions. There are three stages in R-CNN such as feature extraction from region proposals, object classification using Support Vector Machine (SVM) and bounding box regression. The main aim of R-CNN is to increase the detection speed. Girshick et al. improved speed by employing a sophisticated search algorithm that created more than 1500 regions in the image rather than covering the entire image. CNN architecture will accept the selected region from the previous step and extract the features from the regions. These features will be classified using SVM. R-CNN algorithm can be used in autonomous vehicles for sensor fusion. Wagner et al. [43] proposed the R-CNN algorithm for fusing images from visible and thermal cameras for pedestrian detection. In their study they have used the R-CNN algorithm for early and late fusion architectures.

3.5 Fast Region-Based Convolutional Neural Network (Fast R-CNN)

Girshick et al. [44] suggested Fast R-CNN to enhance the accuracy and speed. One of the major problems of R-CNN is, object detection is a very slow and very expensive training process. To overcome this a feature map will be generated by processing the input image. According to their research, Fast R-CNN is almost 9 times quicker in training than R-CNN [44]. Quality of detection is very good in Fast R-CNN. This architecture will accept an entire image and a set of object proposals as input. Convolutional and max pooling layers are responsible for making a convolutional feature map by processing the entire image. Then feature vectors will be extracted by Region of Interest (RoI) pooling layer from the object proposals. Which is then fed to fully connected layers and finally which is given to two output layers.

3.6 Faster Region-Based Convolutional Neural Network (Faster R-CNN)

Ren et al. [45] proposed a neural network called Region Proposal Network (RPN), which can be used to predict bounding boxes. Faster R-CNN is a new algorithm developed by merging RPN with R-CNN. In comparison with other CNN networks, Faster R-CNN has better accuracy and fast processing time. Faster R-CNN was used in pedestrian detection [46] which fuses thermal and colour images. In this experiment, they used a variety of fusion architectures, including early fusion, halfway fusion, late fusion, and score fusion.

Table 3 shows comparison between two stage detector algorithms from the experimental results published in research works [33, 43–46].

Table 3 Comparison between two stage detector algorithms

| Parameter | R-CNN | Fast R-CNN | Faster R-CNN |
|---|-------|------------|----------------|
| Time required for training (hrs) | 84 | 9.5 | Not applicable |
| Time required for testing (seconds/image) | 47 | 2.3 | 0.2 |

4 Conclusion

The research growth in the Autonomous vehicle domain is exponentially increasing nowadays. Because it is a multi disciplinary domain involving different subjects like Electronics, Mechanical Engineering, Sensor Management and fusion etc. In this survey we have studied different algorithms used for multi sensor data fusion required in autonomous vehicles. Each algorithm is showing better performance in different conditions, but deep learning algorithms can be more effective for sensor fusion methods. One of the best deep learning algorithms that can be used in autonomous vehicles for environmental perception is CNN. It has a powerful capacity to handle images and it will automatically learn the features. At the same time it is very important to think about the data to be handled by an autonomous vehicle system. Within a period of time the amount of data generated will be very high and huge. So the efficiency of all deep learning algorithms will depend on the extensiveness and quality of the training. Inorder to handle this big data problem fusion between deep learning algorithms and conventional based methods to be used.

References

1. Singh S (2015) Critical reasons for crashes investigated in the national motor vehicle crash causation survey. Traffic Safety Facts Crash Stats. Report No. DOT HS 812 115; National Center for Statistics and Analysis, Washington, DC, USA
2. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016 Ground Vehicle Standard)—SAE Mobilus. https://saemobilus.sae.org/content/j3016_201806
3. Honda launches world's first level 3 self-driving car Plan for just 100 lease sales reflects Japanese automaker's cautious approach. <https://asia.nikkei.com/Business/Automobiles/Honda-launches-world-s-first-level-3-self-driving-car>
4. BMW Takes Self-Driving to Level 3 Automation. <https://www.electronicdesign.com/markets/automotive/article/21136427/bmw-takes-selfdriving-to-level-3-automation>
5. Fayyad J, Jaradat MA, Gruyer D, Najjaran H (2020) Deep learning sensor fusion for autonomous vehicle perception and localization: a review. *Sensors* 20(15):4220. <https://doi.org/10.3390/s20154220>
6. Giacalone J-P, Bourgeois L, Ancora A (2019) Challenges in aggregation of heterogeneous sensors for autonomous driving systems. In: 2019 IEEE sensors applications symposium (SAS) (2019). France, pp 1–5. <https://doi.org/10.1109/SAS.2019.8706005>
7. White FE (1991) Data fusion lexicon, joint directors of laboratories, technical panel for C3, data fusion sub-panel, Naval Ocean Systems Center, San Diego
8. Rasshofer RH, Spies M, Spies H (2011) Influences of weather phenomena on automotive laser radar systems. *Adv Radio Sci* 9:49–60

9. Mees O, Eitel A, Burgard W (2016) Choosing smartly: adaptive multimodal fusion for object detection in changing environments. In: Proceedings of the 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS), Daejeon, Korea, 9–14 Oct 2016; IEEE: Daejeon, Korea, 2016, pp 151–156
10. Redmill KA, Kitajima T, Ozgumer U (2001) Dgps/ins integrated positioning for control of automated vehicle. *Proc IEEE Int Transp Syst* 172–178
11. Bresson G, Alsayed Z, Li Y, Glaser S (2017) Simultaneous localization and mapping: a survey of current trends in autonomous driving. *IEEE Trans Intell Veh Inst Electric Electron Eng XX:1*. <https://doi.org/10.1109/TIV.2017.2749181>. hal-01615897
12. Bresson G, Rahal M-C, Gruyer D, Revilloud, M, Alsayed Z (2016) A cooperative fusion architecture for robust localization: application to autonomous driving. In: Proceedings of the 2016 IEEE 19th international conference on intelligent transportation systems (ITSC), Rio de Janeiro, Brazil, 1–4 Nov 2016, pp 859–866
13. Melotti G, Premebida C, Gonçalves NMDS, Nunes UJ, Faria DR (2018) Multimodal CNN pedestrian classification: a study on combining LIDAR and camera data. In: Proceedings of the 2018 21st international conference on intelligent transportation systems (ITSC), Maui, HI, USA; 4–7 Nov 2018. IEEE, Maui, HI, 2018, pp 3138–3143
14. Kato T, Ninomiya Y, Masaki I (2002) An obstacle detection method by fusion of radar and motion stereo. *IEEE Trans Intell Transp Syst* 3:182–188. <https://doi.org/10.1109/TITS.2002.802932>
15. Du X, Ang MH, Rus D (2017) Car detection for autonomous vehicle: LIDAR and vision fusion approach through deep learning framework. In Proceedings of the 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS), Vancouver, BC, Canada, 24–28 Sept 2017; IEEE, Vancouver, BC, 2017, pp 749–754
16. Alalise MB, Hancke GP (2017) Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter. *Sensors* 17:2164. <https://doi.org/10.3390/s17102164>
17. Xiong Y, Zhang Y, Guo X, Wang C, Shen C, Li J, Tang J, Liu J (2019) Seamless global positioning system/inertial navigation system navigation method based on square-root cubature Kalman filter and random forest regression. *Rev Sci Instrum* 90(1):015101. <https://doi.org/10.1063/1.5079889>
18. Dasarathy BV (1997) Sensor fusion potential exploitation–Innovative architectures and illustrative applications. United States. <https://doi.org/10.1109/5.554206>
19. Khaleghi B, Khamis A, Karray FO, Razavi SN (2013) Multisensor data fusion: a review of the state-of-the-art. *Inf Fusion* 14(1):28–44. <https://doi.org/10.1016/j.inffus.2011.08.001>. (January, 2013)
20. Durrant-Whyte HF, Henderson TC (2008) Multisensor data fusion. In: Siciliano B, Khatib O (eds) *Handbook of robotics*. Springer, pp 585–610
21. Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353
22. Zadeh LA (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst* 1(1):3–28
23. Luo RC, Chang C-C (2012) Multisensor fusion and integration: a review on approaches and its applications in mechatronics. *IEEE Trans Ind Inform* 8:49–60. <https://doi.org/10.1109/TII.2011.2173942>
24. Castanedo F (2013) A review of data fusion techniques. *Sci World J* 2013:1–19. <https://doi.org/10.1155/2013/704504>
25. Pires I, Garcia N, Pombo N, Flórez-Revuelta F (2016) From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices. *Sensors* 16:184. <https://doi.org/10.3390/s16020184>
26. Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. *IEEE Access* 7:53040–53065. <https://doi.org/10.1109/ACCESS.2019.2912200>
27. Oh C-S, Yoon J-M (2019) Hardware acceleration technology for deep-learning in autonomous vehicles. In: 2019 IEEE international conference on big data and smart computing (BigComp), Kyoto, Japan, 2019, pp 1–3. <https://doi.org/10.1109/BIGCOMP.2019.8679433>

28. Zunjani FH, Sen S, Shekhar H, Powale A, Godnaik D, Nandi GC (2018) Intent-based object grasping by a robot using deep learning. In: 2018 IEEE 8th international advance computing conference (IACC), Greater Noida, India, 2018, pp 246–251. <https://doi.org/10.1109/IADCC.2018.8692134>
29. Akyol G, Kantarcı A, Çelik AE, Ak AC (2020) Deep learning based, real-time object detection for autonomous driving. In: 2020 28th signal processing and communications applications conference (SIU) Gaziantep, Turkey, 2020, pp 1–4. <https://doi.org/10.1109/SIU49456.2020.9302500>
30. Che Z, Liu Y (2017) Deep learning solutions to computational phenotyping in health care. In: 2017 IEEE international conference on data mining workshops (ICDMW), New Orleans, LA, USA, 2017, pp 1100–1109. <https://doi.org/10.1109/ICDMW.2017.156>
31. Lee S-H, Yoon S-H, Kim H-W (2021) Prediction of online video advertising inventory based on TV programs: a deep learning approach. *IEEE Access* 9:22516–22527. <https://doi.org/10.1109/ACCESS.2021.3056115>
32. Convolutional Neural Networks, Chapter. <https://www.accessengineeringlibrary.com/content/book/9781260462296/chapter/chapter11>
33. Girshick R, Donahue J, Darrell T, Malik J (2013) Rich feature hierarchies for accurate object detection and semantic segmentation. [arXiv:1311.2524](https://arxiv.org/abs/1311.2524)
34. Redmon J, Divvala S, Girshick R, Farhadi A (2015) You only look once: unified, real-time object detection. [arXiv:1506.02640](https://arxiv.org/abs/1506.02640)
35. Redmon J, Farhadi A (2017) YOLO9000: Better, faster, stronger. In: Proceedings of the 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017, pp 6517–6525
36. Redmon J, Farhadi A (2018) YOLOv3: an incremental improvement. [arXiv:1804.02767](https://arxiv.org/abs/1804.02767)
37. Bochkovskiy A, Wang C-Y, Hong-Yuan Mark Liao. Yolov4: optimal speed and accuracy of object detection[EB/OL]. (2020–04–23). [arXiv: 2004.10934](https://arxiv.org/abs/2004.10934)
38. Asvadi A, Garrote L, Premebida C, Peixoto P, Nunes U (2017) Multimodal vehicle detection: fusing 3D-LIDAR and color camera data. *Pattern Recogn Lett* 115. <https://doi.org/10.1016/j.patrec.2017.09.038>
39. Wang H, Lou X, Cai Y, Li Y, Chen L (2019) Real-time vehicle detection algorithm based on vision and lidar point cloud fusion. *J Sens* 2019, Article ID 8473980, 9 pages. <https://doi.org/10.1155/2019/8473980>
40. Han J, Liao Y, Zhang J, Wang S, Li S (2018) Target Fusion Detection of LiDAR and camera based on the improved YOLO algorithm. *Mathematics* 6:213. <https://doi.org/10.3390/math6100213>
41. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC (2015) SSD: single shot multibox detector. [arXiv:20151512.02325](https://arxiv.org/abs/20151512.02325)
42. Kim J, Choi J, Kim Y, Koh J, Chung CC, Choi JW. (2018) Robust camera lidar sensor fusion via deep gated information fusion network. In: Proceedings of the 2018 IEEE intelligent vehicles symposium (IV), Changshu, China, 26–30 June 2018. IEEE, Changshu, China, 2018, pp 1620–1625
43. Wagner J, Fischer V, Herman M, Behnke S (2016) Multispectral pedestrian detection using deep fusion convolutional neural networks. In: Proceedings of the ESANN, Bruges, Belgium, 27–29 April 2016
44. Girshick R (2015) Fast R-CNN. In: Proceedings of the 2015 IEEE international conference on computer vision (ICCV), Santiago, Chile, 7–13 Dec 2015, pp 1440–1448
45. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. [arXiv:1506.01497](https://arxiv.org/abs/1506.01497)
46. Liu J, Zhang S, Wang S, Metaxas D (2016) Multispectral deep neural networks for pedestrian detection. [arXiv:1611.02644](https://arxiv.org/abs/1611.02644)

A Novel Neural Network Based Model for Diabetes Prediction Using Multilayer Perceptron and Jrip Classifier



B. Sreedevi, Durga Karthik, J. Glory Thephoral, M. Jeya Pandian, and G. Revathy

Abstract People should be conscious on good health and hygiene. Diabetes is another epidemic that threatens not only elderly but even youngsters due to sedentary lifestyles. Managing blood sugar level is vital to avoid further health complications, particularly for Type 2 Diabetes. Early prediction will help people to change their eating habits. Dataset containing Age, Gender, Height, Weight, Insulin Level, Fat Level for various age groups were collected for prediction. The JRip technique was employed to generate rules for the above parameters and the data was tested using Multi-Layer Perceptron (MLP) for improved accuracy. MLP model with 10-fold Jrip cross validation predicted the class labels with 99.8% accuracy.

Keywords Diabetes · JRip · Multi-layer perceptron · Prediction · Accuracy

1 Introduction

According to World Health Organization 36 million deaths are due to non-communicable diseases such as cardiovascular disorders, cancer, diabetes, etc., Pancreas secretes insulin, which allows our cells to absorb glucose from food [1, 2]. In people with Type 2 Diabetes, insulin is not secreted, and it results in improper

B. Sreedevi · D. Karthik (✉) · J. Glory Thephoral · M. Jeya Pandian
Department of Computer Science and Engineering, SRC, SASTRA Deemed University,
Kumbakonam, India
e-mail: durgakarthik@src.sastra.edu

B. Sreedevi
e-mail: sreedevi@src.sastra.edu

J. Glory Thephoral
e-mail: gloryjs81@src.sastra.edu

M. Jeya Pandian
e-mail: jeyapandian@src.sastra.edu

G. Revathy
Department of Computer science and Engineering, SRC, SASTRA Deemed univeristy,
Kumbakonam, India

functioning of cells. This is called insulin resistance that leads to increased blood glucose level (hyperglycemia). Prolonged hyperglycemia leads to weakening of blood vessels, vision problems, and sometimes loss of vision [3].

International Diabetes Federation (IDF) has predicted that in the next decade diabetes may increase to 592 million. Healthy improvement in lifestyle can manage and prevent diabetes. Diabetes can be controlled or prevented with proper diet, physical work, consistent screening and treatment [4, 5]. Hence it is a need of hour for early prediction of diabetes for our society.

Machine learning techniques provides various methodology for early predictions against diseases or natural calamities. The collected data can be preprocessed [6], analyzed [7] and the results can be used for classification [8], clustering, prediction [9, 10] etc., Data analytics include various measures such as accuracy, recall, precision as metrics to validate the analyzed data.

1.1 Existing Work

Convolution neural network was used in early diabetic classification for eyes by using deep feature extraction method [11]. Image classification can be done using various convolutional techniques in a better manner [12]. Type 2 diabetic prediction was done using hybrid approaches by combining various neural network algorithms [13]. Iris image features were used for early detection of diabetics by combining neural network algorithms [14]. Recently machine learning models are used for predicting diabetes in a more accurate manner [15]. Clustering methods are more common in detecting diabetes disease in an extended way [16].

2 Proposed Work

The work aims at collecting data from various hospitals in around Kumbakonam regions. The data will be first applied to classify the data as diabetic or non-diabetic using JRIP rules generated and on the MLP algorithm. The results will be compared for the various statistical feature. The model will be constructed based on the statistical results such as accuracy, precision, kappa statistics etc.

2.1 Data Collection

The diabetes dataset for the year 2014 to 2019 was collected from Kumbakonam, Government Hospital and Swamimalai, Primary Health Centre (PHC), Tamilnadu. The dataset contains 1160 records and each with the following attributes: Age, Gender, Height, Weight, BMI, Cholesterol level, insulin level.

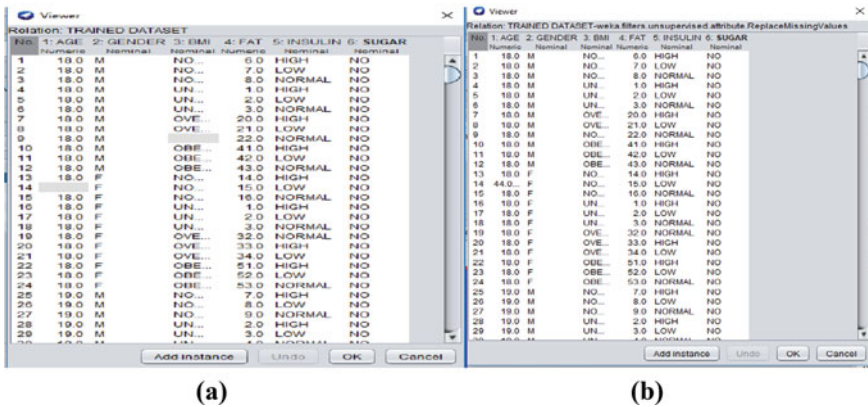


Fig. 1 a Dataset with missing values. b Mean Filtering on dataset

2.2 Data Preprocessing

The cholesterol level was changed to Fat, random insulin level numeric value was converted to low (<120 mg/l), high (>140 mg/l) or normal (120–140) and BMI converted to Normal, Overweight, Under Weight for rule generation. The obtained data contains missing values, mean filters were applied to replace them. The sample data shown in Fig. 1a containing missing value (blue boxes). Mean filters were applied using Weka tool to fill the missing values as shown in Fig. 1b.

3 Model Formulation and Experimental Work

A neural model will be formulated based on the results obtained from JRIP and MLP model. An algorithm is designed for the Novel Diabetic Prediction Model [NDPM].

3.1 NDPM Algorithm

Step 1: Start.

Step 2: Preprocess the collected data and apply PCA for identifying principal components.

Step 3: Generate rules on the identified components using JRIP rule miner to classify if a person is diabetic or not.

Step 4: Use Multi-layer perceptron to identify the class labels on the collected data.

Step 5: Create a model for early diabetic detection with higher accuracy and other statistical features by comparing the results from JRIP and MLP Classifiers.

Step 6: Stop.

The above algorithm was used in creating a model. JRip classifier algorithm is used to generate the model for prediction. The algorithm uses incremental reduced error and generates the classes by adopting a set of rules initially. The MLP employed 10-fold cross validation on the dataset for iterative pruning the dataset by randomly selecting 10 parts of which 9 will be used for training and 1 for testing. This method has been adopted recently for its simplicity and on its accuracy in generating the class label. Hence from various methods Jrip and MLP algorithm were employed in efficient model creation.

First step after preprocessing is to generate the rules for early detection of diabetes using JRIP classifier. Principal component analysis is used to identify important parameters from the available dataset. The important variables identified are age, BMI, fat and insulin level.

The Jrip rules are generated using the four components as given below:

The rules extracted from the JRip Classifier are as follows:

- **Rule 1:** (AGE > = 51) and (INSULIN LEVEL = LOW) = > DIABETES = YES (160.0/0.0)

Rule 1 states that if the person age is above 50 and the insulin level is low then the patient will have diabetes.

- **Rule 2:** (AGE > = 51) and (INSULIN LEVEL = NORMAL) = > DIABETES = YES (160.0/2.0)

Rule 2 states that if the patient age is above 50 and if he/she has insulin level normal then the patient will have diabetes.

- **Rule 3:** (INSULIN LEVEL = LOW) and (FAT > = 28) and (AGE > = 36) = > DIABETES = YES (55.0/1.0)

Rule 3 states that if the patient age is greater than 35 and if the insulin level is low and fat level is greater than 28 then the patient will have diabetes.

- **Rule 4:** (INSULIN LEVEL = LOW) and (FAT > = 49) and (AGE > = 26) = > DIABETES = YES (21.0/1.0)

Rule 4 states that if the patient age is greater than 26 and if the insulin level is low and the fat level is also high, then the patient will have diabetes.

- **Rule 5:** (INSULIN LEVEL = LOW) and (AGE > = 37) and (BMI = OVERWEIGHT) = > DIABETES = YES (7.0/0.0)

Rule 5 states that for the patient with age greater than 37 if their BMI is overweight and insulin level is low then the patient will have diabetes.

- **Rule 6:** (INSULIN LEVEL = LOW) and (AGE > = 46) and (BMI = NORMAL) = > DIABETES = YES (8.0/0.0)

Rule 6 states that if the patient is at the age greater than 46 and with BMI level as normal and insulin level is low then the patient will have diabetes.

- **Rule 7:** DIABETES = NO (861.0/4.0)

Rule 7 states that the patient may not have diabetes as all the above rule fails.

The above rules were applied on the collected data for early prediction of diabetes.

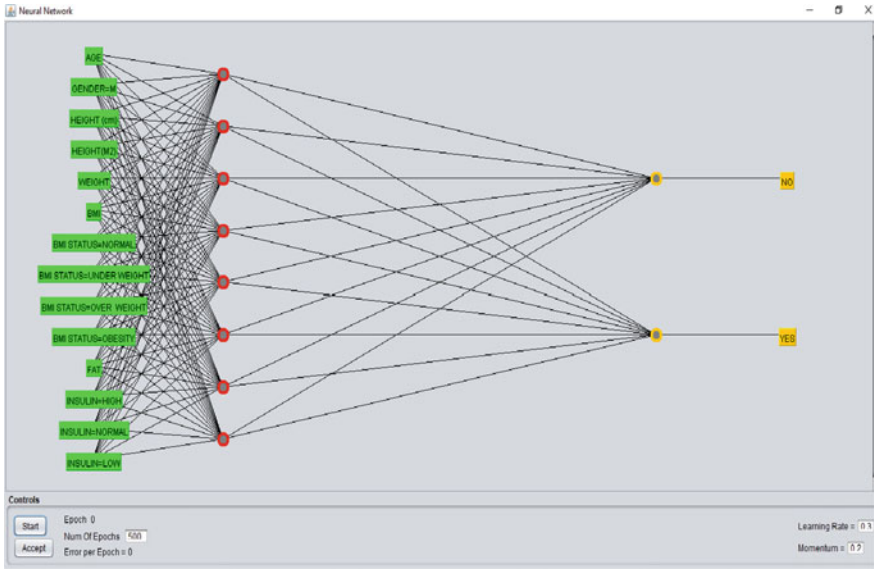


Fig. 2 Visualization of MLP results for diabetes dataset

The next step is to classify the data using MLP tool. Weka tool was used to implement a multilayer perceptron (input, hidden and output layers), which is nothing but a feed-forward network. Each layer has several numbers of nodes. The flow of information is through forward direction. The input neuron performs independent computation and passes the value to output layer through hidden layer. The Fig. 2 shows the classification of patient data on MLP.

MLP algorithm can automatically classify the data set into class labels. The given dataset was classified in to two class labels, one being diabetic and another nondiabetic class label.

4 Results and Discussion

The generated rules from Jrip classifier were applied on the collected dataset. The statistical results from Jrip on the data is shown in Table 1.

Table 1 Statistical metrics of JRip classifier

| | TP rate | FP rate | Precision | Recall | F-Measure | MCC | ROC area | PRC area | Class | Kappa statistic |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|-----------------|
| Weighted Avg: | 0.993 | 0.019 | 0.991 | 0.993 | 0.992 | 0.975 | 0.772 | 0.933 | No | 0.9748 |
| | 0.981 | 0.007 | 0.985 | 0.981 | 0.983 | 0.975 | 0.957 | 0.983 | Yes | |
| | 0.989 | 0.015 | 0.989 | 0.989 | 0.989 | 0.975 | 0.814 | 0.949 | | |

Table 2 Average standard metrics values of multilayer perceptron

| | TP rate | FP rate | Precision | Recall | F-Measure | MCC | ROC area | PRC area | Class | Kappa statistic |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|-----------------|
| Weighted Avg: | 0.998 | 0.031 | 0.999 | 0.998 | 0.999 | 0.953 | 1.000 | 1.000 | No | 0.952 |
| | 0.969 | 0.002 | 0.939 | 0.969 | 0.954 | 0.953 | 1.000 | 0.991 | Yes | |
| | 0.997 | 0.030 | 0.997 | 0.997 | 0.997 | 0.953 | 1.000 | 1.000 | | |

Table 3 Confusion matrix for MLP result

| Class label | True positive | False positive |
|-------------|---------------|----------------|
| Yes | 87 | 57 |
| No | 6 | 1010 |

The results from JRIP algorithm using Weka tool reveal that the precision for class no (non-diabetic) is 99% and for yes (diabetic) is 98.5%. Average precision and recall for both the class are 98.9%. Similarly F measure is 98.9% with kappa statistics 0.9748, that is near to 1.

Similarly MLP algorithm with 10-fold cross validation was applied on the collected data and the average statistical metrics is given Table 2.

The results from MLP algorithm reveal that precision is 99.9% for class label no (meaning nondiabetic) and 93.9% diabetic class. Recall, F-Measure is 99.7% and Kappa statistic is 0.952.

The confusion matrix using MLP is given below in Table 3 for the collected data.

True positive for yes (Diabetes predicted) class label is 87 and false positive is 57. Similarly true negative for no (not diabetic) is 6 and 1010 is false negative (Chances for diabetic) from the data set.

5 Conclusion

The Multilayer perceptron classification has 99.8% accuracy and JRip classifier rules has 99.3% accuracy for the class labels. The accuracy level changes in MLP are based on the number of hidden layers on the network. Kappa statistics is an important metrics that determines how good a class label in classifier. If Kappa statistics are near to 1, the rules can be accepted, for Jrip (0.97) and for MLP (0.95). Hence Jrip rule miner can be adopted for predicting Type 2 diabetes as it has a good accuracy and Kappa value. The work can be extended to include more data by combining other hospitals from nearby regions to improve the accuracy of the work. Similarly, the model can be used for classifying other health issues such cancer detection, and thyroid detection.

References

1. Joshi TN, Chawan PM (2018) Diabetes prediction using machine learning techniques. *Int J Eng Res Appl* 8(1):(part-II). (January 2018)
2. Anant KA, Ghorpade T, Jethani V (2017) Diabetic retinopathy detection through image mining for type 2 diabetes. In: 2017 ICCCI-2017. (January 5, 2017)
3. Khalil RM, Al-Jumaily A (2017) Machine learning based prediction of depression among type 2 diabetic patients. In: International conference on intelligent systems and knowledge engineering
4. Ali Alzubi A (2018) Big data analytic diabetics using map reduce and classification techniques. Springer Science +Business Media, LLC, part of Springer Nature 2018
5. Li C-N, Chang S-H (2015) A cloud based type-2 diabetes mellitus lifestyle self-management system. Springer International Publishing Switzerland
6. Karthik D, VijayaRekha K, Sekar S (2014) Profiling water quality using multivariate chemometric method. *Polish J Environ Stud* 23(2):573–576
7. Karthik D, VijayaRekha K (2014) Multivariate data mining techniques for assessing water potability. *Rasayan J Chem* 7(3):256–259
8. Karthik D, VijayaRekha K, RagaSubha S (2014) Classifying the suitability of River Cauvery water for drinking using Data mining classifiers and Authenticating using fuzzy logic, *Int J ChemTech Res* 7(5):2203–2207
9. Karthik D, Vijayarekha K, Arun A (2018) Printing defect identification in pharmaceutical blisters using image processing. *Asian J Pharm Clin Res* 11(3):210–211. (Mar 2018). <https://doi.org/10.22159/ajpcr.2018.v11i3.23407>
10. Karthik D, Karthikeyan P, Kalaivani S, Vijayarekha K (2019) Identifying efficient road safety prediction model using data mining classifiers. *Int J Innov Technol Explor Eng (IJITEE)* 8(10). ISSN:2278-3075. (August 2019)
11. Sungheetha A, Sharma R (2021) Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network. *J Trends Comput Sci Smart Technol (TCSST)* 3(02):81–94
12. Tripathi M (2021) Analysis of convolutional neural network based image classification techniques. *J Innov Image Process (JIIP)* 3(02):100–117
13. Singh AD, Valarmathi B, Srinivasa Gupta N (2019) Prediction of type 2 diabetes using hybrid algorithm. In: International conference on innovative data communication technologies and application. Springer, Cham, pp 809–823
14. Yashodhara PHAHK, Ranasinghe DDM (2021) Early detection of diabetes by iris image analysis. In: Inventive computation and information technologies. Springer, Singapore, pp 615–632
15. Sharma A, Guleria K, Goyal N (2021) Prediction of diabetes disease using machine learning model. In International conference on communication, computing and electronics systems, p 683
16. Nivetha S, Valarmathi B, Santhi K, Chellatamilan T (2019) Detection of type 2 diabetes using clustering methods–balanced and imbalanced pima indian extended dataset. In: International conference on computer networks, big data and IoT. Springer, Cham, pp 610–619

Development of a Neighborhood Based Adaptive Heterogeneous Oversampling Ensemble Classifier for Imbalanced Binary Class Datasets



S. Santha Subbulaxmi and G. Arumugam

Abstract Class imbalance prevails in many real-word datasets. In this paper, a Neighborhood based Adaptive Heterogeneous Oversampling Ensemble Classifier method is proposed to handle class imbalance in datasets. The proposed method adopts an oversampling approach to create a set of balanced representative training datasets. Several base classifiers are built based on those training datasets, and an adaptive heterogeneous ensemble classifier is created. The proposed method is examined with five datasets, and examination results are compared with popular oversampling algorithms. The comparison revealed that proposed method is able to achieve better performance results.

Keywords Imbalanced data · Classification · Ensemble classifier · Heterogeneous ensemble · Multiple classifiers

1 Introduction

Class imbalance prevails in many real-word datasets. Class imbalance occurs when there exists an uneven number of instances between the classes. Class which have less representation in the dataset is called as minority class and remaining classes are called as majority class. The traditional classification algorithms used to be generalize more on the majority class instances. It result in increased misclassification errors in minority class and affects the performance measure “Accuracy”. So the traditional performance measurement approach should be changed for classifying the imbalanced datasets. The difficulties in learning the imbalanced datasets prevail in many application domains like Health, Business, Industry, and Finance.

S. Santha Subbulaxmi (✉)

Department of Computer Science, Madurai Kamaraj University, Madurai, Tamil Nadu, India
e-mail: santhamd3@gmail.com

G. Arumugam

Department of Computer Science, School of Information Technology, Madurai Kamaraj University, Tamil Nadu, Madurai, India

In this paper, a Neighborhood based Adaptive Heterogeneous Oversampling Ensemble classifier is proposed to address class imbalance issue in binary class datasets. The proposed method acquires balanced representative training datasets by oversampling the minority instances in their neighborhood groups and builds base classifiers. An adaptive heterogeneous ensemble classifier is created based on the built base classifiers. Thus the proposed ensemble classifier train multiple classifiers, generalize their decisions and improve the performance capability. The paper is organized as follows: Sect. 2 presents the related literature. Section 3 presents the proposed method. In Sect. 4, the experimental setup, dataset details and experiment results are discussed. Section 5 concludes with the the closing remarks.

2 Related Works

Researchers proposed many methods for handling the class imbalance. The methods can be broadly categorized into.

- Data preprocessing methods
- Algorithmic methods
- Cost-sensitive methods

Data preprocessing methods can be further divided into data cleaning, data sampling, and feature engineering methods. The data cleaning methods tend to improve the classifier performance by removing the instances which are found to hinder some in learning the data distribution. [1] Tomek link is a popular data cleaning method [2]. Condensed nearest neighbor (CNN) and [3] Edited Nearest Neighbor (ENN) are few of the other data cleaning methods. The data sampling methods can be further divided into Undersampling methods, Oversampling methods or Hybrid sampling methods. Undersampling methods balance the data distribution by removing or selecting informative instances in majority classes. Oversampling methods replicates or creates new synthetic instances in minority class. Hybridsampling methods applies both undersampling and oversampling. The Random Undersampling method reduces the majority class instances in random manner and balances the data distribution. The [4] SMOTE, [5] ADASYN, [6] SMOGN, [7–9] Gaussian Noise are the popular oversampling algorithms. The Features Selection methods focuses on the subset of the features that contributes to performance. [10] ReliefF is a popular feature selection method that solves imbalanced data classification. The algorithm level methods do not change the data distribution but they focus on changing the learning behavior of the classifier. [11] twin hyper-sphere SVM, [12] Hellinger distance based C4.4, [13] gravitational based algorithm for Nearest Neighbor classifiers are a few of the popular algorithm level methods. The cost sensitive methods aim at reducing the misclassification cost. [14] MetaCost algorithm is a popular cost sensitive method. [15–17] Researchers experimented and proposed few of the Decision tree methods which are cost sensitive. [18] Researchers experimented on training the Artificial Neural Networks to be cost sensitive and able to

achieve good performance results. [19] Easy ensemble and Balancecascade are a few of the ensemble based classification methods to solve the imbalanced data classification problem. [20] ACO sampling, [21] CHC sampling are a few of meta-heuristic supported imbalanced data classification algorithms. [22] The parameters of SVM is optimized based on the optimization algorithms, GSA, ABC, DE for nonlinear prediction.

3 Proposed Work

The proposed ensemble algorithms that deals with binary class imbalanced dataset is presented in this section. This ensemble method receives training, validation, and testing datasets as inputs. “N” numbers of representative training datasets are generated by oversampling the minority instances in their neighborhood groups in different percentages. N is a hyperparameter. The hyperparameter “Classifiers for training” lists the classification algorithms. For each training dataset, the base classifiers are built based on “Classifiers for training” and the best base classifier among those trained classifiers is selected as “adaptive”. The hyperparameter “C” is used to prune the trained adaptive base classifiers based on their performance. An ensemble classifier is built based on these pruned adaptive base classifiers by using the majority voting method. The proposed ensemble method is presented in Fig. 1 and described as below:

3.1 Partition Training Dataset

The input, training dataset is partitioned into majority (Ma) class dataset and minority (Mi) class dataset.

3.2 Group Minority Instances Based on Neighborhood

For each Minority instance, 5 nearest neighbors (NN) are identified. HVDM (Heterogeneous Value Difference Metric) distance is used in KNN (K Nearest Neighbor) query. Among the 5 nearest neighbors, the number of instances belongs to minority class (MN) is calculated. The Minority neighborhood score (MNS) of the minority instance is estimated from the below equation:

$$\text{MNS} = \text{MN}/5$$

Based on the MNS value, each minority instance is grouped in a neighborhood group. The four neighborhood groups fixed in the method are Safe group, Borderline

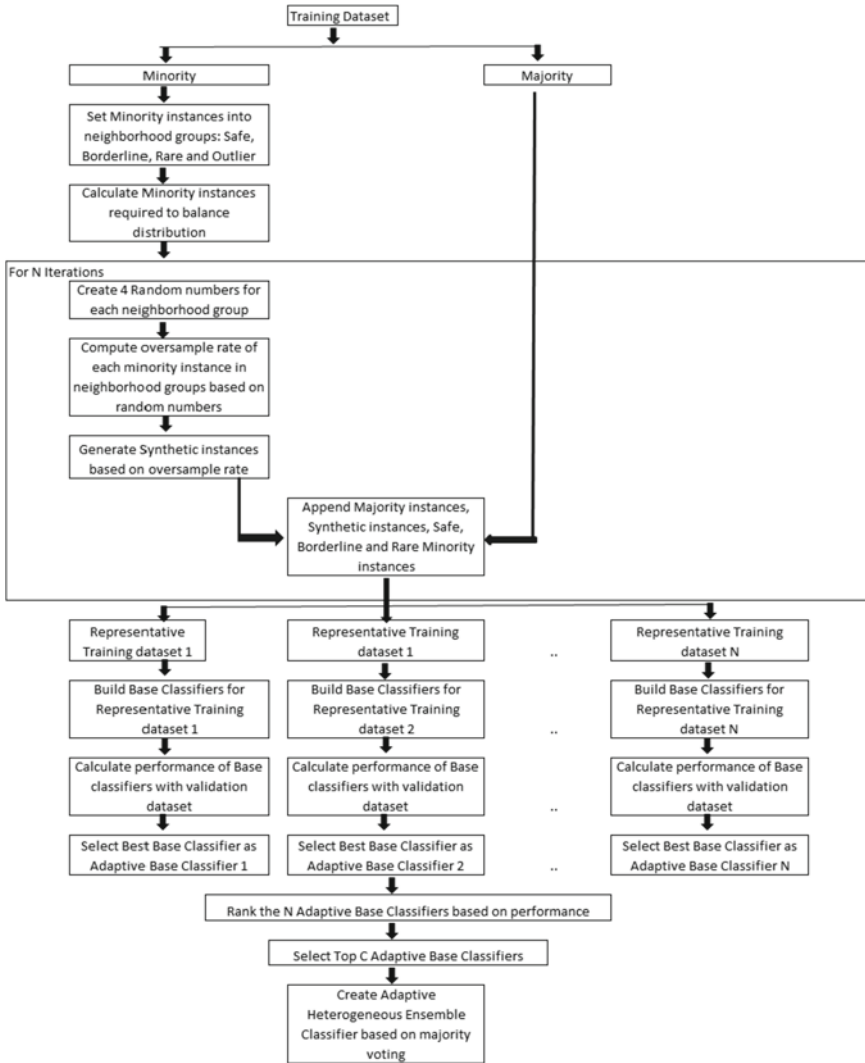


Fig. 1 Pictorial diagram of neighborhood based adaptive heterogeneous oversampling ensemble classifier

group, Rare group, and Outlier group. If MNS is greater than equal to 0.8, the minority instance is set as Safe group. If MNS is between 0.5 and 0.8, the minority instance is set as Borderline group. If MNS is between 0.2 and 0.5, the minority instance is set as Rare group. If MNS is below 0.2, the minority instance is set as Outlier group.

3.3 Create Representative Training Datasets

Representative training datasets are created by oversampling the minority class instances. New synthetic instances are created in minority feature space and the data distribution is balanced. The number of synthetic minority instances required (NR) is calculated as:

NR = Number of instances in Ma–Number of minority instances in Safe, Borderline, Rare groups.

The minority instances are oversampled randomly from the neighborhood groups: Safe, Borderline, Rare, and Outlier. Four random numbers (R1, R2, R3, R4) are generated between 0 and 1. These four random numbers facilitate to derive the percentages of the minority instances to be oversampled from each neighborhood group to balance the distribution. The random number R1 is for Safe neighborhood group, R2 is for Borderline neighborhood group, R3 is for rare neighborhood group and R4 is for Outlier neighborhood group. If any one of the neighborhood group is empty, then the corresponding random number is set as zero. The four random numbers are then divided by their sum and reassigned respectively. Now adding the four random numbers will be equal to 1.

The oversampling rate of each minority instance in the neighborhood group is calculated based on the following equations:

$$GN1 = (NR * R1)/\text{No. of minority instances in safe group.}$$

$$GN2 = (NR * R2)/\text{No. of minority instances in Borderline group.}$$

$$GN3 = (NR * R3)/\text{No. of minority instances in Rare group.}$$

$$GN4 = (NR * R4)/\text{No. of minority instances in Outlier group.}$$

The values of GN1, GN2, GN3, and GN4 are rounded to the nearest integer. These GN1, GN2, GN3, GN4 random numbers are used to create new synthetic minority instances in the neighborhood groups in different percentages. Each minority instance in safe neighborhood group will create GN1 number of new synthetic instances. Each minority instance in borderline neighborhood group will create GN2 number of new synthetic instances. Each minority instance in rare neighborhood group will create GN3 number of new synthetic instances. Each minority instance in outlier neighborhood group will create GN4 number of new synthetic instances..

The synthetic minority instance creation process is described as follows: To create new synthetic instances, each minority instance will search and select the nearest instance by using the HVDM distance. The nearest instance may be a minority instance or a majority instance. If the nearest instance is a minority instance, the numeric features of the new synthetic instances will be created based on the minority and its nearest instance. The difference of numeric features between the said minority instance and its nearest instance is calculated and multiplied with a random number between 0 and 1. This randomized difference is then added with the numeric features of the said minority instance to create the numeric features of new synthetic instances. The values of the nominal features of new synthetic instances will be set randomly either as the nominal features of the said minority instance or as in nominal features of the nearest neighbor. If the nearest instance is a majority instance, the numeric

features of the new synthetic instances will be created by adding Gaussian noise in the numeric features of the said minority instance. The nominal features of the new synthetic instances will have the same values as in the nominal features of the said minority instance.

After the synthetic data generation process, a representative training dataset is created by appending all the new synthetic instances along with the safe, borderline, rare groups of instances in the minority class and then with the majority class instances. Using a similar approach, N number of representative training datasets are created.

3.4 Create an Adaptive Heterogeneous Ensemble Classifier

For each representative training dataset, several base classifiers are built based on the value of hyperparameter “Classifiers for training”. The performances of the trained base classifiers are evaluated based on the validation dataset. The trained base classifiers are measured with the performance metrics, AUC and G-Mean. Among the trained base classifiers, the base classifier which has the highest AUC+G-Mean value is selected as the best base classifier and set as adaptive. Using a similar approach, N number of adaptive base classifiers are generated and they are sorted in descending order based on the AUC+G-Mean performance score and ranked. To create the adaptive heterogeneous ensemble, the top C number of adaptive base classifiers are bagged with the majority voting method where C is the hyperparameter.

4 Experimental Framework and Results

In this section, the experimental framework of the proposed method, dataset details and experiment results are described. R software environment (ver. 4.0.3) is used to implement the proposed method and conduct the experiments. C5.0, Naïve Bayes, Random Forest are set as the “Classifiers for training”. They are implemented by the R package “Caret”, “e1071”, “KlaR”, “randomforest”. Hyperparameter N is set as 100. Hyperparameter “C” is set as 50. The proposed method is compared with few of the popular oversampling methods: SMOGN classifier, Gaussian Noise classifier, ADASYN classifier and they are implemented using the R package “UBL”. Balanced datasets are created by using these algorithms and Random Forest is used to build the classifiers for comparison.

The benchmarking datasets used in the experiments are obtained from the [23] KEEL (Knowledge Extraction based on Evolutionary Learning) data repository. Table 1 summarizes the five dataset details. The datasets are partitioned into training (60%), validation (20%), and test (20%) datasets. G-Mean and AUC are used as Performance evaluation metrics. It is measured from R package “PROC”.

Table 1 Summarization of dataset details

| S.No | Dataset | Attributes | Instances | Imbalance ratio |
|------|------------------------|------------|-----------|-----------------|
| 1 | Glass-0-1-2-3_vs_4-5-6 | 9 | 214 | 3.2 |
| 2 | New-thyroid2 | 5 | 215 | 5.14 |
| 3 | Vehicle2 | 18 | 846 | 2.88 |
| 4 | Ecoli-0-1-4-6_vs_5 | 6 | 280 | 13 |
| 5 | Poker-8-9_vs_6 | 10 | 1485 | 58.4 |

Experiments results of the proposed algorithm’s performance results and the popular oversampling algorithms (SMOEN classifier, Gaussian Noise classifier, ADASYN classifier) are measured and tabulated. The G-Mean value of the proposed algorithm and benchmarking algorithms are shown in Table 2. The proposed algorithm achieve good G-Mean scores for all the five datasets. The G-Mean values of the proposed algorithm are between 0.894427191 and 0.983531374 for two datasets; and they are 1 for three datasets. The experiment results revealed that the proposed method achieve good balance in classifying both majority and minority classes. While comparing with the benchmarked ensembles, the proposed algorithm outperforms them. The AUC value of the proposed algorithm and benchmarking algorithms are tabulated in Table 3. The proposed algorithm is able to achieve better AUC scores for all the five datasets. The AUC values of proposed algorithm are between 0.983531374

Table 2 G-mean value of proposed method and other oversampling algorithms

| S.No | Dataset | Proposed method | SMOEN classifier | Gaussian noise classifier | ADASYN classifier |
|------|------------------------|-----------------|------------------|---------------------------|-------------------|
| 1 | Glass-0-1-2-3_vs_4-5-6 | 1 | 0.974679 | 0.974679 | 0.974679 |
| 2 | New-thyroid2 | 1 | 0.984251 | 0.984251 | 0.984251 |
| 3 | Vehicle2 | 0.983499 | 0.971128 | 0.970698 | 0.970698 |
| 4 | Ecoli-0-1-4-6_vs_5 | 1 | 0.912871 | 0.989743 | 0.989743 |
| 5 | Poker-8-9_vs_6 | 0.894427 | 0.447214 | 0.774597 | 0.774597 |

Table 3 AUC value of proposed method and other oversampling algorithms

| S.No | Dataset | Proposed method | SMOEN classifier | Gaussian noise classifier | ADASYN classifier |
|------|------------------------|-----------------|------------------|---------------------------|-------------------|
| 1 | Glass-0-1-2-3_vs_4-5-6 | 1 | 0.95 | 0.95 | 0.95 |
| 2 | New-thyroid2 | 1 | 0.944444 | 0.944444 | 0.944444 |
| 3 | Vehicle2 | 0.983531 | 0.979025 | 0.950159 | 0.950159 |
| 4 | Ecoli-0-1-4-6_vs_5 | 1 | 0.99 | 0.928571 | 0.928571 |
| 5 | Poker-8-9_vs_6 | 0.998322 | 0.993355 | 0.996656 | 0.996656 |

and 0.998322148 for two datasets and 1 for three datasets. The AUC results revealed the quality of the proposed algorithm. It is evident that the proposed method performs better while comparing with other benchmarking methods.

5 Conclusion

The proposed method has a unique approach for oversampling and uses multiple classifiers to build the ensemble. The minority instances in each neighborhood group are oversampled in random proportions, and representative training datasets are created. Several base classifiers are built, pruned, and a heterogeneous ensemble is created using the majority voting method. The proposed method is evaluated and compared with popular oversampling methods. The results showed that proposed method can handle skewed data distribution in a better manner.

References

1. Tomek I (1976) Two modifications of CNN. *IEEE Trans Syst Man Cybern* 6: 769–772
2. Cover T, Hart P (1967) Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 13(1):21–27
3. Gates G (1972) The reduced nearest neighbor rule (Corresp.). *IEEE Trans Inf Theory* 18(3):431–433
4. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 16:321–357
5. He H, Bai Y, Garcia EA, Li S (2008) ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence), pp 1322–1328. (Jun 1)
6. Branco P, Torgo L, Ribeiro RP (2017) SMOGN: a pre-processing approach for imbalanced regression. In: First international workshop on learning with imbalanced domains: theory and applications. PMLR, pp 36–50
7. Lee SS (1999) Regularization in skewed binary classification. *Comput Stat* (2):277–292. (Jul 14)
8. Lee SS (2000) Noisy replication in skewed binary classification. *Comput Stat Data Anal* 34(2):165–191. (Aug 28)
9. Branco P, Ribeiro RP, Torgo L (2016) UBL: an R package for utility-based learning. [arXiv:1604.08079](https://arxiv.org/abs/1604.08079). (Apr 2016)
10. Kononenko I (1994) Estimating attributes: analysis and extensions of RELIEF. In: European conference on machine learning, pp 171–182
11. Xu Y, Yang Z, Zhang Y, Pan X, Wang L (2016) A maximum margin and minimum volume hyper-spheres machine with pinball loss for imbalanced data classification. *Knowl-Based Syst* 95:75–85
12. Cieslak DA, Hoens TR, Chawla NV, Kegelmeyer WP (2012) Hellinger distance decision trees are robust and skew-insensitive. *Data Min Knowl Disc* 24(1):136–158
13. Cano A, Zafra A, Ventura S (2013) Weighted data gravitation classification for standard and imbalanced data. *IEEE Trans Cybern* 43(6):1672–1687
14. Domingos P (1999) Metacost: a general method for making classifiers cost-sensitive. *KDD* 99:155–164

15. Liu M, Xu C, Luo Y, Xu C, Wen Y, Tao D (2017) Cost-sensitive feature selection by optimizing f-measures. *IEEE Trans Image Process* 27(3):1323–1335
16. Liu Z, Ma C, Gao C, Yang H, Lan R, Luo X (2018) Cost-sensitive collaborative representation based classification via probability estimation with addressing the class imbalance. *Multimed Tools Appl* 77(9):10835–10851
17. Ling CX, Yang Q, Wang J, Zhang S (2004) Decision trees with minimal costs. In: *Proceedings of the twenty-first international conference on Machine learning*, pp 69
18. Zhou ZH, Liu XY (2005) Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Trans Knowl Data Eng* 18(1):63–77
19. Liu XY, Wu J, Zhou ZH (2008) Exploratory undersampling for class-imbalance learning. *IEEE Trans Syst Man Cybern Part B (Cybern)* 39(2):539–550
20. Yu H, Ni J, Zhao J (2013) ACOSampling: An ant colony optimization-based undersampling method for classifying imbalanced DNA microarray data. *Neurocomputing* 101:309–318
21. Drown DJ, Khoshgoftaar TM, Seliya N (2009) Evolutionary sampling and software quality modeling of high-assurance systems. *IEEE Trans Syst Man Cybern Part A Syst Hum* 39(5):1097–1107
22. Raj JS, Vijitha Ananthi J (2019) Recurrent neural networks and nonlinear prediction in support vector machines. *J Soft Comput Paradigm (JSCP)* 1(01):33–40
23. Alcalá-Fdez J, Fernández A, Luengo J, Derrac J, García S, Sánchez L, Herrera F (2011) KEEL data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. *J Multi-Valued Logic Soft Comput* 17:255–287

Real Time Facial Emotions Detection of Multiple Faces Using Deep Learning



Ankita Kshirsagar, Neetesh Gupta, and Bhupendra Verma

Abstract Facial expression is a gesture developed with the facial muscles that communicates the individual's emotional expression. As Artificial Intelligence(AI) and Computer Vision technology advances, machines must learn to recognize human emotions in order to communicate effectively. Using the Deep Convolutional Neural Network (DCNN) technique, face expression are examined to detect emotion using deep learning. The proposed system is trained and evaluated in real time using the ADFES-BIV dataset. In the training phase, transfer learning technique is used to extract features from pictures as well as live video and a haar-cascade classifier is employed to detect the multiple faces in a frame. The Facial Action Coding System detects five universal emotions: happy, sad, neutral, surprise and anger and some micro emotions such as contempt, embarrass, fear, disgust and pride also can be detected by a person's face. It can recognize many facial emotions at the same time using deep learning. The system achieved a training accuracy of 81.67%.

Keywords AI · Computer vision · Deep learning · Facial emotion · Image · Live video · Multiple faces · Real time · Transfer learning · ADFES-BIV

1 Introduction

Facial expression is a way of expressing individual's feelings and emotions without speaking. Facial Expression is a non-verbal communication. People around the world can connect without knowing each other's language through facial expressions and gestures. Developing an automatic Facial Expression Recognition System (FERS) for facial activity analysis is a constant source of inspiration for image processing researchers. Machine Learning (ML) is a basic component of AI, and Emotion Recognition in photos and videos is the most popular use of AI. In future, robots/machines will be required to comprehend human emotions to communicate effectively. Since the use of emotion recognition has grown over the last decade, there has been a lot

A. Kshirsagar (✉) · N. Gupta · B. Verma
Technocrats Institute of Technology, Bhopal, India
e-mail: ankita.kshirsagar@gmail.com

of research in the field of picture processing. Using Transfer Learning, this research study focuses on the deep learning notion of recognizing a person's emotions in real time as well as through images. The standard Facial Expression Recognition (FER) process consists of three major steps: image preprocessing, feature extraction, and expression characterization. The placement of the facial areas is taken into account when detecting faces and the output must be vital and useful for future analysis to increase the learnt model's accuracy. Due to change in emotions, face characteristics alter frequently and as a result, comparable components in diverse representations of emotions have different placements, which aid in accurately re-cognizing the emotions.

2 Literature Review

Face-regions-based facial expression recognition [1]: Author developed an Interface technique to recognize facial landmarks using basic states of emotion recognition. The ROIs are then enlarged and partitioned into blocks in the preprocessing stage before executing the feature extraction to create a face feature as a multi-class Support Vector Machine for emotion detection. The dataset used is CK, and accuracy is 89.85%. Deep Networks are used in a new method for automatic face expression recognition and classification [2]. Author employed a three-layer CNN network, followed by max pooling and ReLU, to recognize emotions. The FER2013 dataset was used to train and the RaFD dataset for testing. The accuracy achieved is 68%. Online Facial Expression Recognition system [3]: An online facial expression recognition framework based on a customized display using based on personalized photos. The network was trained using a Personalized Gallery of 9 people, with 85% accuracy for familiar people & 65% for strangers. The features that include SIFT, HG, and MFs were retrieved after preprocessing. GEMEP-FERA dataset is used, and overall accuracy is 80%. Recognize the temporal periods of facial motions completely automatically [4]. Deep Learning for Detecting Human Emotions from Facial Expressions [5, 12, 13]: A dataset including a variety of facial expressions to train the machine. As a classifier, a 3-layered CNN is used which have the highest accuracy rate when compared to others. The dataset was FER2013, and the accuracy was 79.8%.

Kottursamy employs a variety of deep learning algorithms for emotion detection and sentiment analysis. To dig into the forecasting of children's conduct based on their emotions in greater depth. [6] Proposed research by Shakya tries to accurately distinguish faces from a distance. The Gaussian derivative filter reduces the feature size in the storage element, improving the recognition ratio's speed [7, 14, 15]. A supervised learning is validated with haar cascade and Fisher face classifier on COHN-KANADE dataset with 65% identification rate, to create an eight-class facial expression classification system (7 basic emotions along with neutral) [8].

3 Problem Statement

Most of the researchers have focused FER research in analyzing 7 types of basic human emotions for a single human face in static image only. This study paper demonstrates the real time and static facial expression identification of multiple faces in a frame. It can also identify the ten different types of human emotions in a frame. Such system can be used in medical, monitoring, supervising, education, security etc. purposes. The emotion detection of multiple faces in real time achieved with the Deep Learning technology. Details are described in Sects. 4–6.

4 Methodology

In the proposed method, system is trained with Transfer Learning (TL) by using the ADFES-BIV dataset [653 images with 10 emotions (Happy, Sad, Anger, Neutral, Embarrass, Fear, Contempt, Pride, Disgust and Surprise).]TL for AlexNet is used to train the dataset by fine-tuning the pre-trained model. The model is then used to predict emotions in real time via the webcam or static images provided manually once trained.

4.1 Dataset

The model is trained by using the Facial Expression Recognition (ADFES-BIV) dataset which can be referring here [10]. 13 frames were extracted from the video and converted to gray scale. To train model, 653 images of ten emotions are used (Fig. 1).



Fig. 1 Sample photographs taken from ADFES-BIV Dataset [11]

Table 1 Sample count taken from each class for the training and testing process

| Types of emotions | Total no. of images | No. of images used for training | No. of images used for test |
|-------------------|---------------------|---------------------------------|-----------------------------|
| Happy | 98 | 78 | 20 |
| Angry | 59 | 47 | 12 |
| Disgust | 53 | 42 | 11 |
| Embarrass | 58 | 46 | 12 |
| Fear | 61 | 48 | 13 |
| Pride | 55 | 44 | 11 |
| Surprise | 64 | 51 | 13 |
| Neutral | 87 | 69 | 18 |
| Contempt | 56 | 44 | 12 |
| Sad | 62 | 49 | 13 |

Proposed model can detect 10 emotions. Out of 653 images, 80% are used for training while 20% are used for validation, Please find details in the below table (Table 1).

4.2 Convolutional Neural Network

The model employs the Convolutional Neural Network (CNN) method, which provides the highest level of accuracy when compared to others. CNNs are commonly used in image/object/face recognition, among other applications. AlexNet is CNN architecture had eight layers: the first five were convolutional layers, followed by max-pooling layers in some cases, and the final three were fully linked layers. It employed the non-saturating ReLU activation function, which outperformed tanh and sigmoid in terms of training performance. AlexNet also developed new ways to reduce overfitting (Fig. 2).

Maxpooling The pooling layer is a component of a CNN, and max pooling is the most well-known way of pooling. It selects the greatest number of features from the feature map. As a result, the feature map created after max pooling layer contains the most relevant features from the prior features map.

Exponential Linear Unit (ELU) Elu was employed to set the threshold value in suggested model. It is an activation function that gives accurate results and converge the cost to zero faster. In comparison to other activation functions, ELU has an additional constant alpha, which must be positive. Except for the negative inputs, it's comparable to RELU. When ELU's output equals alpha, it begins to smooth slowly, whereas RELU smoothes quickly. RELU is a non-linear activation function $f(x) = \max(x, 0)$.

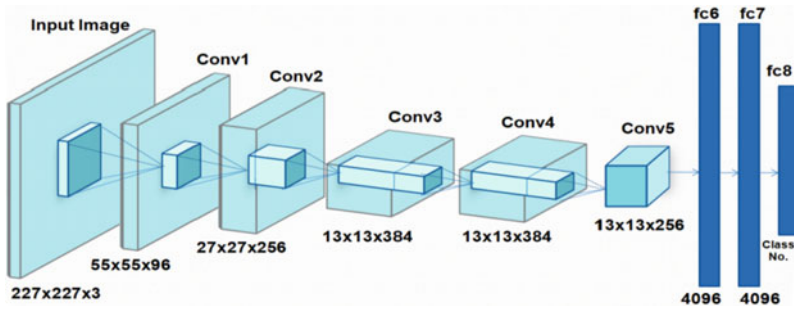


Fig. 2 Architecture of AlexNet

SoftMax layer is utilized as the last layer at the conclusion of the network in Deep Learning, notably in CNN, to produce the actual probability scores for each section or class label. It’s quite valuable because it transforms to a normalized frequency distribution, which can then be used as an input by other systems. Just before the output layer, it’s implemented.

Fully-Connected Layer Each node in a Fully-Linked Layer is connected to all nodes in the preceding layer, whereas the convolutional layer creates local connections.

4.3 Transfer Learning

AlexNet was trained on over a million photos and can categorise them into 1000 different item categories. For a wide spectrum of pictures, the network learnt rich feature representations. The network receives an image as input and returns a label for each object in the picture, as well as probabilities for each of the object categories.

Transfer learning is frequently used in deep learning applications. A pre-trained network can be selected and used it as a starting point to learn a new task. Fine-tuning a network with transfer learning is usually much faster and easier than training a network with randomly initialized weights from scratch. Features learned by transfer learning can be transferred to a new task using a smaller number of training images. By fine tuning pre-trained AlexNet i.e. replacing fully connected, classification and Softmax layers [9] (Fig. 3).

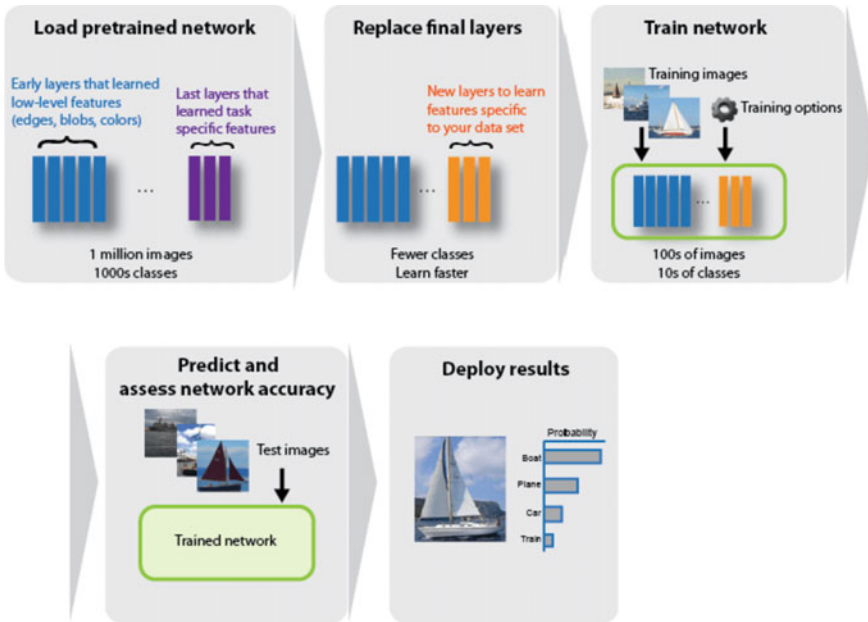


Fig.3 Process of transfer learning [9]

5 Implementation

As per proposed model explained in the Fig. 4, first stage is training process to train the network with pre-trained model by replacing some layers of AlexNet network with some training options. Once model is trained with given data then Face detection, facial feature extraction and emotion prediction of multiple faces in live feed or through webcam is achieved. Performed steps are elaborated below:

5.1 Pre-processing of Data

To create a dataset, 13 frames were extracted from each input video of ADFES and converted to grey-scale. Histogram Equalization was applied to the frames to adjust the contrast and get an image clearer than the original.

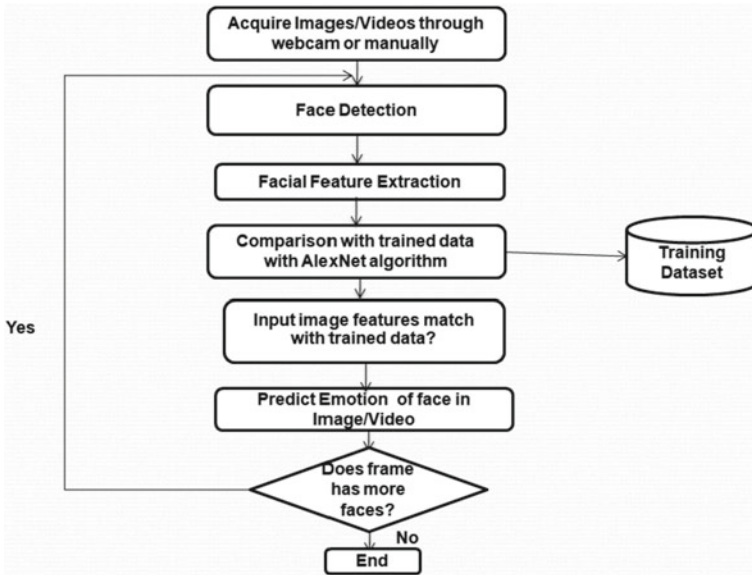


Fig.4 Flowchart for real time facial emotion detection system for multiple faces

5.2 Training Model

To Train model with particular dataset below steps are performed:

Load Data: Above set of images are copied to a Train folder which is pre-processed in $227 \times 227 \times 3$ size as AlexNet requires input size in the same format.

Replacement of Layers: Replace the last 3 layers Fully Connected, Softmax and Classification layer.

Training Dataset: Dataset is trained with pre-learned AlexNet and results are saved.

5.3 Face Detection

Faces are detected in the live video frames or static images using haar cascade classifier to detect multiple faces in image.

5.4 Face Extraction

Key characteristics such as eyes, brows, and lips are retrieved from the facial picture when the face is recognised. Dimensionality reduction is linked to feature extraction. Feature extraction begins with the raw data and builds derived features to understand both useful and non-redundant features.

5.5 Resize Detected Faces

In the previous process, faces regions are detected in a live video taken from webcam are cropped to extract an actual face region. It will be cropped in size $227 \times 227 \times 3$ to extort most accurate facial features to predict emotion.

5.6 Emotion Prediction

Using pre-trained model, on basis of the extracted facial feature emotion is predicted. Later, if there are more faces in the live feed taken through webcam then it detects through Viola Jones algorithm and above process is repeated. Same process is followed when until webcam is ON. In case of static images, above process is followed and program is ended once all faces are detected and emotions are predicted for others.

Below steps are performed to evaluate methodology:

Step 1: For Image - Select image from the directory or switch on Webcam.

Step 2: Extract a frame from real time video.

Step 3: Detect a Face from the input image/frame using facial features.

Step 4a: If human face is not detected then ignore frame of video and extract next frame.

Step 4b: If human face is not detected in an image then warning message is displayed and process ends.

Step 5: Crop detected face from an image or a frame.

Step6: Resize cropped face to the $227 \times 227 \times 3$ dimensions.

Step 7: Classify facial features.

Step 8: Predict emotion of the face in an image.

Step 9: Check if more faces are in an image then repeat steps from Step 2.

Step 10: When no more faces are left for prediction then loop ends and emotions of all faces are displayed in an image/video (Fig. 5).

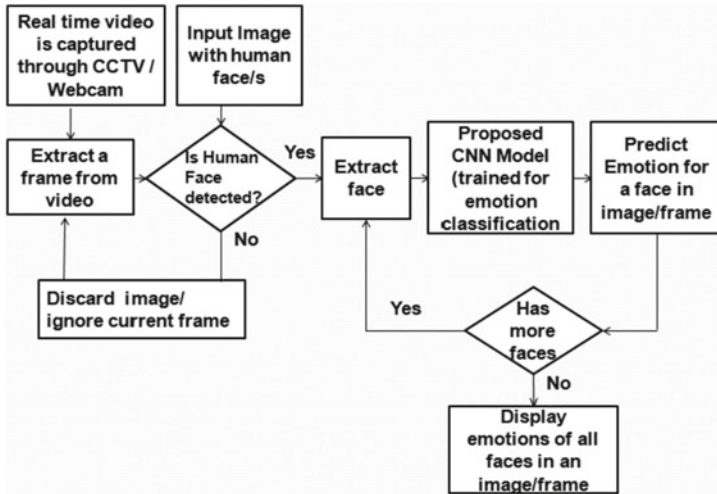


Fig. 5 Block diagram of overall system

6 Experiment Result

The proposed model is trained for 40 epochs at a learning rate of 0.0001 which took 1 h 49 min to complete training including validation loss calibrations in MATLAB 2020a. This model has used pre-trained AlexNet and the Haar-cascade classifier to categorize the emotions captured live from the webcam. Happy, Sad, Angry, Surprise, Pride, Disgust, Contempt, Embarrass, Fear and Neutral are the ten emotions that proposed model can detect. The model’s accuracy during training is 81.67%. CNN has been run numerous times to assess the model’s accuracy on a device with a CPU and 4 GB of RAM. To improve accuracy, overfitting issue has been handled by data augmentation and dropout layers method. Below are the accuracies based on architecture, using Transfer Learning (Table 2).

Precision explains how many of the correctly predicted cases actually turned out to be positive. Recall explains how many of the actual positive cases we were able to predict correctly with proposed model. F1 Score would be a preferable statistic to utilise to find a balance between Precision and Recall and there is an unequal class distribution (Table 3).

Table 2 Comparison table with similar methods reported in the literature

| Paper | Method used | Database | Accuracy (%) | Static image | Real-time video | Multiple faces |
|----------------------|-------------------------|---------------------|--------------|--------------|-----------------|----------------|
| Lekdiou [1] | Support vector machine | CK | 89.85 | Yes | No | No |
| Salunke [2] | CNN | FER2013 & RaFD | 68 | Yes | No | No |
| Hong[3] | Elastic graph matching | Personalized photos | 65 | No | Yes | No |
| Babajee[5] | CNN | FER2013 | 79.80 | Yes | No | No |
| Proposed methodology | CNN + Transfer Learning | ADFES | 81.67 | Yes | Yes | Yes |

Table 3 Precision, Recall and F1 score details of proposed method

| Model | Accuracy | Precision | Recall | F1 Score |
|----------------------|----------|-----------|--------|----------|
| Proposed methodology | 81.67% | 0.85 | 0.83 | 0.84 |

An $N \times N$ matrix is used to evaluate the performance of a classification model, where N is the number of target classes, known as confusion matrix. This matrix compares the actual goal values to the machine learning model's predictions and provides comprehensive picture of how well classification model is working and the types of errors it makes. Here is the Confusion Matrix of the proposed model which defines the performance of training model (Fig. 6).

Below is the training graph where training and validation accuracies are plotted. Dark blue line indicates the smooth training graph and light blue line indicates normal training progress (Fig. 7).

Below are testing results from the webcam video/live feed and static image (Figs. 8, 9 and 10).

Confusion Matrix

| | | | | | | | | | | | | |
|---------------------|---------------------|--------------|----------------|--------------|--------------|----------------|--------------|----------------|--------------|--------------|----------------|----------------|
| | Angry | Contempt | Disgusted | Embarrass | Fear | Happy | Neutral | Pride | Sad | Surprised | | |
| Output Class | Angry | 4 6.7% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% | |
| | Contempt | 0 0.0% | 3 5.0% | 0 0.0% | 0 0.0% | 1 1.7% | 0 0.0% | 2 3.3% | 0 0.0% | 0 0.0% | 50.0% 50.0% | |
| | Disgusted | 0 0.0% | 1 1.7% | 7 11.7% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 87.5% 12.5% | |
| | Embarrass | 0 0.0% | 0 0.0% | 0 0.0% | 5 8.3% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% | |
| | Fear | 0 0.0% | 1 1.7% | 0 0.0% | 0 0.0% | 3 5.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 75.0% 25.0% | |
| | Happy | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 7 11.7% | 0 0.0% | 0 0.0% | 0 0.0% | 100% 0.0% | |
| | Neutral | 0 0.0% | 2 3.3% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 5 8.3% | 0 0.0% | 0 0.0% | 71.4% 28.6% | |
| | Pride | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 1 1.7% | 5 8.3% | 0 0.0% | 83.3% 16.7% | |
| | Sad | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 4 6.7% | 100% 0.0% | |
| | Surprised | 0 0.0% | 1 1.7% | 0 0.0% | 0 0.0% | 2 3.3% | 0 0.0% | 0 0.0% | 0 0.0% | 0 0.0% | 6 10.0% | 66.7% 33.3% |
| | | 100% 0.0% | 37.5% 62.5% | 100% 0.0% | 100% 0.0% | 50.0% 50.0% | 100% 0.0% | 62.5% 37.5% | 100% 0.0% | 100% 0.0% | 100% 0.0% | 81.7% 18.3% |
| | Angry | Contempt | Disgusted | Embarrass | Fear | Happy | Neutral | Pride | Sad | Surprised | | |
| | Target Class | | | | | | | | | | | |

Fig. 6 Confusion matrix post training ADFES-BIV dataset

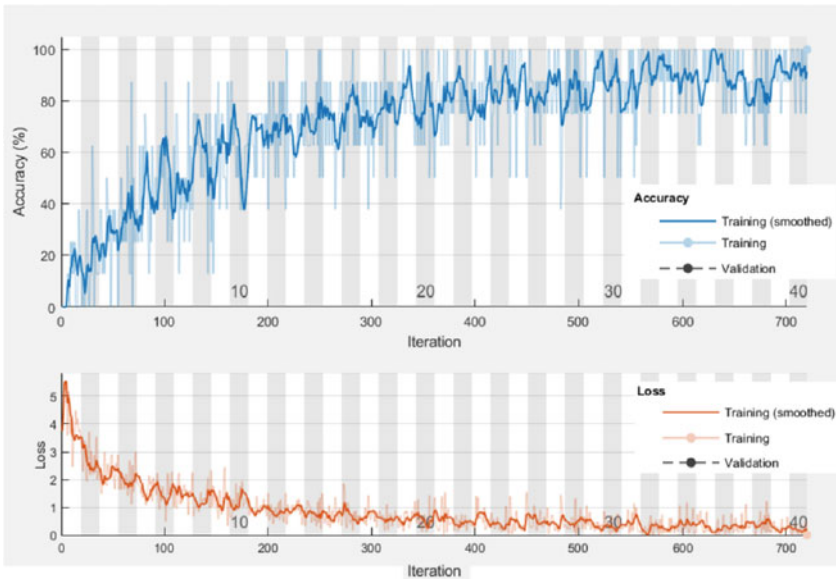


Fig. 7 Training progress graph



Fig. 8 Multiple faces detected in a frame and recognized as ‘Happy’ emotions in an image

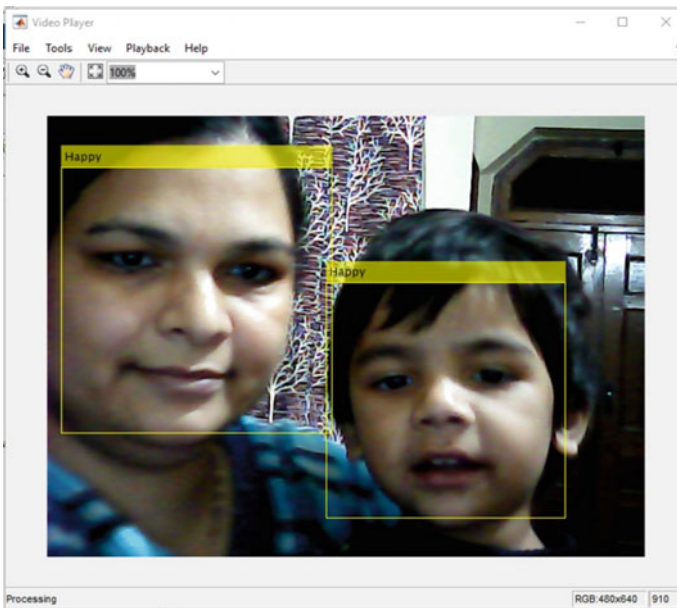


Fig. 9 Two faces detected in a frame of real time video and emotions detected as ‘Happy’

7 Conclusion

FERS (Automatic Facial Expression Recognition System) offers a wide range of applications, ranging from security to engaging with people who are verbally challenged. With the proposed approach system can detect emotions of the multiple faces real time as well as from static images with improved accuracy. During testing it has successfully identified the 10 types of emotions in a single frame.

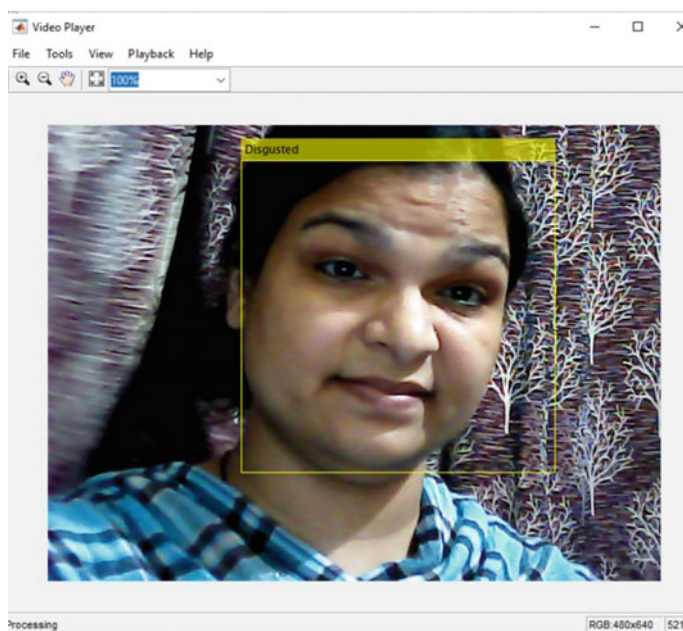


Fig. 10 A face detected in a frame of real time video and recognized “Disgusted” emotion

8 Future Work

The proposed system cannot detect emotions when face is tilted or at different angle due to lack of the database. This approach can be employed on larger datasets in the future. With the help of different machine learning algorithms model can be enhanced so that system can learn in real time and produce result on run time.

References

1. Khadija Lekdioui, Rochdi Messoussi, Youness Chaabi, Raja Touahni.: “Facial Expression recognition using face-regions.” 3RD International conference on advanced technologies for signal and image processing–Atsip’2017, May 22–24, 2017, Fez Morocco, pp. 1–6.
2. Vibha. V. Salunke, Dr. C.G. Patil.: “A new approach for Automatic Face Emotion Recognition and Classification based on deep networks.” 2017 (ICCUBEA), Pune, 2017, pp. 1–5.
3. Hai Hong and C. Von Der Malsburg.: “Online Facial Expression Recognition using Personalized Gallery.” 3 rd IEEE International Conference on Automatic Face and Gesture Recognition, Nara, 1998, pp. 354–359.
4. Michel F. Valstar and Maja Pantic.: “Fully automatic recognition of the temporal phases of facial actions.” IEEE Transactions on Systems, Man and Cybernetics- Part B: Cybernetics, Vol. 42, No. 1, pp. 28–43, February 2012.
5. P. Babajee, G. Suddul, S. Armoogam and R. Foogoa.: “Identifying Human Emotions from Facial Expressions with Deep Learning.” (ZINC), Novi Sad, Serbia, Year: 2020, pp. 36–39.

6. Kottursamy, Kottilingam.: "A review on finding efficient approach to detect customer emotion analysis using deep learning analysis." *Journal of Trends in Computer Science and Smart Technology* 3, no. 2 (2021): 95–113.
7. Shakya, Subarna.: "Multi Distance Face Recognition of Eye Localization with Modified Gaussian Derivative Filter." *Journal of Innovative Image Processing* 3, no. 3 (2021): 240–254
8. Suneeta, V. B., P. Purushottam, K. Prashantkumar, S. Sachin, and M. Supreet.: "Facial expression recognition using supervised learning." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 275–285. Springer, Cham, 2019.
9. <https://in.mathworks.com/help/deeplearning/ug/transfer-learning-using-alexnet.html>.
10. <https://aice.uva.nl/research-tools/adfes-biv/adfes-biv.html>.
11. https://www.researchgate.net/publication/51587980_Moving_Faces_Looking_Places_Validation_of_the_Amsterdam_Dynamic_Facial_Expression_Set_ADFES.
12. Yogita, Neetesh Gupta.: "Integrity auditing with attribute based ECMRSA algorithm for cloud data outsourcing" 3rd International Conference on Intelligent Sustainable Systems, ICISS 2020, pp. 1284–1289.
13. S. Raj, K. Singh, N. Gupta, B. Verma, S. Karsoliya.: "High Accuracy of Hybrid IDS System using Evidence Theory and SVM ML Technique", *International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, 2021, pp. 1261–1264.
14. A. Nagar, N.K., Gupta, U. Singh.: "Precedency with round robin technique for load balancing in cloud computing", *Lecture Notes on Data Engineering and Communications Technologies*, 2020, 46, pp. 461–475.
15. K. Gajbhiye, N.K., Gupta.: "Real Time Twitter Sentiment Analysis for Product Reviews Using Naive Bayes Classifier", *Lecture Notes on Data Engineering and Communications Technologies*, 2020, 46, pp. 342–350.

Mobile Application for Online Pharmacy: A-Pharma App



P. S. JosephNg, A. A. A. Al-Maari, K. Y. Phan, J. T. Lim, and E. H. Lim

Abstract The global spread of diseases has impacted people negatively causing them to be worried about their health which in turn makes them visit the physicians or pharmacies frequently. Online pharmacies are one of the technologies that will be in high demand in the future. Though there are online pharmacy apps available, these apps could not provide the facility to check the availability of medicines in different pharmacies. As an impact, users seem to visit physically and check various stores, which is time consuming. To overcome this inconvenience, this study has proposed to develop a mobile app, A-Pharma App which is an online medical centre. Once the data has been acquired regarding the online pharmacy applications using survey and interview, the results obtained have been utilized to build the features of the proposed application. The presented application provides consultation with the pharmacist and people can get their medicines delivered at their doorstep. The application attempts to satisfy the user's requirements regarding their health concern to a great extent. Furthermore, this paper determines that the proposed A-Pharma App provides all the features necessary for an online pharmacy compared to the other similarly existing apps.

P. S. JosephNg (✉)

Faculty of Data Science and Information Technology, INTI International University Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

e-mail: joseph.ng@newinti.edu.my

A. A. A. Al-Maari

Institution of Computer Science and Digital Innovation, UCSI University, UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia

K. Y. Phan · J. T. Lim · E. H. Lim

Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia

e-mail: phanky@utar.edu.my

J. T. Lim

e-mail: limjt@utar.edu.my

E. H. Lim

e-mail: ehlim@utar.edu.my

Keywords Pharmacy · Online pharmacy · Pharmacist · Patient · Delivery boy · Medicine · Order · Location · Chat

1 Introduction

Healthy lifestyle is the primary factor that every person is most concerned about. A healthy person has a big role in building his country, rebuilding society, providing its services, and doing his work to the fullest [1]. Since health is a very essential need, there are many hospitals, health centers, and pharmacies that have been established to provide health services and medicines [1]. The best interaction strategy for doctors in medication decisions, for pharmacists to administer medicines was the handwritten prescription. With the development of technology, there is also a concern for health on the technical side, which is, the patients have an opportunity to consult the doctor online [2, 3]. Nowadays, life has relied on technology in almost every field to make life more flexible and easier, particularly after the COVID-19 pandemic [3, 4]. The world concerned with its health, followed the protocol to stay at home to reduce the spread of COVID-19 and thus technology effectively played a great role in the most important areas of life such as education, shopping, and health [4, 5]. Also, mobile technology has played a significant role in healthcare, for example, providing individuals with information to track their health conditions, providing healthcare drug tips, and providing medical professionals with resources to monitor their patients [6, 7].

When a patient wants to get his/her medicine, he or she has to visit the pharmacy which often has a queue, with a long waiting time to get the prescribed drugs [10, 11]. Therefore, since mobile communication technology is increasingly evolving, mobile apps can be used to address certain problems and inconveniences for patients. Smart technologies for mobile applications, which consists of the use of Near Field Communication (NFC) Technology and Android mobile application have been developed [12, 13]. An application developed for the patients must be introduced with accurate knowledge to enjoy great benefits [8].

Pharmacists and the people suffer due to the hassle of long waiting hours in the pharmacy. Much of the dissatisfaction of the technicians are due to the number patients waiting to receive their medicine and the patients end up spending too much money on items they do not require. Though there are online platforms available for medical services, they do not provide the communication with the pharmacist, hence the patient may get confused of the brand names because the same brand may exist, but with different drugs. [14, 15]. Patients must visit the pharmacy and take the medicines at the proper time [10]. However, illness make them feel lethargic to visit a doctor or even go to the medicals to get the necessary medicines, which makes the disease worse and makes them weaker [9]. The proposed application will help the user to take proper medicines, in the right quantity and at the appropriate time.

Smartphones have now become a necessity in people's lives. Since everyone has a smartphone with which it is feasible to communicate or order products using

Table 1 Comparison of features in contemporary mobile applications

| Features | Pharm easy | Care n cure | Belshifa |
|-----------------------------------|------------|-------------|----------|
| Location (Search pharmacy nearby) | No | No | No |
| Pharmacist consultation | No | Yes | No |
| Upload prescription | Yes | No | No |
| Home delivery | Yes | Yes | Yes |
| Track the order | Yes | Yes | No |
| Online Payment | Yes | No | Yes |

application, an online mobile application, A-pharma App has been developed. This application presents the patients a good feature, which is, the user can locate the nearby pharmacy and can contact the pharmacist by chatting or calling and describing his necessities without visiting the pharmacy or consulting the pharmacist physically or can upload the prescription to purchase the medicines [9, 10]. Moreover, it helps the patient to order their prescription safely and efficiently by direct contact with a particular pharmacy [11].

There are a few already existing online pharma applications. PharmEasy is the most widely used mobile application among these apps. This application can be also used to get medication and provide pharmacy operations [17]. Other comparable mobile applications include Care n Cure and Belshifa [18]. To further explain the comparability of these applications, the table of comparison is illustrated below.

Table 1 shows that the existing applications have home delivery service. However, PharmEasy has the most features among the others, and has the upload prescription as its main feature [13]. On the other hand, some applications do not provide a track of the order or online payment. Another common attribute that all apps possess is home delivery option to get the order. In most cases, patients do not have a prescription. So they may have to consult a pharmacist to get their appropriate medicine. People conscious on time prefer to contact the nearby pharmacy. However, most of the applications do not provide this feature [14].

2 Research Questions

RQ1: What effect do mobile apps have, when patients buy their medications?

RQ2: How to reduce the cost of purchasing the medicine from the pharmacy?

RQ3: How could technology enhance the reputation of the pharmacy and encourage patients to buy online?

3 Research Objective

ROI: To save and reduce the time with easy handling for a patient to get the medicines quickly with the mobile application.

To assist users, save their time and effort, to facilitate users in checking the availability of medicines in different stores, the mobile application comes out as a platform. It could give more convenience for people whenever and wherever they are.

RO2: To save money and keep the cost lower.

To help users to reduce the cost consumption, since users can only order the required medicines and do not need to travel physically looking for the medicals.

RO3: To create a technologically advanced app that increases the reputation of the app and pharmacies.

To develop a mobile application that encourages people to describe their critical condition to the pharmacist or can upload their prescription.

4 Problem Statement and Hypothesis

The goal of this study is to develop an online pharmacy based on location to provide medicine that could reduce cost, effort, and time, without the need of going to the pharmacy directly. The hypothesis for every criterion extracted from the objectives are summarized in Fig. 1.

H1: Patients will save time by using A-Pharma app.

Providing application based on location will allow users to use one application instead of using different apps for different needs i.e., ordering online from the pharmacy and finding the nearby pharmacy. One of the most important feature is that



Fig. 1 Research model

time may be saved by purchasing medicines online because all of the medication may not be available at the same medical shop. Sometimes medicine may not be easily accessible in the market, but the customers may require it immediately; therefore, they must visit each store to inquire about the drug’s availability which consumes time.

H2: Patients can reduce cost using A-Pharma app.

By using an online application to order specific medicines, users need not go to the pharmacy and spend money because the prescription is available online. While the patient going to buy a particular medicine from the pharmacy, may end up buying another unintended product. On other hand, travel cost can be brought down especially if the medicine is unavailable nearby and requires long drive looking for another store.

H3: Pharmacies branding will be improved using A-Pharma app.

Implementing communication features with the pharmacist in the app will make the app advanced. It will bring comfort to the patient and relative privacy, which might allow patients to be encouraged to inquire about uncomfortable issues.

5 Methodology

Based on the study questions stated previously, a few surveys have been performed and interview questions have been framed. The survey approach includes Likert scale questions. The interview questions delve further into the client’s point of view. During the data collecting phase, all surveys and interviews are conducted. The data are gathered using the approach indicated in Table 2 [16–26].

Based on Table 2, The research dimension is illustrated in a sequential design that explains each stage of data collection. Qualitative reasoning is used to generalize the opinions of a large number of respondents from all over the world. Using survey as the tool, allowed the research team to collect more data easily. The use of interviews allowed the study team to delve into greater detail about their opinions and rationale. The survey has been completed by seventy-four respondents. The interview consisted of twenty-five respondents. All the data obtained during the interview have been utilized to confirm the reliability of the survey responses.

Table 2 Research methodology

| | |
|----------------------|---|
| Research dimension | Explanatory sequential design |
| Research methodology | Quantitative and qualitative perceptive |
| Research method | Survey and interview |

6 Results and Findings

According to Fig. 2, the percentage of the respondents who think the pharmacy app's cost is sustainable in the long run is 90% (16.2% + 54.1% + 18% = 90%). Patients when looking for a particular medicine, may have to travel quite a distance that costs for the travel. As a result, the online pharmacy app is considered sustainable, due to less or no wastage of fuel, since the app gives options from a variety of pharmacies. This system saves cost in the long run by showing more than one pharmacy. Hence many respondents believe that the app saves the cost.

As per the results of data collection shown in Fig. 3, a total of 72% of individuals who responded to the survey, agree that online pharmacy apps based on location reduces the time consumed with less effort from the patient. People use this app to look for the availability of the medication in many pharmacies shown on the map. When the user clicks on the red mark on the map, it displays all locations for the availability of that medication. Moreover, the search function lessens the time by displaying suggestions before typing out the name of the medicine or the pharmacy. On other hand, the delivery boy delivers swiftly the order to the user's doorstep. Using the app, reduces the time consumption, and is more suitable for people who regularly visit the pharmacy.

As seen in Fig. 4, 72% (55% + 22% = 72%) of the respondents agree or strongly agree that the app gives good feedback for the pharmacies. Due to the implementation of an online pharmacy app, the branding for the registered pharmacies is

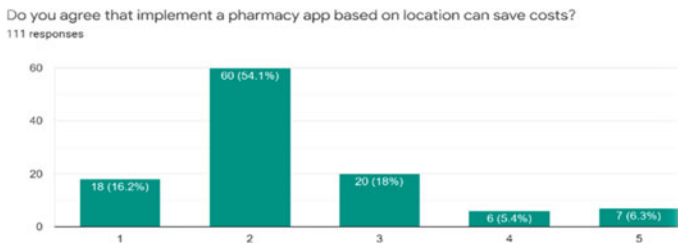


Fig. 2 Survey results related to cost



Fig. 3 Survey results related to efficiency

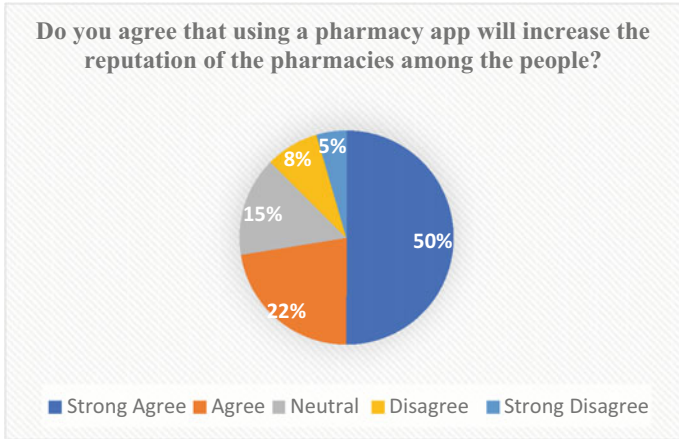


Fig. 4 Survey results related to branding

considerably improved. Using applications is a major area of technological study. When the app is being introduced for fulfilling pharmacy needs, and utilized in place of the conventional direct visit to the pharmacy, it improves the pharmacy’s professionalism. Furthermore, it demonstrates that the pharmacy stays up with the newest technology, which improves the registered pharmacy’s public image in terms of technology adoption. As a result, it provides a higher reputation to the pharmacy based on the service provided in the app. Moreover, a pharmacy that is using the app attracts people to order online. Meanwhile, it contributes to the pharmacy being a contemporary one with excellent facilities.

The ease of the mobile application as an online pharmacy has received positive feedback from interview respondents. From the feedback, this feature has a high positive perspective. Table 3 shows some important questions and responses regarding the cost and efficiency of the app while Table 4 shows more details about the branding of the app.

From the analysis, all respondents agree that using this app reduces the patient’s efforts since all process can be done online. Moreover, the features provided such as delivering the medicine, consulting either doctor or pharmacist, and locating the nearby pharmacy, attract the user to use the app and increase the reputation of the pharmacy. Thus, this app will be popular in the future and many pharmacies will register to increase their profit. The functions of the app are shown in Fig. 5.

As shown in Fig. 5, the main function of the app is the location page where the map displays with the read mark all the nearby pharmacies in the user’s location. Moreover, when the user clicks on the red mark, it shows the name of the pharmacy, and redirects to the selected pharmacy’s page. Then the user can view the category of the medicines and click on the required medicine to view the medicine details. By clicking on the button “ADD TO CART” the medicine can be added and ordered from the pharmacy. Moreover, another main function is the search option. When the

Table 3 Interview results related to efficiency and cost

| Question | Output |
|---|--|
| <i>In what ways do you think, using a pharmacy app will save the cost?</i> | The majority of the respondents agree that by reducing the drive cost to and from the store, and at times the travel cost in search of a particular drug, the Pharmacy app can save the cost |
| <i>Do you think that a pharmacy app can save costs? Why?</i> | Most of the answers are 'yes'. App reduces transportation costs such as driving or calling a taxi |
| <i>How do you think implementing a pharmacy app based on location will ensure a fast and secure ordering process?</i> | When asked about the efficiency of the app, the people said, ordering through app is just a click away and the medicines are delivered in a few minutes, which is faster than going to the pharmacy |
| <i>Do you think implementing a pharmacy app can save time for patients? And how can it reduce the time consumption?</i> | The responses are "yes" because there is no need to wait in the queue, no need to buy from alternate store when the required drug is not available, no need to travel since the availability of medicine can be checked beforehand |

Table 4 Interview results related to branding

| Question | Output |
|--|--|
| <i>How will implementing a pharmacy app, increase the reputation of the pharmacies among the people?</i> | The responses are: Because using chat box patient can contact the pharmacist, ordering by the app will save people's money and time, the app is advanced and faster |
| <i>How would implementing a pharmacy app bring trust to users?</i> | The answers are, 'by providing features to the patient' such as consultation, asking the pharmacist regarding the medication, and uploading the prescription |
| <i>How will implementing a pharmacy app increase the branding of the pharmacies?</i> | People prefer to order their items online without putting much effort, which is why this application receives wide interest as it sends the medicines to the people in just a click, and this will increase the branding of the pharmacies |
| <i>How would implementing a pharmacy app attract the patient to order?</i> | It will be easier for the patient to order his medicines from the app as it won't consume time. The app is easy to use since it is divided into groups and arranged in a manner that it attracts the user |

user begins to type the name of the medication and the app suggests the name of the medication and the name of the pharmacy, as shown in the search page of Fig. 5.

Users who are not sure about the appropriate medicine for their sickness or doubtful of specific details or dosage of the medicines, can chat with the pharmacist

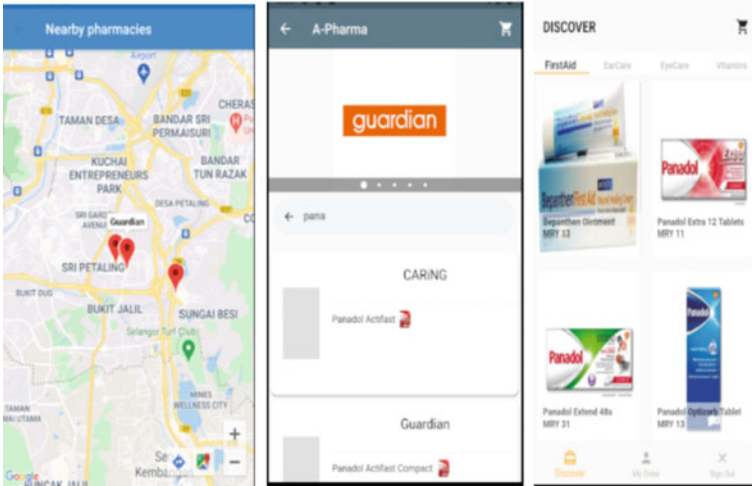


Fig. 5 Location page, search page and list of medicines page

from the selected pharmacy. As shown in Fig. 6, users can view the email of the pharmacist and clicking on the email directs the user to the message page. Moreover, the patients can upload the image of their medicine, prescription, or even their condition and show the pharmacist. The delivery person can also use the chat option to inform the customer when he arrives.

Once the patients have selected their required medicine, their address must be provided. After checking the cost of the medicine, the patient confirms the order that will be saved and sent to the pharmacy as shown in Fig. 7. The pharmacy can see the order details which contain the total price, address, payment method, and



Fig. 6 Chat page and message page

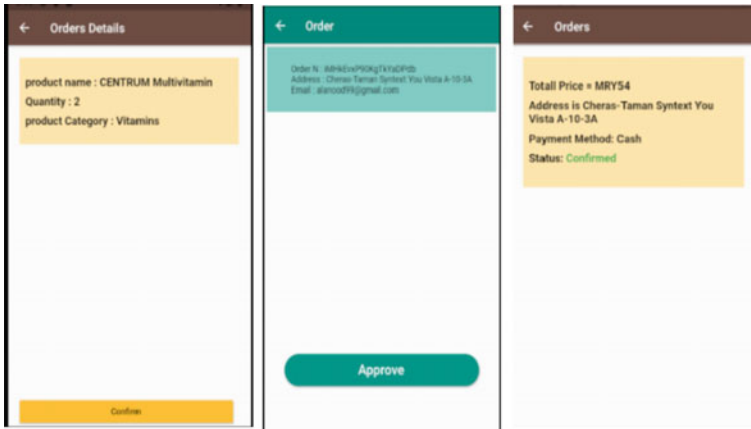


Fig. 7 Order details page

status. When the pharmacy confirms the order, the status is displayed as confirmed in green color. Moreover, when the pharmacy confirms the order, it is displayed on the delivery boy's order page. When the button called "Approve" is selected by the delivery boy, the user can view their order time.

7 Conclusion

Online pharmacy applications are one of the important apps that are mandatory in these days for the satisfaction of the user's health needs. In conclusion, the A-Pharma app that has been developed, provides the online ordering of medicines to be more efficient and cost-effective, and boosts the registered pharmacy image because of the provided features such as consulting the pharmacists, adding various medicines from different pharmacies to the database, delivering the order to the patient's house etc. Pharmacies can register and use this app to make everyone's life easier, especially the patient who regularly visit the pharmacy physically since with a few clicks, the user can check the availability of the medication in all the nearby pharmacies and get their medications without leaving their place which saves their effort, time, and cost.

References

1. T. Le, Z. Andreadakis, A. Kumar, R.G. Román, S. Tollefsen, M. Saville, and S. Mayhew, "The COVID-19 vaccine development landscape", *International Journal of Nature Reviews Drug Discovery*, 19(5), pp.305-306 Apr. 2020.

2. M. R. Laurent and T. J. Vickers, "Seeking Health Information Online: Does Wikipedia Matter?", *Journal of the American Medical Informatics Association.*, 16(4), pp. 471–479, Jul. 2019.
3. N. Ismail, S. Kasim, Y. Yah Jusoh, R. Hassan, and A. Alyani, "Medical appointment application," *Journal of Acta Electron. Malays.*, 1(2) pp. 05–09, Nov. 2017.
4. K. Sharma and R. Sharma, "Online medicines and medical products shopping—a brief study", *International Journal of Management and Applied Science*, 2(2), pp.112–113, Feb.2016.
5. Z. Shaikh, D. P. Doshi, D. N. Gandhi, and D. M. Thakkar, "E-Healthcare Android Application based On Cloud Computing," *Int. J. Recent Innov. Trends Comput. Commun.*, 6(4), pp. pp. 307–310, Apr. 2018.
6. M Loke, PS JosephNg, AS Shibghatullah & HC Eaw (2020), Jomdrone: Data mining financial sense in the property agency, *IEEE Symposium on Industrial Electronics and Application*, Malaysia, 15
7. J. Kaur and Dr. K. S. Mann, "AI-based HealthCare Platform for Real-Time, Predictive and Prescriptive Analytics using Reactive Programming," *International Conference on Computer and Electrical Engineering.*, 933(10), pp. 1–12, Jan. 2018.
8. G.Alex, B.Varghese, J.G.Jose, and A.M.Abraham, "A Modern Health Care System using IOT and Android," *Journal for Research*,2(01).pp.98–101, March.2016.
9. M. A. Mohammed and A. S. K. Bright, "Mobile-Based Medical Health Application - Medi-Chat App", *International journal of scientific & technology research*, 6,(05), pp. 70–76, MAY 2017.
10. S. Surve and S. Perveen, "Descriptive analysis of quality of information of drugs on android-based Indian online pharmacy applications: a preliminary analysis", *Asian Journal of Pharmaceutical and Clinical Research*, 13(5),pp94-96, May 2020.
11. T. E. Kwadwo, K. Kusi, and P. A. Rutherford, "Design and Implementation of Hospital Reservation System on Android", *International Journal of Computer Science and Information Security (IJCSIS)*,17(10), pp. 31–37, October 2019.
12. M. Davies, M. Collings, W. Fletcher, and H. Mujtaba, "Pharmacy Apps: a new frontier on the digital landscape?," *Journal of Pharmacy Pract.*, 12(3), pp. 1-11, Jul. 2016.
13. S. Sawarkar, H. Shaikh, V. Kshirsagar, P. Machale, and R. Shahabade, "An Android Application for Electronic Health Record System", *International Research Journal of Engineering and Technology (IRJET)*,06(03), pp. 2569- 2570, Mar 2019.
14. B. Hawashin and A.M. Mansour, "Healthcare Monitoring System using Near-field communication", *International Journal of systems applications, engineering & development*,12(4), PP 39–44, 2018.
15. Md. A. Majid, M.J.Alam and MD.N.Mustafa, "Smart Doctors Appointment and Prescription System" *IOSR Journal of Computer Engineering (IOSR-JCE)*,19(6), PP 85–91, 2017
16. JosephNg, P. S., & Eaw, H. (2021). Making Financial Sense from EaaS for MSE during Economic Uncertainty. *Future of Information and Communication Conference* (pp. 976–989). Vancouver, Canada: Springer Advances in Intelligent Systems and Computing.
17. JosephNg, P. S. (2021). Economic Turbulence and EaaS Grid Computing. *Int. J. of Business Forecasting and Marketing Intelligence*, 7(1), 33–52
18. JosephNg, P. S., & Eaw, H. C. (2022). Still Technology Acceptance Model? Reborn with Exostructure as a Service. *International Journal of Business Information Systems*, forthcoming.
19. JosephNg, P. S. (2019). EaaS Infrastructure Disruptor for MSE. *International Journal of Business Information System*, 30(3), 373-385.
20. JosephNg, P. S. (2018). EaaS Optimization: Available yet hidden information technology infrastructure inside the medium-size enterprise. *Technological Forecasting and Social Change*, 132(July), 165 - 173.
21. JosephNg, P. S., & Kang, C. M. (2016). Beyond Barebones Cloud Infrastructure Services: Stumbling Competitiveness During Economic Turbulence. *Journal of Science & Technology*, 24(1), 101-121.
22. JosephNg Poh Soon, Kang Chon Moy, Ahmad Kamil Mahmood, Wong See Wan, Phan Koo Yuen, Saw Seow Hui, Lim Jit Theam (2016), EaaS: Available yet Hidden Infrastructure inside MSE, 5th International Conference on Network, Communication, and Computing, Kyoto, Japan, ACM International Conference Proceeding Series, 17–20.

23. PS, J. N., Kang, C. M., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. (2016). Exostructure Services for Infrastructure Resources Optimization. *Journal of Telecommunication, Electronic and Computer Engineering*, 8(4), 65-69.
24. JNP Soon, WS Wan, PK Yuean, LE Heng & LJ Theam (2015). Barebone Cloud IaaS: Revitalisation Disruptive Technology, *International Journal of Business Information Systems*, 18(1), 107–126.
25. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. H. (2014). IaaS cloud optimisation during economic turbulence for Malaysia small and medium enterprises. *International Journal of Business Information Systems*, 16(2), 196–208.
26. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W. Phan, K. Y., & Lim, E. H. (2013). Battles in volatile information and communication technology landscape: The Malaysia small and medium-size enterprise case, *International Journal of Business Information Systems*, 16(2), 196–208.

A Convolutional Neural Network Incorporated Long Short-Term Memory with Autoencoder for Covid-19 Intensity Levels Detection



J. Deepika and J. Akilandeswari

Abstract The intelligent machine assisted diagnostics for the reliable and rapid identification of coronavirus disease (COVID-19) has become a most demanded approach to prevent the novel coronavirus spread during the pandemic and to relieve the strain on the healthcare system. The need for speedy diagnosis necessitates deep learning approaches for predicting the patient's health, and disease severity assessment using Lung Ultrasound (LUS) is the secure, radiation-free, adaptable, and advantageous choice in prediction and detection of novel coronavirus. The suggested model is the convolutional neural network deep layers integrated with recurrent neural networks autoencoder block used to indicate disease intensity ranges from lung ultrasound (LUS) images. The evaluation metric for the proposed model used is the five-fold cross-validation approach. Experimental results for novel proposed model depict through confusion matrix and accuracy-validation curve compared between the traditional convolutional neural network model and united training model consisting of convolutional neural network and long short-term memory (LSTM) based convex probe and linear probe evident that accuracy rate has increased in predicting the intensity levels than the former model. The memory unit incorporated in the training model enables to store, modify, update the temporal features including both of training data and testing data. Convolutional Neural Network (CNN) incorporates an autoencoder block to provide a robust, noise-free classification model in predicting intensity levels.

Keywords Covid -19 · Deep learning · Convolutional neural network (CNN) · Long short-term memory (LSTM) · Lung ultrasound (LUS)

1 Introduction

COVID-19 has been announced as a pandemic which is triggered by the SARS-CoV-2 virus by the World Health Organization (WHO). This pandemic has effects on all

J. Deepika (✉) · J. Akilandeswari
Department of Information Technology, Sona College of Technology, Salem, India
e-mail: jdeepikait@gmail.com

fields such as politics, education, the economy, social issues, the environment, climate change, and has provided huge concerns about health care to tackle the toughest scenario with a disease which is first of its kind [2]. Despite significant technological advances and a large coverage of vaccination campaign being implemented in several countries, the catastrophic condition prevails unabated.

The accurate finding is critical in delivering immediate health assistance to affected patients, as well as assisting government authorities in preventing the spread of the disease [5]. From the outbreak of coronavirus (COVID-19), there has been constant research undergoing to develop a faster method of detecting COVID-19 than the reverse transcription–polymerase chain reaction (RT-PCR) test. Reverse transcription–polymerase chain reaction (RT-PCR) test relies heavily on testing environment and sample collecting processes, and it has a restricted testing capacity [7].

Rapid testing is the most efficient technique for limiting the spread of this easily transmitted disease, prompting researchers to look for a quick diagnostic procedure. Most investigations for COVID-19 detection performed with these following strategy of imaging techniques such as computed tomography (CT) scan, X-Ray, or lung ultrasound (LUS). The wearable devices enable with medical sensors to obtain signals [9] is becoming more common to be employed. The CT scan is one of the radiological imaging procedures that provides lungs 3D image which can detect lung cancer and COVID-19 symptoms at several phases of the illness [10]. The CT showing of images is expensive and reveals harmful rays [12]. X-Ray is appealing approach because of its versatility, elasticity, lesser price, and a significantly faster method [13].

However, the disease's features and respiratory associations at several phases are readily not observable in X-ray pictures because of compression. The originality of these images is less-pixels visibility with different attributes [15]. Lung ultrasound (LUS) images remits clarity with original view avoiding health risks along higher effectiveness. A comparison of COVID-19 detection performance utilising known methods such as CT scan, X-ray, and ultrasound datasets revealed highest prediction accuracy through LUS images than that achieved with a CT scan or an MRI [16] X-ray. A traditional classification network of CNN in addition to the LSTM is proposed to classify the input images into different intensity levels ranging from level 0 to level 3.

2 Related Work

The segmentation methodology, in the form of frames, and video data system of grading primarily used to segregate LUS imaging problems into chosen groups. The LUS utilized to diagnose respiration related disorders, and it provides a more accurate diagnosis of pneumonia than other methods. X-Rays were used to get the results, which were inspected visually by specialists in the domains [17]. Following the start of the COVID-19 epidemic, [18] the RT-PCR test results has impact on the testing environment and test sample collecting procedure [14]. Fast testing process is the

effective method for controlling transmitted disease [11]. The consistent techniques as radiological imaging techniques to identify COVID-19 using the input retrieved from computed tomography (CT) scan, X-Ray, or lung ultrasound (LUS).

Lung ultrasound (LUS) is efficient visual inspection. In comparison to LUS related scrutiny, remaining procedures necessitate need for equipment, that results in contamination and disease spread [11]. Lung ultrasound (LUS) images are utilized to detect disease patterns and different ranges of disease. The intelligent learning techniques used for disease prediction in the resource-constrained context.

Based on the severity prediction, patients may be classified as mild, moderate, extreme, or critically ill, depending on the stages of corona disease development [19]. If a patient's symptoms decrease, lung ultrasound images show a gradual clearance of results. At various stages of the illness, a rise in pattern and consolidation can be observed, which is referred to as ground-glass opacity (GGO).

Screening and identifying Covid-19 in CT images from actual cases of pneumonia is a significant challenge since they have radiographic characteristics. CT images for thin slice acquisition will contain several slices that must be examined manually [20]. False negative results are possible, especially in the initial stages of the disease.

As a result, computerised investigation, and analysis methods are required for distinguishing COVID-19 from commonly acquired pneumonia that can be achieved from examining lung ultrasound images [9]. An effective deep layered architecture for autonomous illness intensity forecast that performs well in a resource-constrained and hospital-independent context.

In the COVID-19 diagnosing procedure, where clinicians are still in high demand and lack of resources scenario this hospital assistance environment plays a vital role in detection of disease intensity. The LSTM classifier provides higher accuracy in detection of sentiments in entire phrase of sentences [1]. Deep Convolutional Neural Network (DCNN) can remit optimized accuracy by using the hyperparameters involved in this model [2]. The principal component analysis (PCA) method is used to extract features using a statistical learning method. In the deep learning domain, it's also known as the KL transformation method. Furthermore, this is an unsupervised method of learning image process [4]. The classification strategy of CNN has optimized accuracy in detection of user emotions [6]. The LSTM classifier also used in stock market price prediction along with other deep learning architectures such as multi-layer perceptron and support vector machines with higher grade of performance evident from experimental results. When the industry's pattern is used over the others, the trial results show that pattern recognition may be done well [3]. The LSTM classifier model related to normalization technique utilized in bitcoin price prediction [5]. The convolution structure uses deep convolution to solve inverse problems, with a specific focus on biological data [8].

3 Deep Learning Model

Deep convolutional neural network (DCNN) process categorization of lung ultrasound (LUS) images based on severity scores ranging from 0 to 3 which is possible to achieve high accuracy with integration of Long Short-Term Memory (LSTM), component efficiently perform processing of temporal aspects from LUS images. The spatial features presence in the frames of lung ultrasound (LUS) video, there can be also temporal features which are efficiently processed by the addition Long Short-Term Memory (LSTM) network.

CNN with deep layers includes an autoencoder block which connect the several parallel convolutional networks at various points which enables the classification process easily to be performed by finding edge dominant features. The four-score illness severity prediction performance improves significantly after including Long Short-Term Memory (LSTM) with cross validation convolutional layers.

3.1 Dataset

The COVID-19 Lung Ultrasound Database (ICLUS-DB) [25] used, has 6 LUS videos, which can be accessed subsequently physical examination is completed. The account request has been approved. The linear and convex probes gathered after subsequent the data acquisition procedure taken. The remaining 21 were obtained with a linear probe. Individual frame scores were obtained from [12], which included all 58 frames. There are 39 videos with a convex probe and 21 videos with a linear probe.

The 14,312 frames predicted using a four-intensity range detection method, which is completed in a timely way. Here, 0 represents the healthiest instance, and 3 represents the worst-case scenario [18], which has been shown to be a beneficial tool for adult respiratory distress syndrome (ARDS) patients.

But this as well can be a prospective replacement for COVID-19's assessment of the intensity levels [20]. The images of convex and linear probes have different patterns and dimensions, training and testing are done separately on them.

Table 1 represents the dataset used in the suggested network which must cope with an unbalanced set of data across several groupings, including bigger proportion

Table 1 Dataset

| Category of data | Videos count | Frames count | | | |
|------------------|--------------|-------------------|-------------------|-------------------|-------------------|
| | | Intensity level 0 | Intensity level 1 | Intensity level 2 | Intensity level 3 |
| Linear | 21 | 1986 | 405 | 1465 | 758 |
| Convex | 39 | 1954 | 1604 | 4281 | 1458 |

with intensity level 0 and level 2 examined in linear images and consisting of convex images.

3.2 Pre-Processing

The initial goal to abstract frames for performing illness intensity level detection. The features undesirable to detect through ultrasound be stable as white patches throughout the ultrasound video’s continual motion. By removing the undesired bits, the target region may be simply recovered and then supplied to categorization model.

3.3 Proposed Novel Convolutional Neural Network

To categorise the provided images of lungs obtained through ultrasound into ranges of intensity levels, categorization training model created on an altered CNN connected to second phase of LSTM. Figure 1 depicts the proposed scheme’s simplified pipeline. The suggested CNN stage in Fig. 1 is shown in the top portion, contains autoencoder block that passes dynamic attributes for distinct convolutional divisions derived through LUS images, as well as a block of separable convolutional branches that extracts edge-dominant features. Later, as indicated in the bottom part of Fig. 1, the pictures are successively transmission done through LSTM for every LUS frame in video to conduct decision making detections.

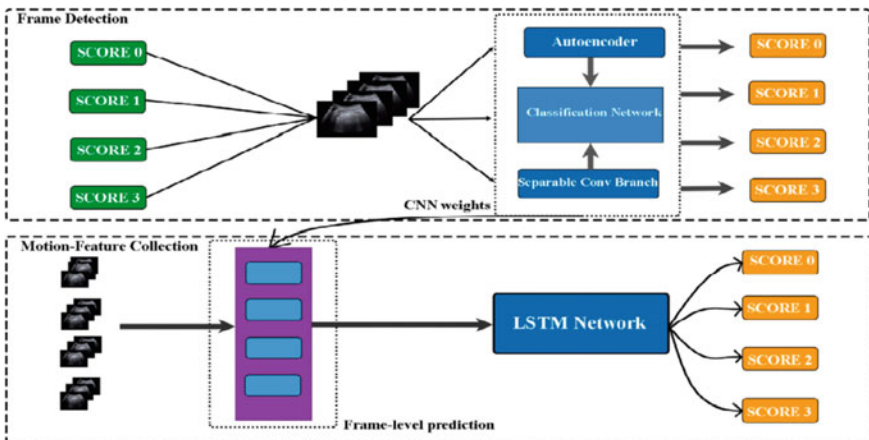


Fig. 1 Pipeline architecture of proposed deep learning model

3.4 Proposed Algorithm

The deep convolutional neural network (DCNN) integrates autoencoder block that feed complicated attributes into separable layers of convolution derived and block of separable convolutional layers that extracts edge-dominant features from the input LUS images.

For each of LUS video, sequential transferring of images to LSTM, which processes the end results detections considering the weight fixed by deep layered CNN with integration of LSTM unit blocks. To decrease noise, the images are initially put through an autoencoder block, which extracts strong attributes that helps in detection of vital qualities segregates between intensity level ranges. The autoencoder used for lessening other outliers and filter vigorous features which helps to identify the finest features between different intensity levels classification.

The individual LUS frames are fed into the deep layered fivefold cross validation CNN model, which is then combined with the LSTM blocks to provide one of the four severity scores as a final prediction of the illness. An autoencoder is used to refine the input images by directing the hidden layers to find the important characteristics in any input images. These features are serially encoded using up sampling before the reconstructed image is displayed. The autoencoder mapping input,

$$x [0, 1]^d \quad (1)$$

into a representation of the input,

$$y [0, 1]^{d'} \quad (2)$$

using the function,

$$y_i = s(Wx_i + b) \quad (3)$$

with the mapping occurring,

$$z_i = s(W'y_i + b') \quad (4)$$

The representation is transferred into a restoration of the same shape in the given input. S in Eq. (3) is an abbreviation for a nonlinear function, such as the sigmoid function. The encoder is the first component, followed by the decoder. This model's parameters have been fine-tuned to reduce the average reconstruction error.

The input images are refined using autoencoder [23] directing into the deep layers to find important characteristics. Raw image dimension of $128 * 128 * 3$ taken as input for encoding process and converted to $16 * 16 * 8$ dimensional image. This layer contains image's most important features. These characteristics are serially decoded using upsampling, resulting in a reassembled output dimension of $128 * 128 * 3$ same dimension of input given.

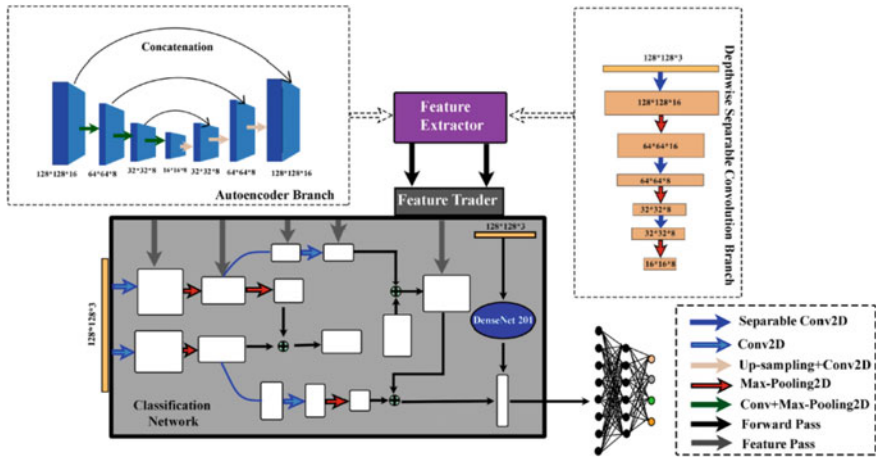


Fig. 2 Deep convolutional neural network (CNN) architecture

The disadvantages with traditional autoencoder is that they eliminate the dominant elements. Concatenation performed between the features which are encoded and the features which are decoded in supervised networks same as in the fully connected network to retain the governing characteristics and maintain the similarity among the input and output. The decoder representation shown in Eq. (5).

$$z_i = s(W'z_i - 1 + b') \oplus y_i, \quad i = 1, 2, 3 \tag{5}$$

where, $z_0 = y_0$.

Notation \oplus in Eq. (5) represents concatenation. The dimension of input image of $128 * 128 * 3$ convoluted into dimension of $128 * 128 * 16$ and result image dimension $128 * 128 * 16$ obtained sent to autoencoder unit. The first three encoding steps in the autoencoder unit depicted in Fig. 2.

The matrix y_0 with dimensions $16 * 16 * 8$ after encoding generated by operations (in centre of diagram), and subsequent three decoding process are performed. The initial step of decoding output (z_1), $s(W' z_0 + b')$ is generated using $z_0 = y_0$ and then concatenated with y_1 . The Encoding component performed at the convolution layer and max-pooling layer.

The convolution process and sampling process are performed for the decoding part in Fig. 2 denoted with arrow heads. y_0 is converted to a $32 * 32 * 8$ dimension image after the convolution and up sampling procedures and subsequently added to encoding branch y_1 performed priorly, belongs to equivalent family as $32 * 32 * 8$ are of equal dimension. This convolution operation, up sampling operation and joining process is in repetition for obtaining the feature image dimension of $128 * 128 * 16$ indicated in Fig. 2. Initial stage of encoded operation, matrix dimension increased

by changing the filter size from 16 to 8. The reformed output of dimension $128 \times 128 \times 3$ is achieved after process of deconvolution.

Convolution and up sampling procedures, y_0 is turned to a $32 \times 32 \times 8$ matrix, added with the encoding process performed branch y_1 resulted dimension of $32 \times 32 \times 8$. This process of convolution, up sampling, and addition performed until a $128 \times 128 \times 16$ feature matrix is created. A loss function reduces the reconstruction error. The cross-entropy loss function utilizes sigmoid/softmax activation function.

The cross-entropy (CE) definition shown in Eq. (6).

$$CE = - \sum_i^N t_i \log f(s)_i \tag{6}$$

N is the count of classes, t is the label, and $f(s)$ is the softmax function shown in Eq. (7).

$$f(s)_i = \exp(s(i)) / \sum_j^N \exp(s(j)) \tag{7}$$

Autoencoder entropy function definition is derived in Eq. (8).

$$L_H(x, z) = - \sum_{k=1}^N [x_k \log(z_k) + (1 - x_k) \log(1 - z_k)] \tag{8}$$

Instead of employing traditional CNN, depth wise separable convolution is used in the CNN-based block. As a replacement for individual kernel of dimension $3 \times 3 \times 3$, three distinct kernels utilized in depth wise separable convolution are employed. The size of each kernel is $3 \times 3 \times 1$. Individual kernel interacts with the others. The layer considered as input has one channel. The input dimension $M \times N \times 3$ are employed. The size of the map produced by convolution is $(M-2) \times (N-2) \times 1$. They are concatenated to achieve $(M-2) \times (N-2) \times S$ dimension. The “S” represents as 3 for value of dimensions. During the 1×1 process of convolution consisting of kernel dimension $1 \times 1 \times 3$ is employed in second stage of the depth wise separable convolution to extend the depth. At the end of K iterations of 1×1 , consisting of dimension $(M-2) \times (N-2) \times K$ is achieved.

Convolutions that can be separable is used to extract low-level characteristics from the input image. Overfitting occurs as the number of parameters increases due to multiple branches. Depth wise separable convolution modifies the parameters.

From various levels of the autoencoder and convolution branches, the dominant features are filtered and sent to the proposed classification training model. The input images are categorized into various paths, each of which combines features from the previous branches. This network’s output compressed, and attributes are added to previous path. For classification of data the connected layers with dimensions 128, 64, and 4 are obtained. Final output images are categorized four different intensity levels.

To reduce the cost, proper optimization is required. As a loss function, categorical cross-entropy utilized minimise damage encountered.

3.5 Long Short-Term Memory Unit Integration

The lung ultrasound (LUS) represented in sequence manner images along with temporal data such as motion data. An optimised technique to input images based on their respective category and predictions performed using the greater number of hidden layers with convolutional neural network (CNN). This technique has the processing capability to filter any data. Memory units are utilized by the long short term memory unit (LSTM) blocks of recurrent neural network model for storing, alter, retrieve detect the temporal features data.

The deep layered convolutional neural network (CNN) is sequence invariant. However recurrent neural networks consider a succession of lung images by encoding temporal features which can improvise the network’s efficacy.

The input $x = (x_1, x_2, \dots, x_T)$, a conventional recurrent neural network processes hidden vector $h = (h_1, h_2, \dots, h_T)$ and the output vector sequence $y = (y_1, y_2, \dots, y_T)$ using Eqs. (9) and (10).

$$h_t = \sigma(W_i h_{xt} + W_{hh} h_{t-1} + b_h) \tag{9}$$

$$y_t = W_{ho} h_t + b_o \tag{10}$$

where, W depicts weight matrices, b for bias, and σ hidden layer activating operation.

The long short term memory (LSTM) block architecture shown in Fig. 3 stores and outputs information using memory cells enables to find complicated temporal correlations and provide dependencies existence among the features. The long short term memory (LSTM) includes following components:

- Input gate
- Forget gate
- Output gate

Input gate includes needed data in cell, forget gate eliminates not required data in memory unit, final component output gate finds only the features from the current cell that are required to identify end prediction. The activation functions used within the cell sigmoid (σ) and hyperbolic tangent (\tanh) functions. By optimising three gates, LSTM regulate weight parameter which includes temporal features data.

The CNN layer’s output is fed into the LSTM at subsequent video features. The output of LSTM of all layers are fed for corresponding layer. The three layers of LSTM, the deep layered CNN’s output is filtered forward in time and upwards. A softmax layer is used to normalise the probability vectors for the intensity prediction levels.

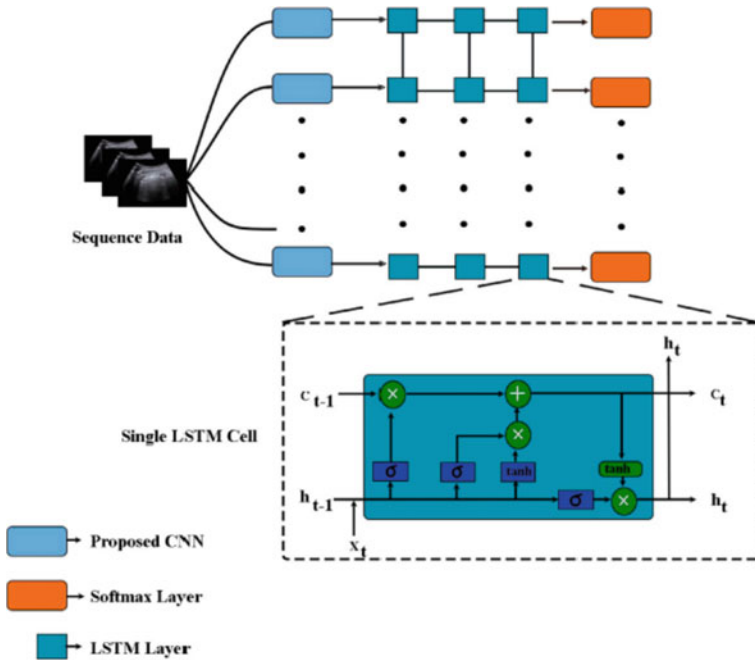


Fig. 3 Long short-term memory (LSTM) integration unit

The one LUS feature frame is provided, then the deep layered CNN weight be influential in predicting the intensity levels. The combined CNN-LSTM weight has capability to detect frames with more precisely on both sequential and temporal features data.

4 Result Analysis

The proposed convolutional neural network (CNN) model integrated along LSTM unit with pre-trained data applied. DenseNet-201 [24] represents convolutional neural network which contains 201 layers deep that enables to be less efficient without LSTM unit proven from the experimental results.

4.1 Training—Testing Technique

The convex probe and linear probe divided as two different data groups consisting of training data and another as testing data. The fivefold cross validation technique implemented with given dataset. Linear probing data, frames from four films (20%)

are maintained in the test set and the 16 (80%) videos are used in the training set for each of the cross-validation steps.

The convex probe data, 8 videos (20%) for testing dataset and 30 videos (80%) are used for training dataset. For both CNN and LSTM, learning rate of 3, a batch size of 64, and 120 epochs are utilized. Several dropout layers are implemented at the end of deep convolutional layers as a result decreases overfitting. The accuracy curve and Loss validation curve is depicted for one of the cross-validation training phases.

A successful match at the training stage is indicated by the lesser gap between the training and testing lines as evident in the Figs. 4 and 5. To address the issue, augmenting process of data is applied to the data used for training with angle of

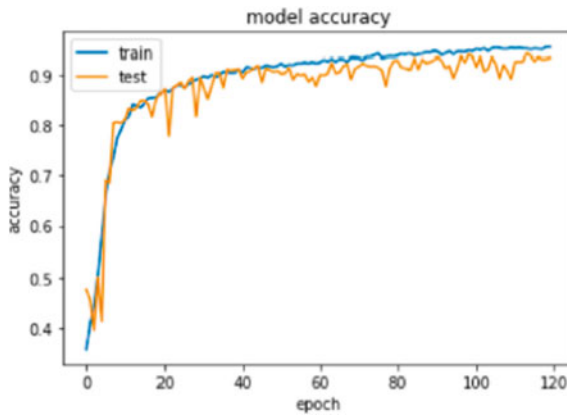


Fig. 4 Accuracy validation curve

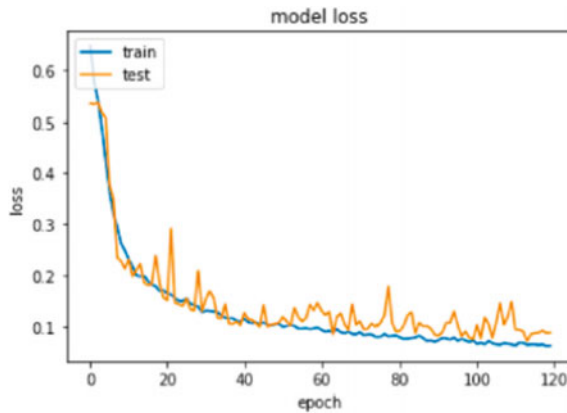


Fig. 5 Loss validation curve

rotating ($0^\circ \pm 360^\circ$), both in position horizontally vertically with shifting angle ($0\% \pm 20\%$) and angle of expanding ($0\% \pm 20\%$) resulting in frame quantity in output.

4.2 Performance Evaluation of Proposed Algorithm Based on Classification

The proposed model's performance is evaluated against three different categories: (1) DenseNet-201 design only, (2) CNN model lacking LSTM unit (3) Proposed CNN model incorporated LSTM autoencoder unit. The evaluation metrics are accuracy, sensitivity, specificity, and F1 score. DenseNet-201 model is applied to the dataset with already fixed weight for previous training of ImageNet [21] to categorize LUS images categories that represent the four intensity levels. DenseNet201 has the process gone through 5 cross validation rounds, with 85 percent of data being trained and the remaining 15 percent being tested on unseen test data. In this situation, the overall accuracy obtained in linear probe data is 50.5 percent and 51.5 percent accuracy achieved in convex probe.

After that, instead of the typical DenseNet-201, traditional CNN architecture used with dataset. CNN is executed and all evaluation parameters improve considerably. The proposed hybrid network, which combines CNN and LSTM blocks, is then used to produce the optimal outcome. Table 2 summarises the incremental improvement in outcomes following the implementation of the suggested network, with error and 95 percent confidence level. Overall performance achieved combining the proposed CNN and the LSTM autoencoder unit with 21% higher in accuracy over DenseNet-201 architecture and 12% higher from traditional CNN without LSTM unit.

Table 2 Performance comparison based on different training models

| Category of data | Training model | Performance evaluation parameters | | | |
|------------------|-----------------------|-----------------------------------|-------------------|-------------------|-------------------|
| | | Accuracy | Sensitivity | Specificity | F1 Score |
| Linear | Dense net 201 | 0.565 ± 0.009 | 0.565 ± 0.009 | 0.868 ± 0.018 | 0.47 ± 0.110 |
| | CNN | 0.600 ± 0.021 | 0.600 ± 0.021 | 0.918 ± 0.023 | 0.712 ± 0.143 |
| | Proposed CNN and LSTM | 0.991 ± 0.078 | 0.991 ± 0.078 | 0.951 ± 0.024 | 0.796 ± 0.027 |
| Convex | Dense net 201 | 0.525 ± 0.049 | 0.525 ± 0.049 | 0.682 ± 0.072 | 0.525 ± 0.046 |
| | CNN | 0.650 ± 0.060 | 0.650 ± 0.060 | 0.726 ± 0.087 | 0.546 ± 0.015 |
| | Proposed CNN and LSTM | 0.787 ± 0.082 | 0.787 ± 0.082 | 0.718 ± 0.120 | 0.696 ± 0.024 |

Table 3 Proposed CNN + LSTM model 5- fold cross validation performance results

| Category of data | Performance evaluation parameters | 5- fold cross validation | | | | | Median value |
|------------------|-----------------------------------|--------------------------|--------------|--------------|--------------|--------------|---------------|
| | | K value as 1 | K value as 2 | K value as 3 | K value as 4 | K value as 5 | |
| Linear | Accuracy | 0.880 | 0.824 | 0.697 | 0.891 | 0.891 | 0.891 ± 0.048 |
| | Sensitivity | 0.880 | 0.824 | 0.697 | 0.891 | 0.891 | 0.891 ± 0.048 |
| | Specificity | 0.927 | 0.912 | 0.744 | 0.823 | 0.823 | 0.911 ± 0.024 |
| | F1 Score | 0.960 | 0.839 | 0.661 | 0.760 | 0.770 | 0.796 ± 0.067 |
| Convex | Accuracy | 0.729 | 0.678 | 0.632 | 0.686 | 0.710 | 0.667 ± 0.022 |
| | Sensitivity | 0.729 | 0.678 | 0.632 | 0.686 | 0.710 | 0.667 ± 0.022 |
| | Specificity | 0.600 | 0.920 | 0.927 | 0.858 | 0.769 | 0.778 ± 0.150 |
| | F1 Score | 0.760 | 0.628 | 0.632 | 0.645 | 0.685 | 0.626 ± 0.024 |

Suggested model examined with fivefold cross validation method. Table 3 shows the detailed results for CNN + LSTM autoencoder unit, including each of fivefold cross-validation steps. At a 95% confidence level, the median values displayed. Each validation stage performance is persistent with an accuracy of linear probe data 79.1% and convex probe data 67.7%.

The images obtained with the linear probe were better compared to convex. For achieving higher quality of images, transducers that are linear utilized [22], resulting in improved detection presentation obtained. Although all types of probes can detect lung consolidations, the category of linear has highest optimizing effectiveness.

The visualisation representation of the consistent enhancement proven by implementing the deep layered CNN model with LSTM unit in two LUS recordings from convex probe and linear probe recorded. The matrices show that the proposed integration method improves the results significantly which are evident in Figs. 6, 7, 8 and 9.

The proposed CNN incorporated LSTM model has capacity of reliably recognising majority of intensity level 2, that are previously mis predicted with the result of intensity level 0 and intensity level 1 utilizing CNN training model alone, in linear

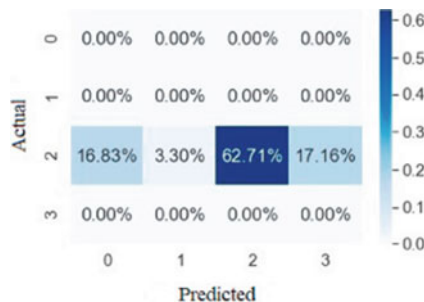


Fig. 6 Confusion matrix—traditional CNN (Convex)



Fig. 7 Confusion matrix—traditional CNN (Linear)



Fig. 8 Confusion matrix—proposed CNN + LSTM (Convex)



Fig. 9 Confusion matrix—Proposed CNN + LSTM (Linear)

probe data shown in Figs. 6 and 7. In the CNN incorporated LSTM network, the count of false negatives in intensity level 3 similarly decreased to some extent. Similarly, the result count of false negatives has decreased to major extent with implementation of proposed CNN integrated LSTM autoencoder block shown in Fig. 8 and Fig. 9.

Figure 10 illustrates both images are covid-19 intensity level 2, both images expected to be of intensity level 1 and intensity level 3. In a healthy lung, ultrasound imaging produces A-lines, that are lines horizontally positioned parallelly to pleural

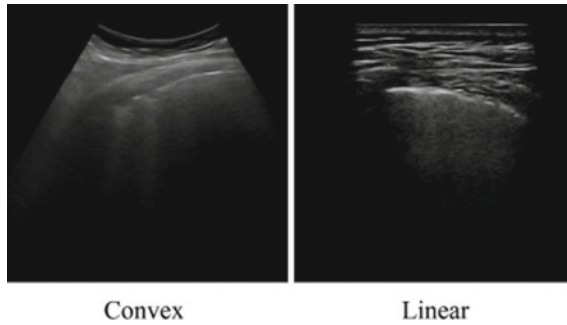


Fig. 10 Lung ultrasound (LUS) images

line. B-lines are positioned vertically that represent a variety of lung pathologies [24]. These vertical abnormalities damage pleural lines, which perfectly continuously placed in normal healthy lung but represent increasingly opaque due to these artefacts.

The proposed deep layered convolutional layered network incorporated with long short-term memory unit predicts the intensity level as 2 with higher precision also with consuming less interval of time because its capable of predicting tiny variations in the images.

5 Conclusion

The proposed frame-based illness severity prediction with the deep convolutional neural network layers and recurrent neural networks predicts intensity levels with ranging from 0 to 3 from the input of lung ultrasound frames. The deep convolutional neural network model with LSTM unit can process the intensity ranges of the lung ultrasound (LUS) images with higher rate of prediction in comparison with traditional convolutional neural network model without LSTM unit.

The proposed deep CNN model required on average less time to make an ultrasound movie with a duration of 100 milli seconds along with a suitable machine configuration and the test for diagnosing the disease should take less than 6 ms to complete. As a result, it can be used in conjunction with point-of-care equipment such as ultrasonography.

In a real-time scenario, machines can forecast a patient's status of disease prediction and severity. The hybrid model will consider both temporal and spatial information, and it will be able to predict ranges of severity with high rate of precision. The proposed deep layers of CNN model capable of predicting the score for a specific frame. This process of identifying the patient's immediate status could be of tremendous use to the medical fraternity in the current pandemic situation. Furthermore, the future enhancement of the proposed work could be to design the architecture even more with deep layers segmentation to accomplish a optimized clinical product.

References

1. Chen, Joy Iong-Zong, and Kong-Long Lai. "Deep Convolution Neural Network Model for Credit Card Fraud Detection and Alert." *Journal of Artificial Intelligence* 3, no. 02 (2021): 101–112.
2. Tripathi, Milan. "Sentiment Analysis of Nepali COVID19 Tweets Using NB, SVM AND LSTM." *Journal of Artificial Intelligence* 3, no. 03 (2021): 151-168.
3. Kottursamy, Kottilingam. "A review on finding efficient approach to detect customer emotion analysis using deep learning analysis." *Journal of Trends in Computer Science and Smart Technology* (2021) pp95–113.
4. Anand, C. "Comparison of Stock Price Prediction Models using Pre-trained Neural Networks." *Journal of Ubiquitous Computing and Communication Technologies (UCCT)* 3, no. 02 (2021): 122-134.
5. Andi, Hari Krishnan. "An Accurate Bitcoin Price Prediction using logistic regression with LSTM Machine Learning model." *Journal of Soft Computing Paradigm* 3, no. 3 (2021): 205-217
6. Jacob, I. Jeena, and P. Ebby Darney. "Design of Deep Learning Algorithm for IoT Application by Image based Recognition." *Journal of ISMAC* 3, no. 03 (2021): 276–290.
7. Vijayakumar, T. "Posed Inverse Problem Rectification Using Novel Deep Convolutional Neural Network." *Journal of Innovative Image Processing (JIIP)* 2, no. 03 (2020): 121-127
8. Dhaya, R. "Deep net model for detection of covid-19 using radiographs based on roc analysis." *Journal of Innovative Image Processing (JIIP)* 2, no. 03 (2020): 135-140.
9. Shayan Hassantabar, Novati Stefano, Vishweshwar Ghanakota, Alessandra Ferrari, Gregory N. Nicola, Raffaele Bruno, Ignazio R. Marino, Kenza Hamidouche, Niraj K. Jha, Covid Deep: SARS-CoV-2/COVID-19 Test Based on Wearable Medical Sensors and Efficient Neural Networks, 2020.
10. Ai Tao, Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen, Wenzhi Lv, Tao Qian, Ziyong Sun, Liming Xia, Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: a Report of 1014 Cases, *Radiology*, 2020, p. 200642.
11. Michael Chung, Bernheim Adam, Xueyan Mei, Ning Zhang, Mingqian Huang, Xianjun Zeng, Jiufa Cui, Wenjian Xu, Yang Yang, Zahi A. Fayad, et al., CT imaging features of 2019 Novel Coronavirus (2019-nCoV), *Radiology* 295 (1) (2020) 202–207.
12. S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli, E. Peschiera, R. Trevisan, G. Maschietto, E. Torri, R. Inchingolo, A. Smargiassi, G. Soldati, P. Rota, A. Passerini, R.J.G. van Sloun, E. Ricci, L. Demi, Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound, *IEEE Trans. Med. Imag.* 39 (8) (2020) 2676–2687.
13. Tanvir Mahmud, Md Awsafur Rahman, Shaikh Anowarul Fattah, CovXNet: a multidilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization, *Comput. Biol. Med.* 122 (2020) 103869
14. Shayan Hassantabar, Mohsen Ahmadi, Abbas Sharifi, Diagnosis, and detection of infected tissue of COVID-19 patients based on lung X-ray image using convolutional neural network approaches, *Chaos, Solitons & Fractals* 140 (2020) 110170
15. Li Fan, Dong Li, Huadan Xue, Longjiang Zhang, Zaiyi Liu, Bing Zhang, Lina Zhang, Wenjie Yang, Baojun Xie, Xiaoyi Duan, et al., Progress and prospect on imaging diagnosis of COVID-19, *Chinese Journal of Academic Radiology* (2020) 1–10.
16. M.J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, N. Shukla, COVID-19 detection through transfer learning using multimodal imaging data, *IEEE Access* 8 (2020) 149808–149824.
17. Yogendra Amatya, Jordan Rupp, Frances M. Russell, Jason Saunders, Brian Bales, Darlene R. House, Diagnostic use of lung ultrasound compared to chest radiograph for suspected pneumonia in a resource-limited setting, *Int. J. Emerg. Med.* 11 (1) (2018).

18. Gino Soldati, Andrea Smargiassi, Riccardo Inchingolo, Danilo Buonsenso, Tiziano Perrone, Domenica Federica Briganti, Stefano Perlini, Elena Torri, Alberto Mariani, Elisa Eleonora Mossolani, Francesco Tursi, Federico Mento,Libertario Demi, Is there a role for lung ultrasound during the COVID-19 pandemic? *J. Ultrasound Med.* 39 (7) (2020) 1459–1462.
19. Andrea Smargiassi, Gino Soldati, Elena Torri, Federico Mento, Domenico Milardi, Paola Del Giacomo, Giuseppe De Matteis, Maria Livia Burzo, Anna Rita Larici, Maurizio Pompili, Libertario Demi, and Riccardo Inchingolo, Lung ultrasound for COVID-19 patchy pneumonia,” *J. Ultrasound Med.*
20. Danilo Buonsenso, Davide Pata, Antonio Chiaretti, COVID-19 outbreak: less stethoscope, more ultrasound, *The Lancet Respiratory Medicine* 8 (5) (2020) e27
21. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems 25*, in: F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (Eds.), Curran Associates, Inc., 2012, pp. 1097–1105.
22. R. Ketelaars, E. Gülpinar, T. Roes, M. Kuut, G.J. van Geffen, which ultrasound transducer type is best for diagnosing pneumothorax? *Crit. Ultrasound J.* 10 (1) (2018) 27.
23. Gao Huang, Zhuang Liu, Laurens van der Maaten, Q. Kilian, Weinberger, densely connected convolutional networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.
24. Daniel A. Lichtenstein, Gilbert A. Mezi'ere, Jean-François Lagoueyte, Philippe Biderman, Ivan Goldstein, Agn'és Gepner, A-lines and B-lines: lung ultrasound as a bedside tool for predicting pulmonary artery occlusion pressure in the critically ill, *Chest* 136 (4) (2009) 1014–1020.
25. Italian COVID-19 Lung Ultrasound Data Base, 2020.

Survey on Attendance System Using Face Recognition



D. Pradeep, A. Bhuvaneswari, M. Nandhini, A. Roshini Begum,
and N. Swetha

Abstract In this framework, face is distinguished and perceived while the video is real time, to recognize the individual amongst difficulties. This recommends comparing people on either pictures or sequential recordings. For video based face acknowledgment, it is tough to accomplish the comparative levels of general execution. There are a few drawbacks in video based face acknowledgment, when face identification depends on the verification with existing face pictures. The drawbacks are indicated by the following: At first, CCTV cameras get pictures which are generally of low norm. Second, picture assurance is generally decreased for video frameworks. Third, face picture varieties, which incorporate light, appearing, stance, obstruction, and development, are more serious in video real time. Therefore this review concentrates on different face position and learning calculations to break down the proficiency in recognition framework.

Keywords Video face recognition · Face detection learning algorithm · Vector construction · Deep learning algorithm

1 Introduction

Face recognition is a way of figuring out or verifying the identification of a person by their faces. Face recognition properties may be used to find individuals in images, motion images, or in real-time. Law enforcement authorities use cellular devices to cognize people at the check points. But face popularity records are liable to blunders that can involve public for mischiefs they have not committed. As per study, the facial recognition applications are at times terrible at spotting mixed ethnic such as African-Americans, and disguised youth, where frequently misidentifying or failing to recognize, disparately impacts the positive group [1–3]. In a powerful sense, multi-view face recognition is the most effective method, which means the scenario in which a number of cameras simultaneously capture the scene and an algorithm

D. Pradeep (✉) · A. Bhuvaneswari · M. Nandhini · A. Roshini Begum · N. Swetha
Department of Computer Science and Engineering, M. Kumarasamy College of Engineering,
639113 Karur, Tamilnadu, India
e-mail: pradeepd.cse@mkce.ac.in

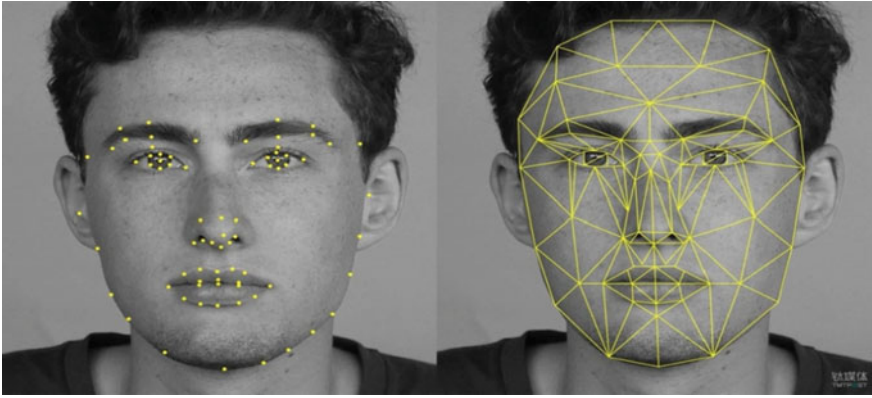


Fig. 1 Face features detection layout

collectively uses the captured images/motion pictures. However, it is time consuming to distinguish faces in a variety of postures. This ambiguity has no bearing on the appeal of photographs; in terms of pose variations, a collection of shots concurrently serious about several cameras and people delighted by a unique digital camera at various perspective angles, both are the same. The two scenarios, however, diverge when it comes to video records. When a multi-advanced camera permits the need of various viewpoint data without warning, the danger of getting similar realities with one camera is extremely high. In non-agreeable standing bundles, such varieties, alongside assessment, become fundamental [4–6]. An assortment of multi-view movies got by synchronized cameras with different view films and a solitary video assortment where the trouble of changing puts a solitary view video for clearness has been emphasized. With the benefit of camera organizations, multi-view reconnaissance movies have arisen to be the most common [7, 8]. At present, numerous different point of view movies face acknowledgment procedures make many single viewpoint films. At the point when several countenances are given for confirmation, they show up inside the assortment for adjusting the face part's exposure in a solitary photo to the equivalent posture and brightening of the other picture. This methodology can even need the stances and brightening circumstances to be imagined for both the face pix [9, 10]. Figure 1 shows the face location format where bounding box is formed around the face to fetch facial features.

2 Related Work

Ayazoglu et al. [11] suggested the reality that, under slight constraints, the 2D trajectories of the spot in the photo planes of every camera are compelled to explore in the identical small areas. This statement enabled to find a linear version that describes all the existing 2D estimations at a given period. In flip, this design was used in

the situation of a changed particle filter out for predicting destiny goal places. If the goal is hindered to a numerous camera, the lost calculations predicted the use of the records that they have to locate each within the subarea spanned by earlier calculations and fulfill epi-polar conditions. Thus, by explicating both varying and geometrical conditions, the technique can actively manage extensive closure, without the demand for acting three-D remodelling, scaled cameras or constraints on sensor division. The accomplishment of the method was demonstrated using numerous hard samples including goals that drastically exchange look and movement fashions while hindered to some cameras. The limitation of the work was the requirement of huge number of cameras and datasets.

Baltieri et al. [12] carried out two dimensional models and, lately, three dimensional models of the human body which are non-articulated. Contrasting to movement capture or movement evaluation methods, the models were not required to be extremely specific but speedy. Nevertheless, model-based localization afforded extra localized and consultant explanations and permitted a valid differentiation of matching frame elements. Issues because of blockages and segmentation mistakes can also be reduced. In this method, a breakthrough by presenting a new original 3-D method to re-identity based on articulated frame fashions has been accomplished. A 3-D version was accepted to match look describers to bones. This hue profoundness as well as bone waves framed using that Microsoft Kinect sensor while the OpenNi libraries are explicated in the form of information. The bone is similarly delicate. The usage of a mastering way allowed to provoke a “bone” set. The acquired signature was exactly related to the actual body structure, accordingly it also let in a characteristic based description, which can be useful in some packages. In addition, a found-out metric is followed which simultaneously acts as a characteristic choice and frame part weighting.

Baltieri et al. [13] provided the design of a whole device for human re-identification, based totally at the mapping of look descriptors to three-D frame fashions known as SARC3D. The adoption of 3-D frame fashions is pretty unique to determine, instead of other computer vision topics, consisting of a sample movement seize and position evaluation. The demanding situations linked with 3-D models depended on the want to obtain accurate human detection, separation and judgement of the 3-D orientation for valid model-to-picture alignment. Global functions which are universal functions were smooth to calculate and the alignment techniques were not required. Moreover, they do not care on the errors of separation or identification. However, the dearth of nearby capabilities ends in deceptive comparisons. The acceptance of 2D models moderately reduces the issue, permitting segment-based comparisons. The upper and lower body parts are compared in a correct way, which increased the clarity of the signature. However, troubles specially associated with specific orientation might also arise.

Barbosa et al. [14] demonstrated the presence of individuals by their clothes, with the reason that altering garments between camera increases or decreases the usefulness of the picture. The RGB-D records offer substantially extra records and is the reason why there has been huge development in this situation, but, alternatively, the recent RGB-D sensors cannot function at the corresponding gap as normal

surveillance cameras; therefore, those specialized in RGB does continue to be an important undertaking. In this paper, a re-identification framework based entirely on the convolutional neural network, which has the aim of addressing the problems stated was presented. The framework is famous for several fantastic traits. Earliest, a schema of a network was not as hard as a complex setup in general. It was based at the inception structure and used as a function extractor. Quantitative experiments sell the preference, giving top scores in reputation phrases. The limitation here was that it took more response time.

Bedagkar-Gala et al. [15] studied more than one individual re-identity across cameras in the context of huge place surveillance and treated matching as a venture trouble based totally on a spatiotemporal appearance version. Biometrics like face or character can be used for comparison however probably was tough to restore because of the resolution or body charge conditions of cameras. For re-identification, appearance models were used. For addressing false comparisons, spatio-temporal affairs were utilized in between cameras. In analysis, each re-ID within the environment having supervising throughout two cameras had accessible set mapping issue, wherein the gallery emerged every year, the inquiry set actively varied for each digital FOV, and every probe inside a hard and fast weren't always a subdivision of gallery.

Yan et al. [16] proposed a singular lively pattern choice technique 'Energetic gaining knowledge' for image type using net pix. The earlier analysis had proven that go-media modeling of variety of media classes was used in evaluating multimedia content. The net photographs were always connected with wealthy textual explanations like captions. When text facts like these aren't present in testing pix, the effectiveness of the textual content characteristics for mastering sturdy separators were displayed, permitting better engaged studying accomplishment of picture category. The recent supervised mastering pattern, specifically mastering the usage of powerful records (LUPI), may be useful for clearing up this issue. In a LUPI situation, along with important functions, there's additionally powerful data present within the training system. Powerful facts were utilized in training and not present in testing. 5 ways to merge the uncertainty rate of these two dividers were given. To confirm the selected samples were consistent, advantage of the range computation was taken, such that the selected samples are very small just like each variant. A purposeful objective function to increase the uncertainty of cross-media and to reduce the uniformity of information was created. Then uncertainty and variability in choosing school pattern was measured. Thus, establishing the exchange-off attributes between the 2 styles of computations was ignored.

Yang et al. [17] applied a brand-new feature choice set of rules, which leveraged the knowledge from associated couple of obligations to improve the performance of the characteristic choice. In the test, the upcoming thoughts had been identified: Sharing records among related duties was useful for supervised studying. However, if a couple of obligations aren't correlated, the overall performance isn't always necessarily progressed. Compared to single assignment mastering, the convenience of multitask gaining knowledge was generally extra seen while having some training samples handy for each mission. When the wide variety of fine schooling information was maximized, the intra-project understanding was enough for schooling,

and as a result accepting inter-task understanding, no longer always assist. The whole achievement refinement of trademark choice changed when unique classifiers were utilized. For instance, KTH dataset was taken, and inspected the accomplishment of different trademark decision calculations when unmistakable separators were utilized. In exceptional, KNN was used as a substitute separator. This is ordinary unmistakable quality precision of the six development characterizations. By matching table and figure, it was proved that when utilizing all qualities for movement ID, the sureness of KNN was declined than that of SVM. Be that as it may, after trademark decision, KNN acquired preferred precision over SVM. One feasible clarification was that SVM can weigh remarkable highlights and in this way the expansion from incorporate assurance is less.

Chang et al. [18] intended to fix the impediments of the current perceive appraisal computations for preposterous solicitation real factors and admonished a compound position k projection estimation to observe bilinear assessment. Consequently, the recommended set of rules and guidelines was named as Compound Rank-alright Projection (CRP) for Bilinear Analysis. Playing out the separation utilizing the customary calculations, the calculation trouble was limited. Contrasted with the old style 2-dimensional direct discriminant assessment techniques, CRP gifts were compromised between the level of opportunity and the evasion of the over-becoming problem. Albeit the customary 2DLDA got right by execution; its redundant streamlining calculation won't consolidate as a result of the uniqueness of the among-class disperse lattice. In an unexpected way, the blend of the system was totally guaranteed. The unwinding of this paper was started as summing up an outline of the old style LDA notwithstanding 2DLDA. However, a new and extraordinary technique which is compound position k projection for bilinear assessment was established. The exploratory results on 5 uncommon datasets were obtained and the finish of the work was discussed. Nevertheless, the model lacks in efficiency.

Luo et al. [19] suggested a structure inclusive of a pair of techniques for both multimedia content examination and recovery. The technique which suggested earliest was, a brand recent transductive rating technique, particularly, rating using Local Regression and Global Alignment (LRGA). Varied from distance-based totally rating strategies, the delivery of the examples inside the complete facts collection was oppressed in LRGA. Comparing with the absolute strategies, simplest the questioning instance was demanded. For every data point, a nearby linear regression version was rent to expect the rating grades of its adjoining points. To assign the surest ranking mark to every fact point, a unified goal characteristic to universally coordinate local linear regression models from every information factor has been suggested. In recovery programs, there was no basic truth in tuning parameters for measurement algorithms like MR. Extensive experiments displayed that these algorithms gain better recovery accomplishment on comparing with the prevailing associated jobs.

Li et al. [20] proposed the calculation named as channel matching neural organization i.e., FPNN for individual re-ID. This profound learning approach had various significant qualities and curiosities contrasted with current works. While training, every one of the central issues were simultaneously improved. Every component

expanded its power while planning with others. All the predominant datasets were quite long, which made it extreme to teach a profound neural local area. Here the dataset had thirteen, 164 pictures of 1300 sixty walkers. The prior datasets just dealt physically cut person on foot photos and depended on ideal identification in appraisal rules. The programmed location in practice sets up enormous wrong arrangement and may seriously affect the achievement of present methods. Here, dataset gave both photographs cut physically and mechanically distinguished control boxes with a nation of the-craftsmanship finder for evaluation.

3 Face Detection Algorithms

Face recognition software use system algorithms to identify distinguishing features of a human's face. These features, such as eye gap or chin structure, are then transformed into a mathematical depiction and matched to data from other face pictures that is in face identification database. A face template is data on a specific face that differs from an image, in that it consists of traits that can be used to recognize one face among many faces in the database.

3.1 Appearance-Based Methods

The methods based on human looks for identification are further divided into two categories: single-shot methods and multiple-shot methods, which are determined by the number of photos utilised to generate the human representation. A single image is utilised to calculate the feature vectors in single-shot methods, whereas a group of images is used to compute the characteristic vectors to address the human in multiple-shot methods. Multiple-shot methods provide more information than single-shot approaches, but they necessitate more complicated algorithms and are more computationally intensive. When no tracking information is available, single-shot approaches are useful [21–25].

3.2 Distance Metric Learning-Based Techniques

Few measurement learning and matching plans were proposed as helpful options for an incredible methodology where man or lady re-distinguishing proof can be obtained. These systems had been intended for concentrating on the quality measurement between look elements of the equivalent individual all through camera sets. These measurements concentrating on procedures are marked into, regulated as limited to unaided, and general difference to local ones. In Person Reid, most positions fall into the scope of regulated global length esteem. Unaided techniques are

particularly mindfulness on work plan and component extraction, and they do not need physically naming preparing tests, notwithstanding, regulated methodologies regularly need the support of physically marked tutoring tests which cause higher general execution. In addition, global measurement perusing methodologies consideration on dominating the vectors of the equivalent polish to be closer on squeezing vectors which are of various illustrations, must be recommended by the Large-Margin Nearest Neighbor measure (LMNN), which has a place with the regulated nearby distance metric concentrating on class, to help improving the traditional KNN type made. Nonetheless, the utilization of the k closest inside-class tests, LMNN has developed to become out to be a time consuming process. As an LMNN adaptation, a moderate improvement into the previously arrived at results through fuse of reject elective for amazing fits known as LMNN-R has been presented. In order to avoid these over fitting issues within LMNN, Information Theoretic Metric Learning (ITML) and Logistic Discriminant Metric Learning (LDML) have been executed as methodology improvement strategies. Considered as a simple and effectively technique, Keep It Simple and Straightforward Metric (KISSME) has been considered taking in length esteem from comparable conditions.

To provide food for the inadequacy, some popular expanding learning methodologies, alongside Self-Organizing Map and Growing Neural Gas, were planned on the reason of neural organization as a portrayal of the unlabeled records, topological shape, and a normal, wherein information can be grouped into explicit directions [26]. To be extra fittingly sound for handling online data, a steady learning strategy has been proposed, alluded to as Self-Organizing Incremental Neural Network (SOINN) which really beats the before referred methods, by allowing the concentration on the significant wide assortment of neural hubs and productively addressing the topological state of plausibility thickness. A remarkable disadvantage of this technique is that, it utilizes Euclidean distance to mark the hole isolating the information and hubs. Notwithstanding, given the commonness intra-style and between class forms, the Mahalanobis metric sounds to be a substitute that most certainly match for dealing with the person Re-ID related issues.

3.3 RGB-Based Approaches

RGBD-based strategies for Person Re-ID has been employed as the RGB look-based man or woman Re-ID assume that, individuals put on the equal garments and take advantage of simplest 2D information, hence a brand new concept based totally on depth has been delivered. In evaluation to RGB statistics, intensity records can preserve more consistent information; though if suffering from apparel alternate and severe brightness due to the fact it is unbiased of hue and keeps extra consistency for a long term. In fact, evoking deep and bone data by using depth cameras (eg. Microsoft Kinect) isn't tough in an indoor environment. Kinect sensors gain intensity value of the distance to the digital camera, of each pixel by using infrared, irrespective of item hue and brightness in tranquil packages. Using depth records, the existence dimension

point cloud and skeleton of someone, may be evoked imparting form and bodily data of his/her frame. Moreover, with depth price of each pixel, people may be more without difficulty segmented from heritage, in order that history have an impact on, that may be largely eliminated. As RGB and depth facts can be received concurrently while Kinect is used. Some Re-ID strategies were evolved to merge depth statistics and RGB look cues, so that one can take out extra peculiar characteristic illustration. These strategies may be divided into two classes, the primary form of method is appearance based methods which integrate look and intensity statistics together, and the second is the proposed a Re-ID technique primarily based on skeletal statistics. Feature descriptors are extracted from person skeletal joints and the final character signature is acquired via concatenating these descriptors [27, 28]. The 2D sort of method depends on geometric capabilities; Re-ID is done through comparing frame shapes as entire point clouds warped to a trendy posture with the explained technique. The technique of anthropometric degree for Re-ID is undertaken and use the 3D location of body joints provided with the aid of skeletal tracker to compute the geometric functions, consisting of limb lengths and ratios.

3.4 HAAR Cascade Classifiers

Digital picture characteristics called Haar-like characteristics are employed in object identification. This algorithm has the OpenCV package for face identification where bounding boxes are formed around the face. This feature increases the accuracy. They are utilized in the first real-time face detector, as well as got their name from their spontaneous resemblance to Haar wavelets. Only high image performance (i.e., RGB pixels per pixel in a photo) has made feature counting more expensive in the past. A HAAR-like characteristic evaluates adjoining rectangular sections in a spotting window at a specified point, adds the pixel strength in every area, and figures out the variation between the total. This distinction is utilized for categorizing image subdivisions. A database with pictures of faces has been considered as an example. This can be a normal perception that in every face, the eye district is hazier than the cheek area. Thus, a simple characteristic of the HAAR to detect face is a pair of neighbor rectangles just above the eye and the cheek area. The location of the rectangles is explained in relation to the acquisition window that enacts as a binding grid on the aim subject i.e., face. The main benefit of a Haar-like characteristics is the computational momentum. HAAR Cascade algorithm generally uses the facial features. To analyze the eye features, circular HAAR features are used.

Rectangular HAAR like features. A common rectangle like HAAR-like characteristic is explained as the variation in the total of resolution of regions within the rectangle, which can be anyplace and measure the genuine capture. This changed list of capabilities is called 2-rectangle feature. Viola explained three rectangle characteristics and Jones explained four rectangle characteristics. The parameters depict some

features of a peculiar region of the picture. Every characteristic type can demonstrate the presence (or vacancy) of some features in the picture, such as edges or modifications in size.

Course classifier. The course divider incorporates a rundown of reaches, wherein each degree incorporates a menu of frail unpracticed people. The gadget identifies objects being referred to, by moving a window over the image. Every level of the categorizer names the particular area characterized through the current area of the window as both decent or awful –which implies that a not set in stone or awful methodology, that the predefined object changed into presently not situated inside the picture. Assuming the marking gives a helpless end yield, then, at that point, the class of this specific area is therefore complete, and the area of the window is moved to the accompanying area [29]. Assuming the naming offers a fine outcome, then, at that point, the area moves to the resulting level of order. The classifier yields a last decision of colossal, when the levels in general, comprehensive of the final remaining one, yield an outcome, declaring that the not set in stone inside the image. A real way that the thing in question is absolutely in the photo and the classifier marks it thusly as an amazing final product. A false way that the naming technique erroneously verifies that the thing is put inside the photograph, even though it isn't. The calculation steps as follows:

$$\text{Feature} = w_1 * \text{RecSum}(r_1) + w_2 * \text{RecSum}(r_2)$$

- Loads can be positive or negative
- Loads are straightforwardly relative to the space
- Determined at each point and scale. It includes a weak classifier.

A weak classifier ($h(x, f, p, \vartheta)$) consists of

- feature (f).
- threshold (ϑ)
- polarity (p), such that

$$h(x, f, p, \vartheta) = \begin{cases} 1 & \text{if } pf(x) < p\vartheta \\ 0 & \text{otherwise} \end{cases}$$

Requirement: Ought to perform better compared with the irregular occurrence.

4 Face Recognition Algorithms

4.1 *Naive Bayes Algorithm*

As the name implies, it suggests that all variables are mutually associated or independent and contribute to the classification. Class restricted independence is the term for this belief. As a result, it's recognized in three ways as the following: Idiot's Bayes or Simple Bayes or Independence Bayes. They project the chance of association with a definite class. When the class variable is given, this divider considers whether the existence (or lack) of a specific characteristic (attribute) of a class is not related to the existence (or vacancy) of any other characteristic, thus identifying the association. Statistical classifiers are Bayesian classifiers. Naïve Bayes calculation is the most durable AI methods which is for determining shower. The Naïve Bayes classifier depends on many policies and regulations among which one is 'Bayes rule' i.e., conditional probability is the likelihood of event or outcome occurrence. It analyzes each attribute individually and assumes that all of them are independent and important [30]. The forecasting of fault-proneness has been extensively utilized by Naïve Bayes classifiers. The benefit of the Naive Bayes Classifier is the demand of a minute quantity of teaching data to gauge the qualities needed for division.

4.2 *Support Vector Machine*

Support Vector Machine (SVM) is an AI procedure that is utilized for both order that is grouping, just as relapse relies on the need. SVMs rely upon deciding the ideal hyper plane for separating a dataset into two divisions. On diverging from more AI strategies, SVM has an undeniable degree of precision. Since they just utilize a small part of preparing focuses, they are more productive. To foresee the extent of precipitation, the SVM calculation is utilized to recognize climate ascribes. SVM's center thought is utilized to isolate twofold information. The methodology that isolates a multiclass issue into various separate paired division occupations is utilized in most of past approaches. Truth be told, these procedures regularly produce indissoluble cases that diminishes the assurance of the segments. Subsequent to acquiring the information and applying the preprocessing strategy, the arrangement or forecast framework can be utilized. The premise of the assessing framework relies upon SVM. Every single example in the showing stage comprises of one "point esteem" and a few "credits". The desire of SVM is to create a plan from the instructing dataset which tracks down the point boundaries of the assessment data. SVM endeavors to observe that hyper plane which expands the edge and diminishes a sum relating to the various wrong division fails at the same time [31, 32]. Nonetheless, the called hyper plane will have the most noteworthy advantage in the midst of the pair of classifications and the breaking point is clarified as an aggregate of the length in the midst of the partitioning hyper plane and the nearby hubs on one or the other piece

of the two classes. Four main ideas can be used to understand SVM classification and, as a result, its predictive ability are:

- The separation hyper plane,
- The SVM with a hard margin
- Kernel function and
- Soft-margin SVM

SVM models were really made to order parts into directly detachable separators.

4.3 Artificial Neural Network

This algorithm is used for face identification. A neural organization is a calculative technique dependent on tremendous arrangements of neural units freely, demonstrating in the specific manner as the human mind settle issues with gigantic arrangement of natural neurons associated with one another utilizing axons. Each neural unit is associated with numerous others. Each extraordinary neural unit might have a capacity called summation work which is for blending its bits of feedbacks boundaries. Such frameworks learn and train without help from anyone else rather than being unequivocally coded and act in regions, in which the outcome or trademark ID is difficult to express in a traditional framework code. A solitary layered or multi-faceted organization of neurons is shaped when a neuron joins with different neurons by means of association connect. A multi-facet ANN contains 3 layers which incorporate one info layer, one result layer, and essentially archived layers [33]. An organization is prepared until the adjustment of loads in a preparation cycle arrives at the very least worth. Subsequent to preparing, a model is approved by checking whether or not it produces precise result. Diverse organizations are talented in helping information, to remember various synaptic loads present in network. The consequence of a neuron at the result layer is sometimes qualified, utilizing an edge work [34]. Notwithstanding the way that this subject includes manufactured neurons, they are address as neurons. The neurotransmitters between neurons called associations where neurons are documented as edges of a coordinated chart and manufactured neurons are recorded as hubs. This technique of making a model is divided into 5 steps:

- Choosing the input and output data.
- Data normalization (both input and output).
- Standardized information is educated with the guide of neural network which is the neural network learning.
- Checking the model's goodness of fit.
- Comparing the predicted output with the desired output.

A layer of processing elements performs independent calculations on the data you receive and transfers the result to another layer. The following layer might do its own assessments and move the result to another layer. Finally, the result of the network is found by the small cluster of one or more processing components. Every refining

component forms its calculation depended on a calculated total of the inputs. The input layer comes at first whereas the output layer i.e., the result layer comes at the end [35, 36]. The archived layers are located amid the layers indicated above. The processing components are the viewed units that are uniform to neurons working in the brain, and so, they are addressed as cells, neuromines, or fabricated neurons.

5 Conclusion

In this paper, survey on several face detection techniques that is face identification and recognition algorithms for attendance system has been presented. Based on the analysis from the previous papers, it is found that the efficiency is low with high response time since large number of datasets were considered as image. Based on this survey, HAAR Cascade with neural network algorithm can provide improved accuracy rate in face recognition while video is streaming. In future, if this algorithm works efficiently, it can be applied in the following ways. It may assist the Police department to arrest the offenders. If the convict's face is found in the security cameras fixed on roads, an immediate alert may be given to the police control room after which the police officers can track them. This algorithm may also help to detect the missing children, women and even pets.

References

1. A. Bedagkar-Gala and S. K. Shah.: Part-based spatiotemporal model for multi-person re-identification: *Pattern Recognition Letters*, vol. 33(14), pp. 1908 – 1915 (2012).
2. Murugesan, M., Thilagamani, S.: Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network: *Journal of Microprocessors and Microsystems*, vol. 79, November (2020).
3. K. Bernardin and R. Stiefelhamen.: Evaluating multiple object tracking performance: the CLEAR MOT metrics: *EURASIP Journal on Image and Video Processing*, (246309), pp. 1–10 (2008).
4. Thilagamani, S., Nandhakumar, C.: Implementing green revolution for organic plant forming using KNN-classification technique: *International Journal of Advanced Science and Technology*, vol. 29(7S), pp. 1707–1712 (2020).
5. M. Bredereck, X. Jiang, M. Korner, and J. Denzler.: Data association for multi-object Tracking-by-Detection in multicamera networks: *Sixth International Conference on Distributed Smart Cameras (ICDSC)*, pp. 1–6, Oct (2012).
6. Thilagamani, S., Shanti, N.: Gaussian and gabor filter approach for object segmentation: *Journal of Computing and Information Science in Engineering*, vol. 14(2), 021006 (2014).
7. Y. Cai and G. Medioni.: Exploring context information for inter-camera multiple target tracking: *IEEE WinterConference on Applications of Computer Vision (WACV)*, pp.761–768, Mar (2014).
8. Pandiaraja Perumal and Suba S: An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography: *Journal of World Review of Science, Technology and Sustainable Development* , vol. 18(1), pp. 51 – 58 (2022).

9. L. Cao, W. Chen, X. Chen, S. Zheng, and K. Huang.: An equalised global graphical model-based approach for multicamera object tracking. ArXiv: 11502.03532 [cs], Feb(2015).
10. Pandiaraja P., Sharmila S.: Optimal routing path for heterogenous vehicular adhoc network: International Journal of Advanced Science and Technology, vol. 29(7), pp. 1762–1771 (2020).
11. M. Ayazoglu, B. Li, C. Dicle, M. Sznaiar, and O. Camps.: Dynamic subspace-based coordinated multicamera tracking: IEEE International Conference on Computer Vision (ICCV), pp.2462–2469, Nov(2011).
12. D. Baltieri, R. Vezzani, and R. Cucchiara.: Learning articulated body models for people re-identification: In Proceedings of the 21st ACM International Conference on Multimedia, MM '13, pp.557–560, New York, NY, USA, ACM (2013).
13. D. Baltieri, R. Vezzani, and R. Cucchiara.: Mapping appearance descriptors on 3d body models for people reidentification: International Journal of Computer Vision, vol. 111(3), pp.345–364 (2015).
14. I. B. Barbosa, M. Cristani, B. Caputo, A. Rognhaugen, and T. Theoharis.: Looking beyond appearances: Synthetic training data for deep cnns in re-identification, arXiv preprint [arXiv:1701.03153](https://arxiv.org/abs/1701.03153) (2017).
15. A. Bedagkar-Gala and S. Shah.: Multiple person reidentification using part based spatio-temporal color appearance model: In Computer Vision Workshops (ICCVWorkshops) 2011 IEEE International Conference on, pp.1721–1728 (2011).
16. Y. Yan, F. Nie, W. Li, C. Gao, Y. Yang, and D. Xu.: Image classification by cross-media active learning with privileged information: IEEE Transactions on Multimedia, vol. 18 (12), pp. 2494–2502 (2016)
17. Y. Yang, Z. Ma, A. G. Hauptmann, and N. Sebe: Feature selection for multimedia analysis by sharing information among multiple tasks: IEEE Transactions on Multimedia, vol.15(3), pp.661–669 (2013).
18. X. Chang, F. Nie, S. Wang, Y. Yang, X. Zhou, and C. Zhang: Compound rank-k projections for bilinear analysis: IEEE Transactions on Neural Networks and Learning Systems, vol. 27(7), pp.1502–1513 (2016).
19. Y. Yang, F. Nie, D. Xu, J. Luo, Y. Zhuang, and Y. Pan: A multimedia retrieval framework based on semi-supervised ranking and relevance feedback: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34(4), pp.723–742 (2012).
20. W. Li, R. Zhao, T. Xiao, and X. Wang: Deepreid- Deep filter pairing neural network for person re-identification: in Proc. CVPR, pp.152–159 (2014).
21. Pandiaraja P., Aravinthan K., Lakshmi Narayanan R., Kaaviya K.S., Madumithra K : Efficient cloud storage using data partition and time based access control with secure aes encryption technique: International Journal of Advanced Science and Technology, vol. 29(7), pp. 1698 – 1706 (2020).
22. K.-W. Chen, C.-C. Lai, P.-J. Lee, C.-S. Chen, and Y.-P. Hung: Adaptive Learning for Target Tracking and True Linking Discovering Across Multiple Non-Overlapping Cameras: IEEE Transactions on Multimedia, vol. 13(4), pp.625–638, Aug (2011).
23. Rajesh Kanna, P., Santhi, P.: Unified Deep Learning approach for Efficient Intrusion Detection System using Integrated Spatial–Temporal Features: Knowledge-Based Systems, vol. 226 (2021).
24. X. Chen, K. Huang, and T. Tan: Direction-based stochastic matching for pedestrian recognition in non-overlapping cameras: 18th IEEE International Conference on ImageProcessing (ICIP), pp.2065–2068, Sept (2011).
25. Santhi, P., Mahalakshmi, G.: Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction: International Journal of Engineering and Advanced Technology, vol. 8(4), pp. 839–846 (2019).
26. S. Liao, Y. Hu, X. Zhu, and S. Z. Li: Person re-identification by local maximal occurrence representation and metric learning: in Proc. CVPR, pp. 2197–2206 (2015).
27. Deepa, K., Thilagamani, S.: Segmentation techniques for overlapped latent fingerprint matching: International Journal of Innovative Technology and Exploring Engineering, vol. 8(12), pp. 1849–1852 (2019).

28. L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian: Scalable person re-identification: A benchmark, in Proc. ICCV, pp.1116– 1124 (2015).
29. D. Pradeep, C. Sundar, QAOC: Noval query analysis and ontology-based clustering for data management in Hadoop, vol.108, pp. 849–860 (2020).
30. R. Logeswaran, P. Aarthi, M. Dineshkumar, G. Lakshitha, R.Vikram: Portable Charger for Handheld Devices Using Radio Frequency: International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8(6), pp. 837–839 (2019).
31. Vijayakumar, T.: Synthesis of Palm Print in Feature Fusion Techniques for Multimodal Biometric Recognition System Online Signature: Journal of Innovative Image Processing (JIIP) ,vol. 3(02), pp. 131–143.cnn (2015).
32. Jacob, I. Jeena, and P. EbbyDarney: Design of Deep Learning Algorithm for IoT Application by Image based Recognition. Journal of ISMAC, vol. 3(03), pp. 276–290 (2021).
33. Bharadwaj, Gunjan, and Pooja Pathak.: Visual Attendance Recording System for Classroom Using Facial Features: In Innovative Data Communication Technologies and Application, pp. 583–589. Springer, Singapore, (2021).
34. Jadhav, Sharad R., Bhushan U. Joshi, and Aakash K. Jadhav: Attendance System Using Face Recognition for Academic Education: In Computer Networks and Inventive Communication Technologies, pp. 431–436. Springer, Singapore, (2021).
35. Raj, Suraj, and SaikatBasu: Attendance Automation Using Computer Vision and Biometrics-Based Authentication-A Review: Computer Networks and Inventive Communication Technologies : pp.757–767 (2021).
36. Niharika, M., & Sree, B. K.: IoT Based Attendance Management System Using Google Assistant: In International conference on Computer Networks, Big data and IoT , pp. 21-31. Springer, Cham, December (2019).

A Spectral-Spatial Classification of Hyperspectral Image Using Domain Transform Interpolated Convolution Filter



M. Preethi, C. Velayutham, and S. Arumugaperumal

Abstract Hyperspectral imaging technology is a very useful technique in remote sensing and hyperspectral sensors receive more than hundred spectral bands from the electromagnetic spectrum. Hyperspectral image (HSI) is produced by each pixel of vector which is used to identify the unique materials of objects. Due to becoming large spectral bands, the classification of image is a more arduous task. To confront a problem, filtering techniques have become a hot research topic. The filtering techniques are employed for improving the HSI classification. This work discusses spatial feature extraction by using Domain Transform Interpolated Convolution Filter (DTICF). There are three steps for the proposed techniques: First, the spatial features are obtained by using DTICF. Second, those obtained spatial in the first processing and spectral bands of HSI are provided to design 3-Dimensional-Convolutional Neural Network (3D-CNN) with and Bi-directional-Long Short Term Memory (Bi-LSTM) framework. In the third step, the obtained probabilistic classification map of 3D-CNN- Bi-LSTM is smoothed by Markov Random Field (MRF) efficiently. Based on the experimental results, two different hyperspectral images prove that DTICF-3D-CNN-Bi-LSTM approach is more important and provides good classification results compared to other classification approaches.

Keywords Bi directional-long short term memory · Deep learning · Domain transform interpolated convolution filter

M. Preethi (✉)

Department of Computer Science, S.T. Hindu College, Affiliated to Manonmaniam Sundaranar University, Nagercoil, TamilNadu, India

e-mail: preethiindira12@gmail.com

C. Velayutham

Department of Computer Science, Aditanar College, Tiruchendur, TamilNadu, India

S. Arumugaperumal

Department of Computer Science, S.T. Hindu College, Nagercoil, TamilNadu, India

1 Introduction

Compared to RGB image, HSI is a most noteworthy technique in remote sensing [30]. HSI accommodate a high number of spectral information and large number of land cover classes [18]. Due to large no of land cover classes and spectral information, HSI classification is a major challenging task. HSI is utilized in many applications such as medical imaging, microscopy or endoscopy, precision agriculture, mineralogy, and food inspection. Last few years, machine learning and deep learning is employed for HSI classification [8, 10]. There are types of feature extraction such as spectral based features and spectral-spatial features based extraction. Spectral features are first extracted by some feature extraction methods [4] such as Principal Component Analysis (PCA) [6], Independent Component Analysis (ICA) [19], and Linear Discriminant Analysis (LCA) [1]. Then, the obtained features are applied to learn the classifier [3]. In spectral-spatial-based methods, texture features [15] and structure features [7] are extracted and combined by utilizing composite kernels [12]. However, the obtained features are hand-crafted.

Recently [3, 23, 24], deep learning methods are utilized for image processing [35] such as image classification [33], image segmentation [21] and object detection [34]. Compared to machine learning methods, CNN [14] has been employed for obtaining the features of spectral and spatial for HSI classification. Konstantinos Makantasis et al. [17] employed deep learning methods for the HSI classification method which capture the features using CNN and this work utilized a Multi-Layer Perceptron for a classification task. Shaohui Mei et al. [29] proposed a novel five-layer CNN for HSI classification such as batch normalization, dropout, Parametric Rectified Linear Unit (PReLU) activation function. In this proposed framework, spatial context and spectral information are elegantly integrated that is utilized to obtain the features. Haokui Zhang et al. [9] introduced an end-to-end 3-D lightweight CNN. This work contains a deeper network structure, fewer parameters, and lower computation cost, resulting for better classification performance. Qin Xu et al. [27] proposed multiscale convolution from 3D-CNN for obtaining the pair of spectral-spatial features. This proposed method helps to reduce the spatial redundancy. Radhesyam Vaddi et al. [28] introduced new HSI classification techniques which are based on data normalization and CNN. In this work, HSI data are normalized and then Probabilistic Principal Component Analysis (PPCA) and Gabor filtering are used. PPCA is used for obtaining the features which are used to reduce the computational time. Jia et al. [13] developed a 3-dimensional (3-D) Gabor-wavelet for hyperspectral classification which helps to predict the features via 3-D. Kang et al. [16] acquired the spectral features by Gabor filtering to form the fused features for Gabor filtering-based deep network (GFDN). In particular, CLSTM is used for obtaining the spectral features of HSIs which improve the extraction of spatial features using convolutional operators [31]. Dhaya et al. [5] utilized a kalman filter for image retrieval. This filter is assimilated with adaptive feature extraction. This work analysis the feature vector data analysis for transparent that provides higher retrieval rate. Subarna shkya et al. [32] generated

the crop water stress index (CWSI). This work is utilized for crop water management. This proposed work is estimated the amount of water content in a vineyard using high resolution thermal imaging. Nirmal et al. [25] proposed an open set (OS) domain adaptation and generative adversarial network (GAN) for HSI classification. This domain adaptation is utilized for feature space that is a subset of feature space from source domain. This proposed work classified the trained classes and unknown labeled classes.

The paper is designed in the following ways. Proposed methodology is illustrated in Sect. 2. The technical description is expounded in Sect. 3. The experimental results for the proposed method are analyzed in Sect. 4. Section 5 is presented with the conclusion.

2 Proposed Methodology

The main contributions of the proposed work are stated below: We have proposed a new DTICF-3D-CNN-Bi-LSTM HSI classification framework. The proposed method is divided into three processing. In the First processing, the spatial features are excerpted by DTICF. The excerpted features and patch-wise input image with spectral information is integrated. During the second phase, the integrated features are provided in the 3D-CNN framework. The extracted deep features are again fed in the Bi-LSTM network. In the third step, the obtained probabilistic classification map of 3D-CNN- Bi-LSTM is smoothed by Markov Random Field (MRF) efficiently. The proposed DTICF-3D-CNN-Bi-LSTM-MRF shows good classification accuracy with low computational time. In this field, the proposed DTICF-3D-CNN-Bi-LSTM of HSI classification is debated in Fig. 1. First, the HSI data is modified to RGB image. The modified RGB images are transformed to gray-level images. The spatial-based features are extracted by using DTICF.

The spatial features and patch-wise input data of spectral bands are integrated. The integrated features are provided to newly developed 3D-CNN architecture for classification. In this section, 3D-CNN with Bi-LSTM based classification method is explained and discusses how to train the network with deep learned features from HSI. 3D-CNN configuration substantially consists of three blocks of Convolution (c_1, c_2, c_3) and ReLU (R_1, R_2, R_3) layers. The filters used in three sets are $k_1 = 20$, $k_2 = 20$ and $k_3 = 35$ respectively. The extracted features of the HSI $(x_1, y_1, 1)$ are given as input to 3D_CNN. In 3D-CNN, the first convolutional layer c_1 with k_1 filters data becomes (x_1, y_1, k_1) and (x_2, y_2, k_1) . The second convolutional layer c_2 with k_2 filters the data becomes (x_2, y_2, k_2) . and (x_3, y_3, k_2) . Finally we obtain the data (x_3, y_3, k_3) by using third set of Convolution and ReLU layers. The ReLU features are given into Bi-LSTM network to extract features. In the last stage of Bi-LSTM model, we take the input of the ReLU features obtained by 3D-CNN. The final data is categorized by applying a soft-max function. The probability map of 3D-CNN with Bi-LSTM can also be improved by MRF efficiently.

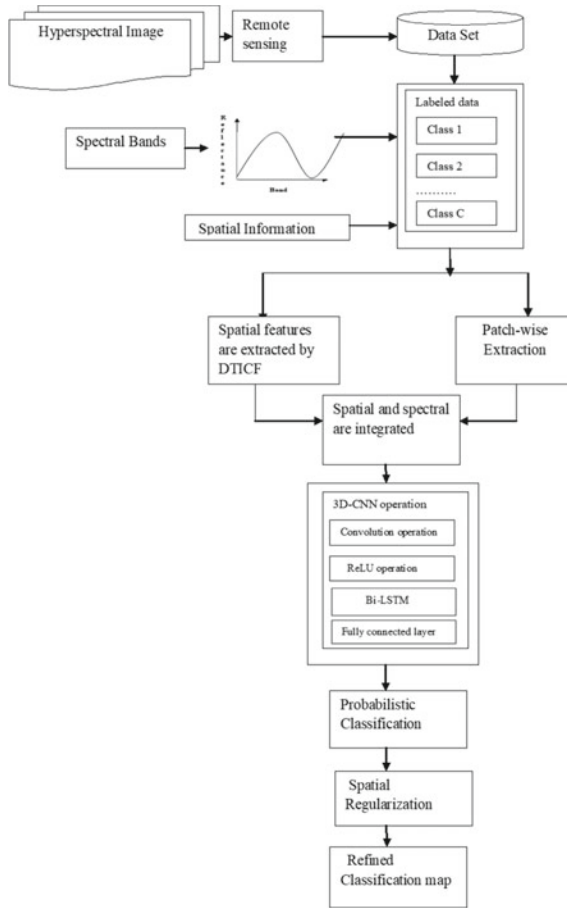


Fig. 1 Structure of the HSI classification

2.1 Domain Transform Interpolated Convolution Filter (DTICF)

Oliveria [26] proposed a DTICF for image filtering which is an edge preserving filtering. This image filtering is used for improving the spatial features. It mainly concentrates the spatial correlation. The spatial correlation contains the reflection intensity between pixels. It is calculated in the following ways:

$$Z_i(u) = \int_{\Omega_w} P_w Q(h(u), x) dx \quad i = 1, 2, \dots, n; u \in \Omega_w \quad (1)$$

In Eq. (1) [27], Filtering P_w is evaluated by the consecutive convolution, where Q is a normalized box kernel and r is the filter radius.

$$Q(h(u), x) = \frac{1}{2r} \delta\{|g(u) - x| \leq r\} \tag{2}$$

$$(E) = \begin{cases} 1 & \text{Eistrue} \\ 0 & \text{other} \end{cases} \tag{3}$$

Substituting Eqs. (5), [27] and (6) into (4):

$$F_i(u) = \frac{1}{2r} \int_{g(u)-r}^{g(u)+r} P_w(x) dx \tag{4}$$

$$G(u) = \int_0^u 1 + \frac{\sigma_s}{\sigma_r} \sum_{l=1}^c |I_k'(x)| dx \tag{5}$$

$$\sigma_r = \sqrt{3\sigma_j} \tag{6}$$

$$\sigma_{jn} = \sigma_s \sqrt{3} \frac{2^{M-n}}{\sqrt{4^M - 1}} \tag{7}$$

2.2 CNN Operation

CNN is used to extract the features from image. It is employed to obtain deep learned features. It's able to handle large size of data [5]. Compared to machine learning, deep learning method provides good performance. In deep learning, CNN is employed for image processing. In this work, 3D-CNN is utilized.

2.2.1 3D Convolution

3D-CNN is employed to extract the features of spatial and spectral information concurrently. In 3D-CNN, 3D kernels are used to compute the weighted sum of pixels as:

$$C_{pqr} = f \left(\sum_{i,j,k} w_{ijk} a_{(p+i)(q+j)(r+k)} + b \right) \tag{8}$$

2.3 Bi-Long Short Term Memory (BI-LSTM)

Bi-LSTM network LSTM was established by Hochreiter and Schmidhuber [11]. This network structure overcomes the problems of RNN [24].

In Bi-LSTM, there are three memory gates such as input, forget and output. The high frequency Intrinsic Mode Function (IMF) is C_{tsh} that is given as an input and h_{i-1} is the output. CS_{t-1} is the input of cell state determined by forget gate f_t using a sigmoid function. It is written in Eq. 10 [11]:

$$f_t = (w_f[h_{i-1}, C_{tsh}] + b_f) \quad (9)$$

Input gate i_t is used to determine the values that are to be updated to CS_t as in Eq. 9 [11]:

$$i_t = (w_i[h_{i-1}, C_{tsh}] + b_i) \quad (10)$$

The output gate o_t are equated in Eq. 10 [10]:

$$CS_t = f_t \odot CS_{t-1} \oplus i_t \odot CS_{t-1} \quad (11)$$

Consequently, the output of LSTM memory cell is written in Eq. 11[11]:

$$h_N = o_t \odot CS_t \quad (12)$$

$$y_d = \text{softmax}(W_o h_N \cdot + b_o) \quad (13)$$

3 The Proposed Method of DTICF-3D-CNN-Bi-LSTM-MRF

In this field, we discuss the proposed method DTICF-3D-CNN-BiLSTM-MRF. The network structure of 3D-CNN-BiLSTM is discussed in Fig. 2.

3.1 The Spatial Feature Extraction by DTICF

First, HSI image is modified to RGB image. Then, RGB image with spatial features and spectral bands from original HSI Data is applied in DTICF. For dataset $D = \{d_1, d_2, \dots, d_s\}$, we utilized the DTICF to capture the features.

$$[g_1, g_2, \dots, g_s] = \text{RGB}(D) \quad (14)$$

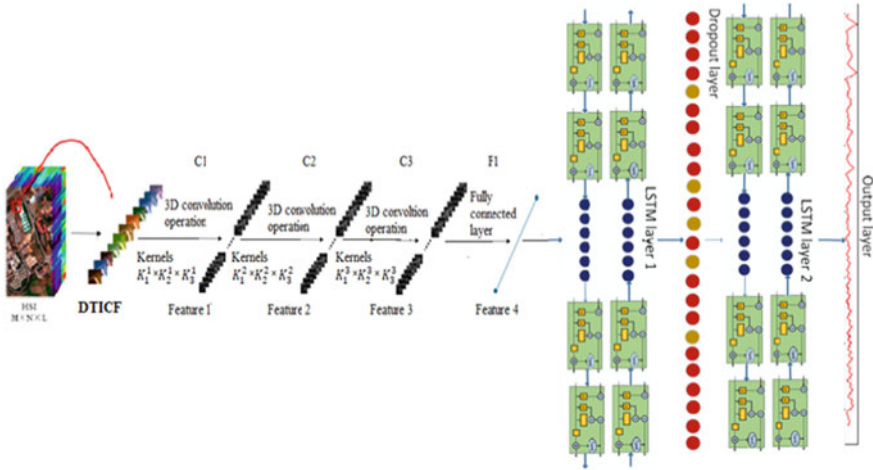


Fig. 2 Network structure of DTICF- 3D-CNN-Bi-LSTM

3.2 Classifying HSI by 3D-CNN-BiLSTM

In this section, the excerpted features are provided to 3D-CNN network which is utilized for obtaining the pair features of spectral-spatial using convolution, pooling and ReLU layers. Finally, ReLU is employed for extracting the features. Bi-LSTM is applied in the sequence features from 3D-CNN. Then, we adopt a soft-max function for classifying the entire feature vector. The obtained probabilistic classification map of 3D-CNN-Bi-LSTM is smoothed by MRF efficiently.

The proposed method analyze the output by using log-likelihood $\log P((y_i | \tilde{y}_i)$ can be given by [11]

$$\hat{y} = \arg \max_{y_i \in K^n} \left\{ \sum_{i=1}^n \sum_{k=1}^K 1\{y_i = k\} \log Q_{ik} + \mu \sum_{i=1}^n \sum_{j \in N(i)} \delta(y_i - y_j) \right\} \tag{15}$$

Algorithm 1 examines the proposed work of our classification algorithm as follows:

Algorithm1. The proposed DTICF-3D-CNN-BiLSTM-MRF Hyperspectral Image Classification

Input: HSI data $\text{img} \in R^{h \times w \times d}$, Ground truth $G \in R^{h \times w \times d}$

Output: Features $\{\{y_1, y_2, \dots, y_n\}\}$

1. Set the convolution size;
 2. Compute the padding size;
 3. Extract the spatial features by using Eq.
 4. // Sigma_s Filter spatial standard deviation;
 5. // Sigma_r Filter range standard deviation;
 6. // num_iteration Number of iteration to perform;
 7. //Joint_image optional image for joint filtering;
 8. function F = IC(img, sigma_s, sigma_r, num_iterations, joint_image)
 9. I = double(img);
 - 10.
 11. if ~exist('num_iterations', 'var')
 12. num_iterations = 3;
 13. end
 14. if exist('joint_image', 'var') && ~isempty(joint_image)
 15. J = double(joint_image);
 16. if (size(I,1) ~= size(J,1)) || (size(I,2) ~= size(J,2))
 17. error('Input and joint images must have equal width and height.');
 18. end
 19. else
 20. J = I;
 21. end
 22. [h w num_joint_channels] = size(J);
 23. //Compute the domain transform;
 24. // Estimate horizontal and vertical partial derivatives using finite differences;
 25. dIcdx = diff(J, 1, 2);
 26. dIcdy = diff(J, 1, 1);
 27. dIdx = zeros(h,w);
 28. dIdy = zeros(h,w);
 29. // Compute the l1-norm distance of neighbor pixels;
 30. for c = 1:num_joint_channels
 31. dIdx(:,2:end) = dIdx(:,2:end) + abs(dIcdx(:, :, c));
 32. dIdy(2:end,:) = dIdy(2:end,:) + abs(dIcdy(:, :, c));
 33. end
 34. for i = 0:num_iterations - 1
 35. // Compute the sigma value for this iteration;
 36. // Compute the radius of the box filter with the desired varianc;
 37. End
 38. End function
 39. While D: 1 → X
 40. Compute the patch for training data;
 41. Compute another patch for testing;
 42.
$$D_{l(k)}^{(k)} = \{(x_1, y_1), \dots, (x_{l(k)}, y_{l(k)})\};$$
 43. Generate feature maps operation using convolution operation Eq. (8) ;
 44. $f(x) = \max(0, x)$; % ReLU operation;
 45. Sequence input by $f(x)$ given Bi-LSTM network using Eq. (9);
 46. Perform forward and backward sequence to generate the feature;
 47. Compute $a = \frac{\exp(o)}{\sum_k \exp(o_k)}$; % Soft-max activation function;
 48. end while
 49. Compute the classification label \hat{y} using Eq. (15).
-

4 Experiments

In this section, there are two hyperspectral dataset used in the proposed DTICF-3D-CNN-BiLSTM-MRF such as Indian pines data and Pavia University data. There are three measurements used for validation such as [1]: Overall accuracy (OA), Average Accuracy (AA), and Statistically kappa measure (k). Indian pines data set contains the spatial dimension size of height and width as 145×145 and 220 spectral reflectance bands. It contains 16 classes. Pavia University dataset contains the spatial dimension size of height and width 610×340 and 103 spectral bands. There are 9 land cover classes in this dataset. The above two data sets are displayed in Fig. 3 and 4.

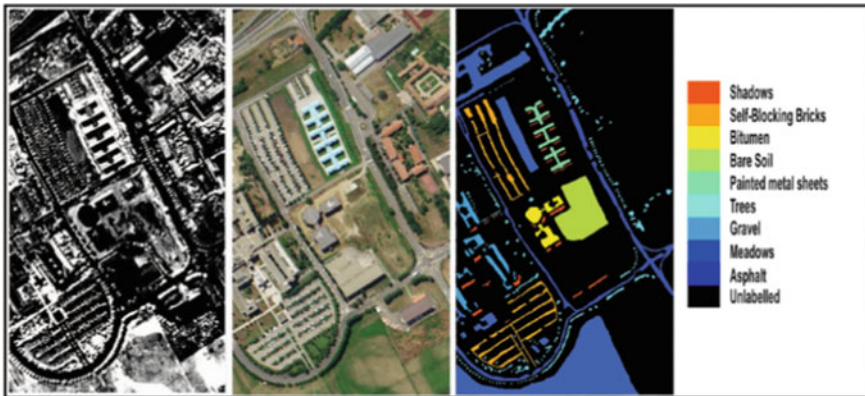


Fig. 3 Original image and Ground Truth of Pavia University dataset

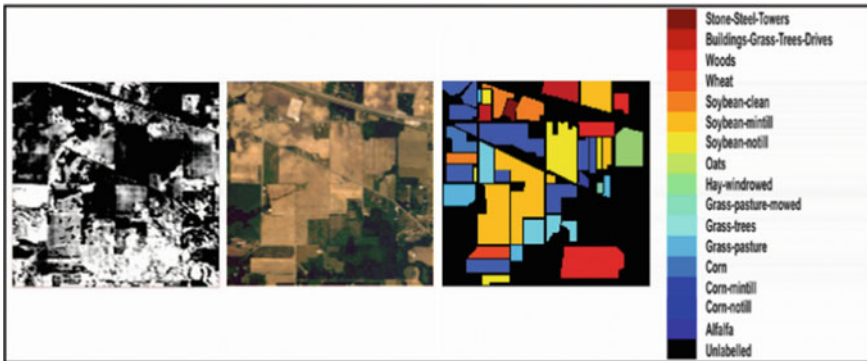


Fig. 4 Original image and Ground Truth of Indian Pines dataset

4.1 The Proposed DTICF-3D-CNN-Bi-LSTM-MRF HSI Classification on Indian Pines Data and Pavia University Data

To experiment the data, we choose 50% of samples for training and another samples for testing. In Table 1, overall accuracy of the propose method DTICF-3D-CNN-Bi-LSTM for HSI classification is discussed. The propose method DTICF-3D-CNN-Bi-LSTM compared to other deep learning based HSI classification method such as CNN, CNN-MRF, 3D-CNN and 3D-CNN-MRF respectively [23]. Compared to CNN, CNN-MRF, 3D-CNN, and 3D-CNN-MRF, our proposed methods achieve good classification accuracy. Classification results are validated in terms of OA. In Table III, our DTICF-3D-CNN-BiLSTM-MRF approach provides the best result compared with other methods.

In Fig. 5 of Indian pines, the classification accuracy of DTICF-3D-CNN-BiLSTM-MRF for HSI classification is showed. In this Fig. 4 of classification accuracy, CNN

Table 1 Overall, average and individual class accuracies (%) and kappa statistics of all competing methods on the Indian pines image test set

| Class | CNN (%) | CNN_MRF (%) | 3DCNN (%) | 3DCNN_MRF (%) | IC_3DCNN_BiLSTM (%) | IC_3DCNN_BiLSTM_MRF (%) |
|-------|---------|-------------|-----------|---------------|---------------------|-------------------------|
| 1 | 34.15 | 31.71 | 69.69 | 78.32 | 93.23 | 91.05 |
| 2 | 89.57 | 89.57 | 89.11 | 90.37 | 91.20 | 96.81 |
| 3 | 88.62 | 90.36 | 91.98 | 91.88 | 92.95 | 98.76 |
| 4 | 95.12 | 98.12 | 97.92 | 99.13 | 98.68 | 98.85 |
| 5 | 94.01 | 94.93 | 96.27 | 95.29 | 96.20 | 99.12 |
| 6 | 95.59 | 95.74 | 94.09 | 96.08 | 96.18 | 98.18 |
| 7 | 76.00 | 76.00 | 79.37 | 77.04 | 88.52 | 93.75 |
| 8 | 98.84 | 98.84 | 98.59 | 99.89 | 99.18 | 98.63 |
| 9 | 100.00 | 100.00 | 99.64 | 100.00 | 98.38 | 99.92 |
| 10 | 91.99 | 94.15 | 93.84 | 95.90 | 92.09 | 98.64 |
| 11 | 95.07 | 96.42 | 89.97 | 97.73 | 98.39 | 97.72 |
| 12 | 87.43 | 89.49 | 91.87 | 94.21 | 92.20 | 99.61 |
| 13 | 98.37 | 99.46 | 98.65 | 100.00 | 98.84 | 99.05 |
| 14 | 97.98 | 98.07 | 98.06 | 99.46 | 98.31 | 98.74 |
| 15 | 89.91 | 92.22 | 95.02 | 96.30 | 97.04 | 99.50 |
| 16 | 98.80 | 98.80 | 99.25 | 99.55 | 99.12 | 99.86 |
| OA | 73.50 | 84.62 | 95.24 | 96.75 | 97.36 | 98.82 |
| AA | 79.65 | 80.26 | 92.71 | 94.45 | 96.16 | 98.05 |
| KA | 72.28 | 83.40 | 94.67 | 95.82 | 96.97 | 98.61 |

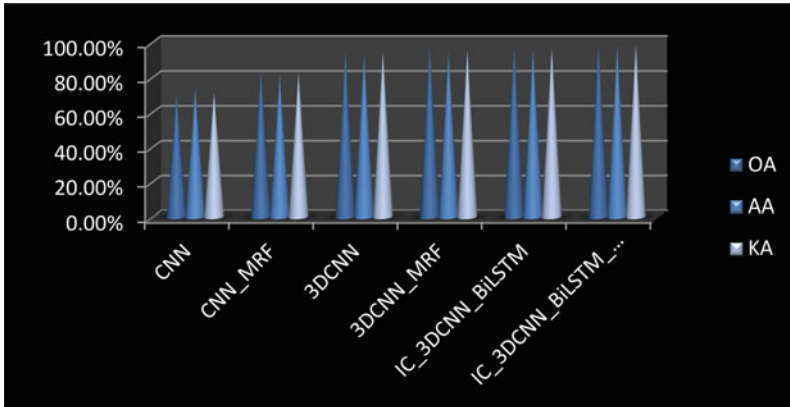


Fig. 5 Classification accuracy of proposed method and other classification methods on Indian pines dataset

and CNN-MRF is 73.50 and 84.62% for HSI classification. 3D-CNN and 3D-CNN-MRF is 95.24 and 96.75% for HSI classification which is larger than that of CNN and CNN-MRF. Finally, the proposed method of DTICF-3D-CNN-BiLSTM-MRF achieves significant higher performance for HSI classification accuracy. In Fig. 6, shows the ground truth, training Map, Testing map, DTICF-3D-CNN-BiLSTM, and DTICF-3D-CNN-BiLSTM-MRF. Finally, our proposed method DTICF-3D-CNN-BiLSTM with MRF has shown good classification performance.

To examine the proposed method, pavia university data is used. In Table 2, discuss the classification results compared to other 11 classification methods.

In Fig. 7 of pavia university, the classification accuracy of DTICF-3D-CNN-BiLSTM-MRF for HSI classification is showed. In this Fig. 4 of classification accuracy, CNN and CNN-MRF is 03.50 and 95.68% for HSI classification. 3D-CNN and 3D-CNN-MRF is 94.37 and 97.33% for HSI classification which is larger than that of CNN and CNN-MRF. Finally, the proposed method of DTICF-3D-CNN-BiLSTM-MRF achieves significant higher performance for HSI classification accuracy. In Fig. 8, shows the ground truth, training Map, Testing map, DTICF-3D-

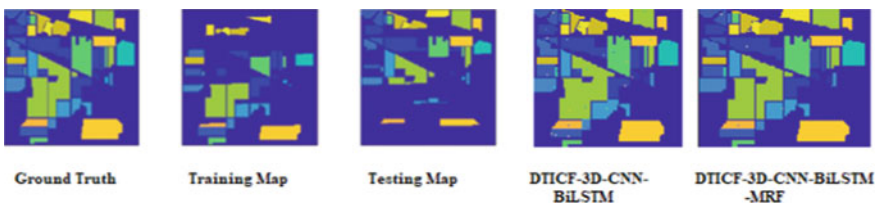


Fig. 6 Classification results obtained by DTICF-3D-CNN-BiLSTM-MRF on the Indian Pines dataset

Table 2 Overall, average and individual class accuracies (%) and kappa statistics of all competing and proposed methods on the pavia university image test set

| Class | CNN (%) | CNN_MRF (%) | 3DCNN (%) | 3DCNN_MRF (%) | IC_3DCNN_BiLSTM (%) | IC_3DCNN_BiLSTM_MRF (%) |
|-------|---------|-------------|-----------|---------------|---------------------|-------------------------|
| 1 | 96.78 | 97.92 | 97.8 | 98.15 | 98.54 | 99.29 |
| 2 | 96.92 | 97.38 | 95.83 | 96.85 | 97.27 | 98.30 |
| 3 | 85.96 | 87.66 | 91.92 | 92.09 | 94.22 | 95.92 |
| 4 | 98.78 | 98.97 | 97.31 | 97.55 | 98.46 | 98.71 |
| 5 | 99.92 | 99.92 | 98.38 | 99.07 | 99.04 | 99.27 |
| 6 | 90.10 | 92.00 | 95.7 | 96.42 | 97.17 | 98.13 |
| 7 | 84.42 | 85.27 | 89.14 | 89.72 | 91.36 | 94.81 |
| 8 | 89.84 | 91.54 | 96.31 | 96.72 | 97.74 | 99.07 |
| 9 | 96.80 | 97.13 | 99.38 | 99.43 | 99.49 | 99.68 |
| OA | 70.50 | 85.68 | 94.37 | 97.33 | 97.93 | 98.67 |
| AA | 70.44 | 84.20 | 95.76 | 96.22 | 97.03 | 98.13 |
| KA | 77.46 | 84.26 | 93.26 | 96.69 | 97.33 | 98.01 |

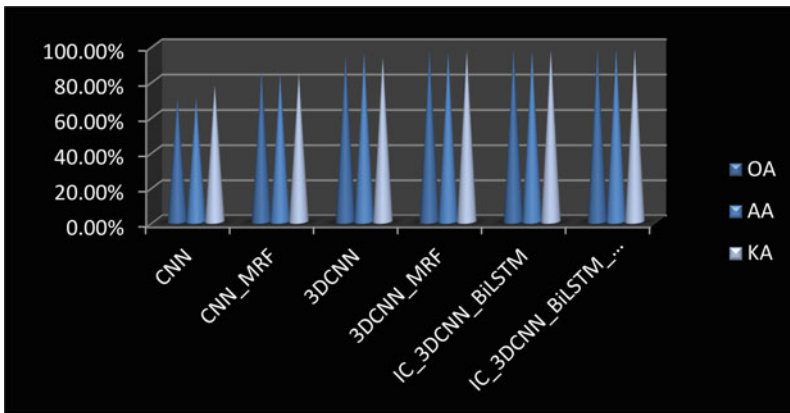


Fig. 7 Classification accuracy of proposed method and other classification methods on Pavia University dataset

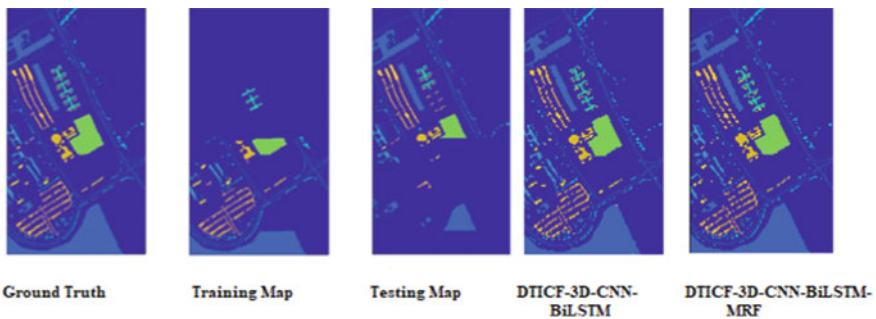


Fig. 8 Classification result by DTICF-3D-CNN-BiLSTM-MRF on the Pavia University dataset

CNN-BiLSTM, and DTICF-3D-CNN-BiLSTM-MRF. Finally, our proposed method DTICF-3D-CNN-BiLSTM with MRF has shown good classification performance.

5 Conclusion

To improve the HSI classification, the proposed 3D-CNN-BiLSTM framework has been proposed that is employed to extract the features of spectral- spatial information in this work. HSI data is converted to RGB image with spatial features. DTICF is applied to the combination of the RGB image with spatial features and raw HSI data. The excerpted features are provided to the 3D-CNN-BiLSTM. Finally, MRF is utilized to improve the features map for smoothing the classification result. The proposed DTICF-3D-CNN-BiLSTM-MRF approach compared with other HSI classification methods. The experimental result clearly viewed that DTICF-3D-CNN-BiLSTM-MRF based HSI classification attained the welfare classification accuracy. In future work, we will concentrate on how to reduce computational time across a variety of HSI datasets.

References

1. Bandos, T., Bruzzone, L., and Camps-Valls, G., "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 862–873, 2009.
2. Chen, Y., Lin, Z., Zhao, X., Wang, G., and Gu, Y., "Deep learning-based classification of hyperspectral data," *IEEE Journal Of Topics In Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094–2107, 2014.
3. Chen, Y., Nasrabadi, N. M., and Tran, T. D., "Hyperspectral image classification via kernel sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 217–231, 2013.
4. Dalla Mura, M., Villa, A., Benediktsson, J. A., Chanussot, J., and Bruzzone, L., "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 542–546, 2011.
5. Dhaya, R., "Analysis of Adaptive Image Retrieval by Transition Kalman Filter Approach based on Intensity parameter," *Journal of Innovative Image Processing*, vol. 03, no. 01, pp. 7-20, 2021.
6. Ellis, D.M., Draper, N.P., and Smith, H.S., "Applied regression analysis." *Applied Statistics*, Vol. 17, no. 1, pp. 83-90, 2014.
7. Fauvel, M., Benediktsson, J., Chanussot, J., and Sveinsson, J., "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3804–3814, 2008.
8. Guo, Y., Cao, H., Han, S., Sun, Y., and Bai, Y., "Spectral-spatial hyperspectral image classification with K-Nearest neighbor and guided filter." *IEEE Access*. Vol. 6, pp. 18582–18591, 2018.
9. Haokui Zhang , Ying Li , Yenan Jiang, Peng Wang, Qiang Shen , and Chunhua Shen, "Hyperspectral Classification Based on Lightweight 3-D-CNN With Transfer Learning," *IEEE Transactions On Geoscience And Remote Sensing*, VOL. 57, NO. 8, pp. 5813-5830, 2019.

10. He, K.M., Sun, J., and Tang, X.O. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 35, no.6, pp.1397–1409, 2013.
11. Hochreiter, S. and Schmidhuber, J., “Long short-term memory. *Neural computation*, vol. 9, no. 8, pp.1735-1780,1997.
12. Hughes, G. F., “On the mean accuracy of statistical pattern recognizers,” *IEEE Transaction Information Theory*, vol. 14, no. 1, pp 55-63, 1968.
13. Hu, W., Huang, Y., Li, W., Zhang, F., and Li, d H., “Deep convolutional neural networks for hyperspectral image classification,” *Journal of Sensors*, vol. 501, pp. 258619, 2015.
14. Jia, S., Shen, L., Zhu, J., and Li, Q., “A 3-D Gabor phase-based coding and matching framework for hyperspectral imagery classification.” *IEEE Transactions on Cybernetics*, Vol. 48, No. 4, pp. 1176–1188, 2018.
15. Jun li., and jose M. Bioucas, “Spectral- spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields,” *IEEE* vol. 50, no.3, 2012.
16. Kang, X., Li, S., and Benediktsson, J. A., “Feature extraction of hyperspectral images with image fusion and recursive filtering.” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3742–3752, 2014.
17. Konstantinos Makantasis, Konstantinos Karantzas, and Anastasios Doulamis, “Deep supervised learning for hyperspectral data classification through convolutional neural networks,” *IEEE*. pp.4959- 4962, 2015.
18. Landgrede, D, A., “Hyperspectral image data analysis, *IEEE signal process*,” Mag. 1053–5888, pp.17–28, 2002.
19. Licciardi, G., Marpu, P. R., Chanussot, J., and Benediktsson, J. A., “Linear versus nonlinear pca for the classification of hyperspectral data based on the extended morphological profiles,” *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 447–451, 2012.
20. Li, J., Bioucas-Dias, J. M., and Plaza, A., “Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4085–4098, 2010.
21. Li, J., Bioucas-Dias, J. M., and Plaza, A., “Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 809–823, 2012.
22. Li, J., Bioucas-Dias, J. M., and Plaza, A., “Hyperspectral image segmentation using new Bayesian approach with active learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp3947-3960, 2011.
23. Li, Y., Zhang, H., and Shen, Q., “Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network,” *Remote Sens*, vol.9, pp. 67-74, 2017.
24. Melgani, F., and Bruzzone, L., “Classification of hyperspectral remote sensing images with support vector machines,” *IEEE Transactions On Geoscience And Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
25. Nirmal, Sowmay, Soman, “Open Set Domain adaptation for Hyperspectral Image Classification Using Generative Adversarial etwork,” *Springer*, pp.819–827, 2021.
26. Oliveira, M.M., and Gastal, E.S., “Domain transform for edge-aware image and video processing.” *ACM Transactions on Graphics (ToG)*. ACM, Vol. 30,no. 4, pp. 69. 2011.
27. Qin Xu , Yong Xiao, Dongyue Wang and Bin Luo, “CSA-MSO3DCNN: Multiscale Octave 3D-CNN with Channel and Spatial Attention for Hyperspectral Image Classification,” *Remote Sens*. 2020.
28. Radhesyam Vaddi, and Prabukumar Manoharan, “Hyperspectral image classification using CNN with spectral and spatial features integration,” *Infrared Physics and Technology*, vol. 107, Elsevier, 2020.
29. Shaohui Mei , Jingyu Ji , Qianqian Bi , Junhui Hou , and Qian Du, “Integrating spectral and spatial information into deep convolutional neural networks for hyperspectral classification,” *IEEE* . 5067- 5070 2016.
30. Shen, L., and Jia, S., “Three-dimensional gabor wavelets for pixel-based hyperspectral imagery classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 5039–5046, 2011.

31. Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., and Woo, W.C., “Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting,” *Advances in Neural Information Processing Systems*, Volume 28, pp. 1049–5258, 2015. Available online: <https://papers.nips.cc/paper/5955-convolutional-lstm-network-a-machine-learning-approach-for-precipitation-nowcasting> (accessed on 10 October 2019).
32. Subarna shakya, “Unmanned Aerial Vehicle with Thermal Imaging for automating water status in vineyard,” *International journal of Electrical Engineering and automation*, vol. 14, pp. 79–91, 2021.
33. Sun, X., Qu, Q., Nasrabadi, N. M. and Tran, T. D. “Structured priors for sparse-representation-based hyperspectral image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1235–1239, 2014.
34. Szegedy, C., Toshev, A., and Erhan, D., “Deep neural networks for object detection,” in *Advances in Neural Information Processing Systems*, pp. 2553–2561, 2013.
35. Yushi Chen , Zhouhan Lin and Xing Zhao, “Deep Learning-Based Classification of Hyperspectral Data,” *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, IEEE, vol.7 no.6, pp.2094-2107, 2014.

Performance Analysis of Various Asymmetric Public-Key Cryptosystem



Amogh Desai , Virang Parekh , Utsav Unadkat ,
and Narendra Shekocar

Abstract Today, public-key cryptosystems have brought a significant change to the world of computers. Cryptosystems like the RSA and the ECC algorithm tend to top the table regarding usage and reliability. This study aims at meticulous and nuanced research on asymmetric public-key cryptosystems and compares them based on performance-based parameters and metrics. This research involves a comprehensive, comparative, and detailed study of the RSA, ElGamal, and the ECC-ElGamal Public-key Cryptosystems. Performance analysis has been performed considering the wall time, CPU, and memory used for both encryption and decryption of the algorithms. Security has also been measured by finding the similarity between plaintext and ciphertext obtained from these algorithms. The study aims to provide definitive results about the performance characteristics of the algorithms surveyed.

Keywords Cryptosystem · Rivest · Shamir · And Adleman (RSA) · Optimal asymmetric encryption padding (OAEP) · ElGamal · Elliptical curve cryptography (ECC) · Performance analysis

1 Introduction

Cryptography is the method of securing and protecting information by converting the original plaintext into indecipherable ciphertext using techniques derived from mathematical concepts. There are two main types of cryptography algorithms: Private-key or symmetric cryptosystems and Public-key or asymmetric cryptosystems.

A. Desai (✉) · V. Parekh (✉) · U. Unadkat · N. Shekocar
Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai,
India
e-mail: amoghd9@gmail.com

V. Parekh
e-mail: virangparekh73@gmail.com

N. Shekocar
e-mail: narendra.shekocar@djsce.ac.in

Private-key cryptosystems use a single private (secret) key for both encryption and decryption. The challenge with this type of cryptography is the key management, distribution issue. Public-key cryptosystems use a pair of keys, namely—Public and Private keys, wherein the former is known to others, and no one knows the latter except for the owner. These keys are generated based on various cryptographic algorithms using different one-way functions, i.e., easy to compute but hard to invert. The effectiveness of these cryptosystems lies in keeping the private key secret and openly distributing the public key without compromising security.

Asymmetric cryptosystems are generally slower than symmetric cryptosystems and are inefficient as they are used only for short messages. But asymmetric cryptosystems are more secure than symmetric cryptosystems as they do not need the secret key being shared over the transmission media which prevents attacks such as Man in Middle Attack (MIM). Another advantage of the asymmetric cryptosystem is that the receiver can verify the integrity of the data, i.e., it came from someone who has a private key, and the message can be trusted. Hence confidentiality, authentication, and non-repudiation of the message are ensured simultaneously. Popular asymmetric cryptography algorithms such as RSA, ECC have ubiquitous applications worldwide in encryption and digital signatures, from credit cards to electronic mails.

Since it was invented in the late 1970s public-key cryptosystems have been fascinating to study. The publication of the Diffie-Hellman key exchange in 1976 and the introduction of the famous RSA (Rivest–Shamir–Adleman) algorithm in 1978 have further boosted interest in it. They have gained popularity and find applications in digital signatures, encryption procedures, and key agreement techniques. New and more effective algorithms have come up like the DSA, ElGamal encryption, and Elliptical Curve Cryptosystems (ECC).

In this work, we present a comparative analysis of three public-key cryptosystems, RSA with OAEP, ElGamal, and ECC-Elgamal, on various metrics and parameters that provide insights on the performance, security, and critical features of the cryptosystems. We have unified and gathered different parameters and tried to paint a complete picture of the algorithms.

2 Existing Systems

A vast number of papers have been researched to get the optimal parameters for the comparison of algorithms. This survey has enabled us to evaluate and draw a comparison between different techniques, avoid common pitfalls, follow best practices, and build upon the best techniques.

Rivest et al. [1] presented the RSA algorithm in a 1978 work that not only introduced a novel approach to asymmetrically encrypt messages, but also introduced the world to the concept of digital signatures, which makes the algorithm so essential and popular in the cryptography industry. RSA is still used to encrypt internet messages and the majority of online apps. Boneh [2], in their work, describes attacks possible

using the underlying mathematical tools employed by the RSA algorithm and gives a set of warning road signs and pitfalls for the vulnerability of the algorithm and guides towards the dangers of misusing the RSA. In their work, Orman et al. [3] take a step towards quantifying the strength of a public-key algorithm based on known parameters and discuss approaches to resist large-scale attacks on the algorithm. They discuss the algorithm's security effect based on changing parameters like the system strength and key strength. While this work seems to take a step towards giving an insightful comparison, the study is only done on the brute force attack on the RSA. The Optimal Asymmetric Encryption Padding (OAEP) was created by Bellare and Rogaway [9] to prevent attacks on RSA. A random padding function is applied to the Plaintext before encryption with RSA, and hence different ciphertexts are obtained every time the message is encrypted. Shoup [11] and Fujisaki et al. [12], find the RSA-OAEP encryption scheme to be IND-CCA2 secure under the assumptions that the RSA problem is hard and that the adversary treats the hash functions instantiating OAEP as random oracles. Brown et al. [10] further prove that RSA-OAEP is even secure when the adversary is not restricted to treat the hash functions as random oracles.

Rabin [4] introduced the Rabin cryptosystem. The Rabin algorithm uses the factorization problem as the trap-door for its effectiveness. Its security is considered as effective as the hardness of the factorization problem. For decryption, the system uses the Chinese remainder theorem (CRT) to decrypt the message to plaintext. As a characteristic of the CRT, the algorithm outputs four different options for the plaintext, out of which one shall be the authentic message expected. Karakra et al. [5] tried to maintain the same security level and reduce the key size of the RSA algorithm by adding a layer of Rabin Cryptosystem to the RSA algorithm. They also implemented Huffman encoding to reduce the file's size and avoid indirect attacks like FOB. This made the encryption and decryption process 45 and 99% faster than the original RSA algorithm while improving the memory efficiency by 54%. It is also more secure against the frequency of block attacks and brute force attacks than RSA.

Koblitz et al. [6] discuss the ECC cryptosystem. It is known to perform better than the RSA system for specific curves. The CPU, memory, and key size are heavily reduced by almost 1.6 times and increase the security by three times. ElGamal [7] proposed yet another asymmetric crypto-system in 1985, which, like the ECC, is based on the problem of discrete logarithms over finite fields.

Raj [14] in their study presents an energy-efficient, privacy-preserving strategy to content exchange. Encryption using public keys and ciphertext policy attributes works effectively. At the end user level, ciphertext decryption minimizes the energy consumption of the secure proxy decryption process. The privacy-preserving content exchange system's routine is improved by the proxy encryption and decryption operation. This solution does not provide any information to unauthorized parties and does not have a substantial impact on end users' energy use.

Satheesh's [15] proposed approach has a major goal which is to validate the concept of accountable OFE. A third party that is accountable and transparent is offered, along with a general model of OFE. To achieve all security criteria, the

proposed architecture relies on numerous well-studied cryptographic algorithms. In this essay, the three types of accountability are explained. According to each OFE protocol, the qualities must be met.

The algorithms employed have been selected with care to ensure a complete and comprehensive view of the existing asymmetric encryption and decryption algorithms. The algorithms mentioned above and included in the final comparison encompass the current state of public-key cryptosystems.

3 Proposed Setup

This section describes the experimental setup for the performance analysis of the various algorithms with their details.

3.1 Existing Algorithm Overview

The RSA algorithm uses the trap-door function of multiplying two large prime numbers for generating the public key. The algorithm's effectiveness lies in the fact that it is tough to factorize two large prime numbers.

The ECC cryptosystem uses the elliptic curve discrete logarithm problem as the trap-door function and its hardness as the effectiveness. This "hopping the curve" phenomenon is what ECC is based on. The ElGamal algorithm, like the ECC algorithm, is based on the problem of discrete logarithms over finite fields. Though the discrete logarithm problem is not always hard and depends on the groups chosen, the hardness of this crypto-system lies in the fact that it is challenging to find a discrete logarithm in a cyclic group [9]. The following elliptic curves are considered secure and have been used in the ECC-ElGamal implementation.

1. E222: It is a 222-bit prime field Edwards curve. The curve equation is as follows—

$$x^2 + y^2 \equiv c^2(1 + dx^2y^2). \quad (1)$$

2. E382: It is a 382-bit prime field Edwards curve. The curve equation is as follows—

$$x^2 + y^2 \equiv c^2(1 + dx^2y^2). \quad (2)$$

3. M383: It is a 383-bit prime field Montgomery curve. The curve equation is as follows—

$$By^2 \equiv x^3 + Ax^2 + x. \quad (3)$$

4. Curve25519: 255-bit prime field Montgomery curve. The curve equation is as follows—

$$By^2 \equiv x^3 + Ax^2 + x. \quad (4)$$

5. P256: It is a 256-bit prime field Weierstrass curve. The curve equation is as follows—

$$y^2 \equiv x^3 + ax + b. \quad (5)$$

6. secp256k1: It is a 256-bit prime field Weierstrass curve. The curve equation is as follows—

$$y^2 \equiv x^3 + ax + b. \quad (6)$$

These curves and the and the algorithm complemented by the high security it offers per key bit size compared to RSA as mentioned by Mallouli et al. [13], makes it one of the optimal algorithms for encryption.

3.2 Experimental Setup

Performance testing of the algorithms mentioned earlier is performed to assess their efficiency and viability for encrypting messages. Benchmarking for encryption and decryption has been performed in Python using the various inbuilt libraries.

1. Execution time—*timeit* library, with five loop repeats.
2. Memory use—*psutil*
3. CPU use—*psutil*

Similarity between the plaintext and ciphertext is computed using the Damerau-Levenshtein Distance, a score representing the minimum number of insertions, deletions, or substitutions of a single character, or transposition of two adjacent characters required to change one word into the other.

Open Source Libraries and Code have been used for cryptographic algorithms. This is done to ensure optimized, tested, bug-free code.

1. RSA—*PyCryptodome* library
2. ElGamal—*RyanRiddle/elgamal* on Github
3. ECC-ElGamal—*ecc-ptycrypto* library

Benchmarking has been performed varying the key size and the message size. Safe, practically secure key sizes and verified curves have been considered to emulate real-world working.

1. Message Size—1 byte, 8 bytes, 64 bytes, 256 bytes, and 512 bytes

2. Key Size—RSA and ElGamal: 1024, 2048, 3072, 4096
3. Curves—ECC-ElGamal algorithm: E222, E382, M383, Curve25519, P256, secp256k1.

Performance testing is performed independently on all the parameters for both encryption and decryption. The message size is varied for each key size to check the maximum encryptable bytes. The final values are rounded to two decimal places, and the graphs are created in Plotly.

4 Experimental Results and Analysis

Benchmarks have been carried out multiple times on various high-powered laptops and desktops to ensure accuracy, but minor errors might exist in the measurements due to background processes.

“-” Indicates that the algorithm could not encrypt/decrypt the message for a given key size, and hence no measurements have been obtained for the same.

4.1 CPU Utilization

The CPU Utilization for the RSA, ElGamal, and ECC-ElGamal algorithms has been calculated using the psutil library in Python. The CPU Utilization is calculated as follows:

```
curr_cpu_use = psutil.getloadavg()[0]/psutil.cpu_count()*100
```

As a tuple, the getloadavg() function returns the average system load during the 15 min. The load indicates the processes that are either running or waiting to run on the CPU.

The results are displayed in Figs. 1, 2 and 3.

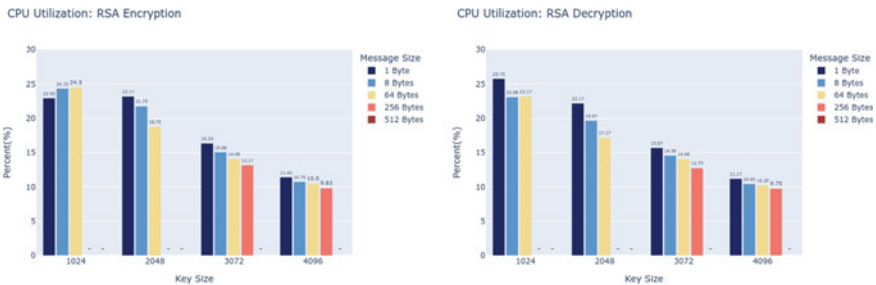


Fig. 1 CPU utilization for RSA

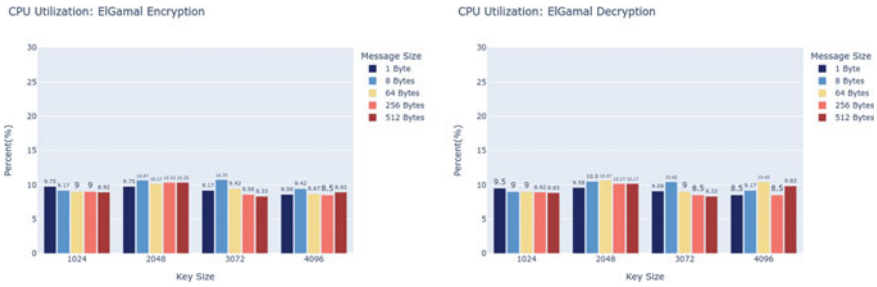


Fig. 2 CPU utilization for ElGamal

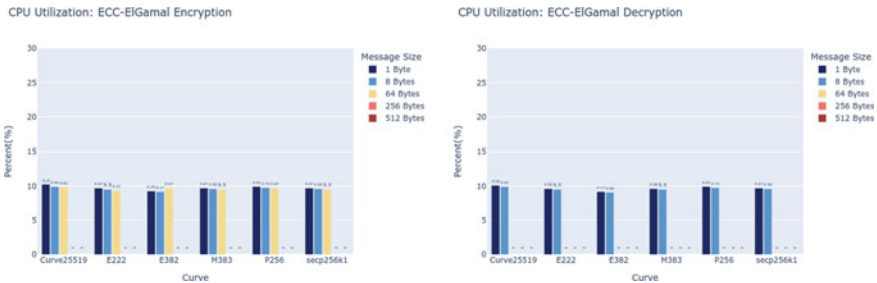


Fig. 3 CPU utilization for ECC-ElGamal

Observations—As shown in Fig. 1 RSA shows a rather peculiar trend of CPU utilization, especially during encryption process. Similarly, Fig. 2 depicts CPU use for the ElGamal method, while Fig. 3 depicts the same for the ECC-ElGamal algorithm, both of which exhibit a fairly steady pattern of values despite changes in key sizes.

We see that overall, larger keys have a lower CPU use. For a constant key size, the CPU use decreases by increasing the size of the message. Decryption consumes less CPU than encryption. Encryption and Decryption values follow similar trends, i.e., they increase and decrease together. RSA has a large discrepancy in CPU usage, 22.92% for a 1024-bit key encrypting a 1-byte message to 9.83% for a 4096-bit key encrypting a 256-byte message. ElGamal consumes the least CPU on average, followed by ECC-ElGamal.

Inference—Buffering and caching might be the reason for Observation 1, 2, and is the most prevalent in the RSA implementation, as observed in 5 above. ElGamal can encrypt larger messages while still consuming less CPU and might be suitable for use-cases with lower CPU requirements.

4.2 Memory Utilization

The Memory Utilization for the RSA, ElGamal, and ECC-ElGamal algorithms has been calculated using the psutil library in Python. The Memory Utilization calculation has been described below.

```
curr_mem = psutil.Process().memory_info().rss / (1024 * 1024)
```

The Process() class is used to represent an process in the OS and the memory_info() function and the rss method return the resident set size of the current process in kilobytes.

The results are displayed in Figs. 4, 5 and 6.

Observations—Fig. 4 depicts the RSA Algorithm’s Memory Usage, which demonstrates a clear upward trend in values with rising key sizes for both encryption and decryption processes, regardless of message size. Figure 5 depicts memory usage for ElGamal and Fig. 6 depicts memory usage for ECC-ElGamal algorithms, which use about the same amount of memory regardless of key or message size.

The memory used increases with an increase in key size and message size. Decryption consumes slightly more memory than encryption. The memory used increases

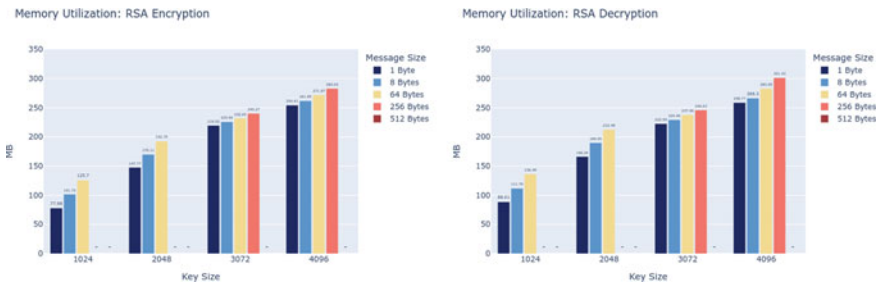


Fig. 4 Memory utilization for RSA

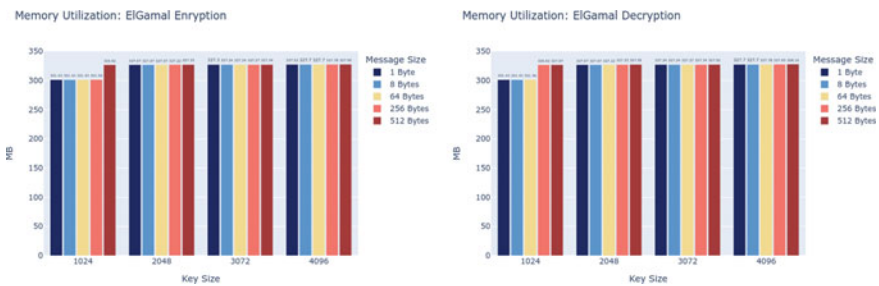


Fig. 5 Memory utilization for ElGamal



Fig. 6 Memory utilization for ECC-EIGamal

the most in RSA and is essentially unchanged in ElGamal and for all curves in ECC-ElGamal; RSA occupies the least memory on average. ElGamal and ECC-ElGamal approximately occupy the same memory.

Inference—RSA uses memory more efficiently as compared to ElGamal, ECC-ElGamal. All the algorithms occupy similar amounts of memory for larger key and message sizes.

4.3 Execution Time

The Execution Time for the RSA, ElGamal, and ECC-ElGamal algorithms has been calculated using the `timeit.perf_counter` method in Python. The calculation done as follows:

```
curr_time = %timeit -r 5 -o -q -c -p 2 benchFunc(*args)
```

`%timeit` measures the execution time of a given code snippet, inline.

The results are displayed in Figs. 7, 8 and 9.

Observations—Except for the 1-byte message, the execution time for RSA is shown in Fig. 7, which is rapidly growing. The execution time for the ElGamal algorithm is shown in Fig. 8, which exhibits a completely growing pattern throughout all



Fig. 7 Execution time for RSA



Fig. 8 Execution time for ElGamal

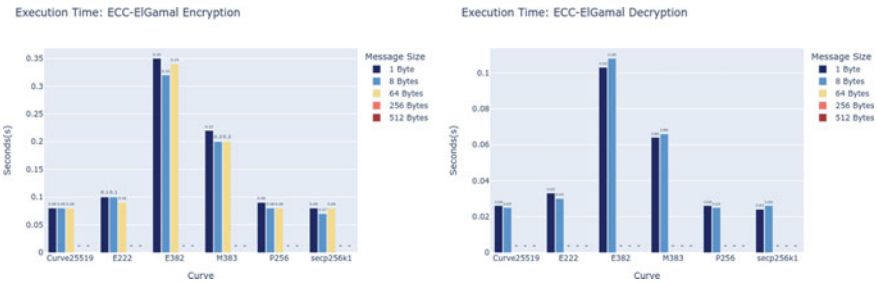


Fig. 9 Execution time for ECC-ElGamal

message sizes; however, there is a significant variation for the 512-byte message. The execution time for ECC-ElGamal is seen in Fig. 9, which first climbs and subsequently decreases.

Decryption time is significantly lesser than encryption time. For a constant key size, the execution time for encryption is variable and does not follow a specific pattern. A more constant trend is observed for the execution time for decryption. Decryption time stays constant for a specific key size in RSA but varies nominally for ElGamal and ECC-ElGamal. Encryption time increases exponentially for 3072 and 4096-bit keys in ElGamal. The E382 curve in ECC-ElGamal takes the most time for decryption among all the other curves.

Inference—Buffering and caching might be the reason for Observation 2 and is the most prevalent in the RSA implementation. ElGamal is inefficient for larger keys as it consumes the most time.

4.4 Plaintext and Ciphertext Similarity

The Similarity Score for the RSA, ElGamal, and ECC-ElGamal algorithms has been calculated using the Damerau-Levenshtein Distance. The results are displayed in Figs. 10, 11 and 12.

Similarity: RSA

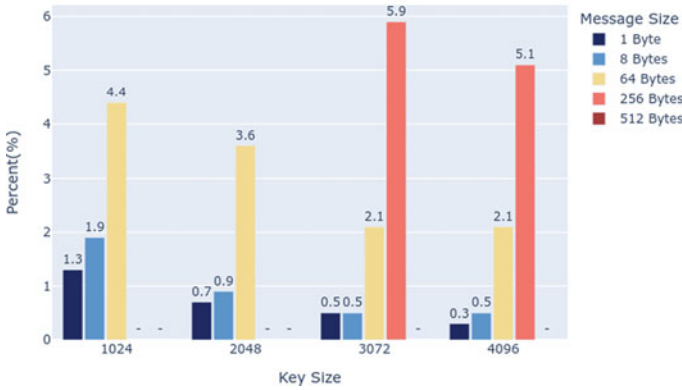


Fig. 10 Similarity score for RSA

Similarity: ElGamal

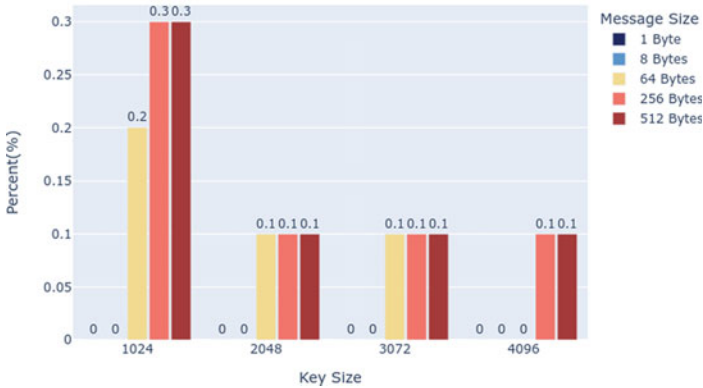


Fig. 11 Similarity score for ElGamal

Observations—Fig. 10 depicts the RSA Algorithm’s Similarity Score, which demonstrates a declining to steady pattern of values. The Similarity Score for ElGamal is shown in Fig. 11, which is basically stable except for the 1024 bit-size. ECC-ElGamal is more or less consistent over a range of curves, as seen in Fig. 12.

As the key size increases, the similarity measure between the plaintext and ciphertext decreases. ElGamal has the lowest similarity measure between the plaintext and ciphertext and is equal to 0.0 for 1 byte and 8-byte message sizes for all key sizes. RSA has the largest similarity index value, greater than 5% for a 256-byte message using a 3072 and 4096-bit key. Curve E382 and M383 have the smallest similarity index among all the curves in ECC-ElGamal because they are 384-bit curves.

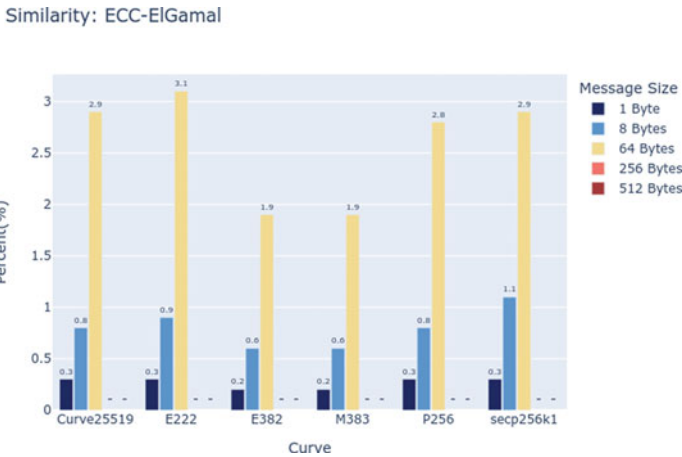


Fig. 12 Similarity score for ECC-ElGamal

Inference—Although the similarity between the plaintext and ciphertext is not the most deterministic measure for verifying the strength of an algorithm, ElGamal is the safest. The similarity measure for all algorithms is low overall, and hence they can be considered safe.

5 Conclusion

This paper compares the RSA-OAEP, ElGamal, and ECC-ElGamal algorithms on performance metrics—CPU utilization, Memory Use, Execution Time for encryption and decryption, and a security metric—similarity between plaintext and ciphertext measured using the Damerau-Levenshtein distance. The results show that RSA consumes the most CPU, 22.92%, and has a high-similarity index, 5%. At the same time, ElGamal occupies the most memory ~328 MB and takes the most time for encryption and decryption, especially for large key sizes, 82 and 1.7 MB, respectively. But ElGamal keys can encrypt 512-byte messages, which is not achievable by other algorithms. ECC-ElGamal consumes almost as much memory as ElGamal but has moderate CPU use, execution time, and similarity index. This, combined with the higher security it provides per key bit size compared to RSA, makes it one of the best encryption methods available. The research currently uses only alphabetical messages for encryption. We wish to extend the research by comparing different types of data for encryption and including more security measures for a broad comparison between the algorithms. Including hybrid-encryption algorithms and further optimizing the existing algorithms will increase the breadth of comparison, benefiting the community.

References

1. Rivest, R., Shamir, A. and Adleman, L., 1983. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 26(1), pp.96–99. (1983)
2. Boneh, Dan. (2002). Twenty Years of Attacks on the RSA Cryptosystem. *NOTICES OF THE AMS*. 46. (2002)
3. Orman, H. and Hoffman, P., 2004. Determining Strengths For Public Keys Used For Exchanging Symmetric Keys. (2004)
4. Rabin, M., 1979. Digitalized Signatures And Public-Key Functions As Intractable As Factorization. (1979)
5. Karakra, A. and Alsadeh, A., 2016. A-RSA: Augmented RSA. 2016 SAI Computing Conference (SAI). (2016)
6. Koblitz, N., Menezes, A. and Vanstone, S., 2000. *Designs, Codes and Cryptography*, 19(2/3), pp.173-193. (2000)
7. Elgamal, T., 1985. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4), pp.469–472. (1985)
8. Tsiounis, Y. and Yung, M., 1998. On the security of ElGamal based encryption. *Public Key Cryptography*, pp.117–134. (1998)
9. Bellare, M. and Rogaway, P., 1993. Random oracles are practical. *Proceedings of the 1st ACM conference on Computer and communications security - CCS '93*. (1993)
10. R. L. Brown, D., 2007. What Hashes Make RSA-OAEP Secure?. [online] Available at: <<https://www.semanticscholar.org/paper/What-Hashes-Make-RSA-OAEP-Secure-Brown/4d96ea89cbb277af3ebfb552ece3be8c038d0797>> [Accessed 29 January 2022]. (2007)
11. Shoup, V., 2002. OAEP Reconsidered. *Journal of Cryptology*, 15(4), pp.223-249. (2002)
12. Fujisaki, E., Okamoto, T., Pointcheval, D. and Stern, J., 2002. RSA-OAEP Is Secure under the RSA Assumption. *Journal of Cryptology*, 17(2), pp.81–104. (2002)
13. Mallouli, F., Hellal, A., Sharief Saeed, N. and Abdulraheem Alzahrani, F., 2019. A Survey on Cryptography: Comparative Study between RSA vs. ECC Algorithms, and RSA vs. El-Gamal Algorithms. 2019 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), (2019)
14. S. Raj, J., 2021. Secure Data Sharing Platform for Portable Social Networks with Power Saving Operation. September 2021, [online] 3(3), pp.250–262. Available at: <<https://irojournals.com/iroismac/article/pdf/3/3/6>>
15. M, S. and .M, D., 2020. Implementation of Multifactor Authentication Using Optimistic Fair Exchange. *Journal of Ubiquitous Computing and Communication Technologies*, [online] 2(2), pp.70–78. Available at: <<https://www.ijournals.com/jucct/V2/I2/02.pdf>>

Design of a Web-Based Platform: Event-Venues Booking and Management System



P. S. JosephNg, S. M. Al-Sofi, K. Y. Phan, J. T. Lim, and S. C. Lai

Abstract Conducting an event require planning multiple tasks and managing finance for the event. Nevertheless, event management and event organization are more of a manual-based works and require a more automated system to produce more efficient and reliable reservations. This research aims to improve the event organization and management by promoting people to find, book a venue, and track the budget expenses through a web-based platform, Event-Venues Booking and Management System, by implementing cloud computing. After the data from a survey and an interview have been obtained and presented, the suggested system features are created based on the gathered results that would meet the needs of the future user. The findings suggest that the online platform assists individuals in arranging and managing their events in a more manageable and timely manner, while lowering the cost and time associated with it. Furthermore, the established system helps both users and suppliers to foster a positive connection between them, as determined by the research questions and objectives.

Keywords Event · Online booking · Budget tracking · Cloud computing · Venues · Search

P. S. JosephNg (✉)

Faculty of Data Science and Information Technology, INTI International University Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

e-mail: joseph.ng@newinti.edu.my

S. M. Al-Sofi

Institution of Computer Science and Digital Innovation, UCSI University UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia

K. Y. Phan · J. T. Lim · S. C. Lai

Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia

e-mail: phanky@utar.edu.my

J. T. Lim

e-mail: limjt@utar.edu.my

S. C. Lai

e-mail: laisc@utar.edu.my

1 Introduction

The internet plays a significant role in forming all business environments since the prior decade [1]. There is a high-speed improvement of online booking in the developing and improving world [2, 3]. Information and Communication Technology (ICT) is fast-growing, thus affecting and changing every business in this world. Environmental information systems hold a significant place in this centenary [4]. The infectious disease (Covid-19), first diagnosed in China in late 2019, has become a pandemic situation [5]. Among the most affected industries are \$1000 billion Event Management or Event Preparation Business and other industries [5]. Implementing modern technology not only does physical work in factories but also makes everyone's life simpler [6]. People can overcome the impact of the current situation by making life easier and safer.

The Web has grown as an essential medium in supporting successful purchasing of services and products. The Web has appeared as one of the numerous substantial channels for online booking distribution [7]. Online reservation systems which are used in regular life, seem to be available everywhere [8]. Online distribution channels of several hospitality-based services such as travel packages, hotel rooms, car rentals and flight reservations have quickly developed, and due to various advantages they contribute to both online e-retailers and users [9]. Consumer satisfaction is an official terminology followed in the business environment [10]. It refers to the consumers' expectations being satisfied by the quality of product or services experienced. This terminology is followed by various successful companies [11]. The potential of achieving e-satisfactory and e-Loyalty significantly depends on online user experience and how users engage and communicate with online booking websites [9]. The online booking website allows online users to get accurate knowledge about the booking facilities [12]. For instance, online users can now compare prices and other things without communicating with the representative [13].

'Event Planning' in Indian meaning is a centre of opportunities, but managing it, is far from sustainable [14]. When it comes for planning an event, a person has to set the occasion that needs to be organized by laying a plan and budget that will cover all the needs for the function [15]. Each event, no matter how complex or straightforward, needs comprehensive organization and planning [16].

One of the essential keys for organizing the event is budget tracking [17]. The current research studies the relationship between budget control and organizational effectiveness [18]. Budget management enhances strategy, teamwork, inspiration, collaboration, and senior management behaviour [19]. The study indicates that management at all stages and periods must ensure that budget occurrences are reviewed to prevent mismanagement and organizational loss [19]. According to the organization, it is said to be successful when it meets its mission and goal while offering a total return to the society [20].

When it comes to events, each person has at least one struggle with searching and booking the available venue for the event [21]. As there are many venues available

for different events, it is not easy to visit each venue to decide and compare. It is time consuming and requires a lot of resources to do that.

There are similar platforms for event management and venue booking such as BOOKiiIT, Celebrations, Zhengzhou University Online Booking System for Sports Venues, Online University Facilities Reservation System and more. BOOKiiIT is an app proposed to solve the problem in room booking related to the academic institute to overcome the problem of managing the booking spaces [22]. The event is an app piled to provide the student with an event by booking the event to attract a large audience. Skedda is an online booking and scheduling service for any space relative to the learning, such as office desks, coworking spaces, meeting rooms, sports venues, professional studios, classrooms, and labs [23]. Celebration is an online venue booking management system that runs on the web [24]. Zhengzhou University Online Booking System for Sports Venues has been designed to solve the unfair use of university sports venues in Zhengzhou University [25]. Online University Facilities Reservation System is another web-based development to solve the inappropriate use of the university facilities [26]. Afrah online is a web app that allows users only to view a venue.

Although online venue booking for events runs in the current era, it has not discovered a complete platform for booking venue and budget management for events, which provide this project to improve and solve the existing problems of the previous study using cloud computing [27]. The proposed system, “” helps the customers to compare and book venues in a short period without the need for travelling. Additionally, the new system provides the user with a budget tracking feature that allows the user to track the expenses regarding the event.

2 Problem Statement and Hypothesis

This project intends to provide an online platform for event venue booking and management. This system will solve the existing problem of organizing and managing the event by providing users with new technology to help with event organization. The hypothesis for each criterion obtained from the objectives have been shown in Fig. 1 below.

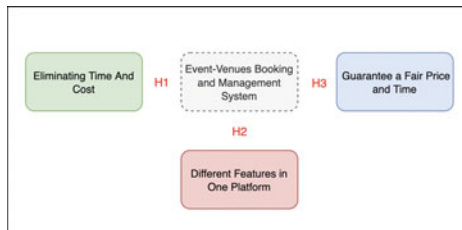


Fig. 1 Research model

H1: By eliminating the time and cost of searching for the venue, users can save more time and money for organizing the event.

To organize an event conventionally, a multi-visit for several places must be required, reflecting the total cost due to transportation, parking, taxi, and other expenses. Furthermore, physical search requires much time as the person needs to visit a venue by arranging the time of the visit for each venue, which may take days or weeks. Using the online platform “Event-Venues Booking and Management System” for booking venues for an event, users save time and money instead of physically searching for the site.

H2: By providing different features in one platform, users can manage and organize the event manageably.

By presenting various features grouped in one platform, users will not be required to use different platforms or applications to achieve a single task. It becomes an alternative for handling different tasks, such as booking, searching, finance tracking, managing booking, and payment by using the ‘Event-Venues Booking and Management System’.

H3: By using the online platform, users can guarantee a fair price when booking the venue.

By utilizing the platform, it can secure its fair price and guarantee the users that the price written on the platform is the same as it is genuine and assured by the Event-Venues Booking and Management System.

3 Methodology

This study practices the mixed method to clarify more effectively, the why and how questions from the target audience. The questions for the survey and interview to collect data, are created based on the objectives of the study. A Likert scale method is used to ask questions throughout the survey. The client’s viewpoint are explored in further depth throughout the interview questionnaire. The target group is random people that have an upcoming event or who have already organized any occasion. Accordingly, the results can be more reliable and more valuable. The following technique, as stated in Table 1, is used to gather the data.

Based on Table 1, the research dimension is demonstrated in the sequential design, which clarifies every data collection step, as shown in Fig. 2.

Based on Fig. 2, qualitative data gathering is applied to generalize various respondents’ opinions by applying a survey as a tool to gather more data efficiently. However, the interview provides an in-depth opinion of the responses. The interview and survey

Table 1 Research methodology [22, 23]

| | |
|----------------------|---------------------------------------|
| Research dimension | Explanatory sequential design |
| Research methodology | Quantitative & qualitative perceptive |
| Research method | Survey & interview |

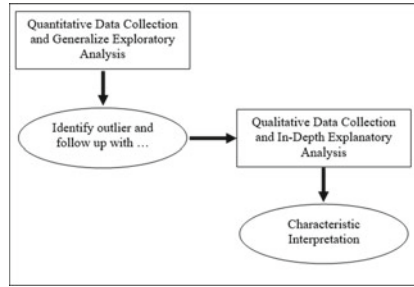


Fig. 2 Research model [28–30]

questions are distributed through social media platforms like WhatsApp, CN, and other platforms. This made it more comfortable to reach out to various people. The survey was taken up by one hundred thirty respondents and the interview was attended by thirty-three people. All the information obtained during the interview is utilized to double-check the accuracy of the survey responses. Lastly, all the data are classified and compiled to reach a single result.

4 Results and Findings

The findings and results of the research interviews and surveys have shown a trend in people’s opinions regarding events organization based on venue booking and budget management. 83.9% of people are concerned about the time and money they spend on the physical searching of venue. A multi-visit for several places is required to conventionally organize an event, reflecting the total cost due to transportation, parking, taxi, and other expenses. Figure 3 demonstrates the opinion of the people regarding the expense, for booking a venue online.

Figure 3 represents the responses of the participants based on how they think about booking a venue online; Does it save money? It shows that 80.8% of people agree. However, 19.2% disagree as they believe that there is no link between booking

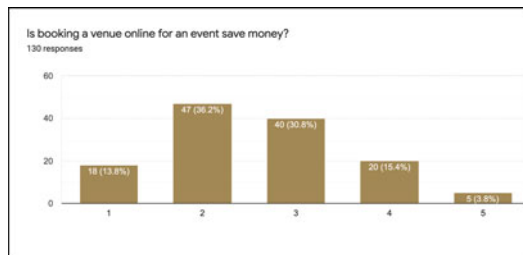


Fig. 3 Survey results related to cost

online and saving money. From the result, majority of the people think that online booking can save money and time as it provides various choices at different prices for different venues without physical visits.

As shown in Fig. 4, almost 84% of the respondents agree that having different features such as booking a venue and tracking the expenses for the event, are more efficient using the online platform. The high percentage of the agreement can be attributable to a variety of factors. Firstly, many people find challenges in using various platforms for different purposes. Besides, using one platform with different services will save memory as it does not require much space for a single application compared to many. Moreover, by presenting different features in one platform, users can handle the event organization in a manageable way. The result confirms that multiple features attract people more than one feature in one place.

Figure 5 reveals the people’s answers about providing detailed descriptions for a venue to help select a venue. A total of 84.6% agree that it provide them with more knowledge to decide which one to choose by presenting detailed descriptions such as price, location, capacity services, and more for the venue. People can compare between the venues which allow them to have a better picture about the venue. Moreover, providing feedback can help the customer, view the people’s experience from the bookings.

The satisfaction of the proposed online platform obtains positive feedback from both interview and survey respondents. With the online platform, users can search,

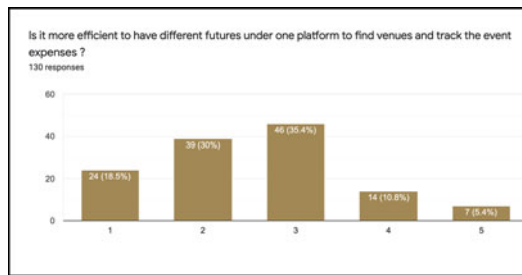


Fig. 4 Survey results related to efficiency

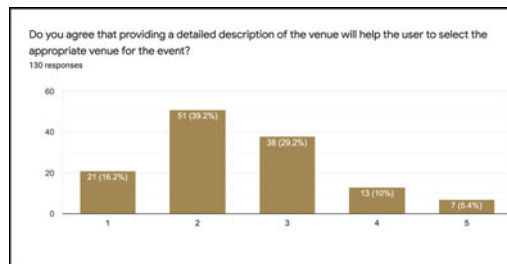


Fig. 5 Survey results related to branding

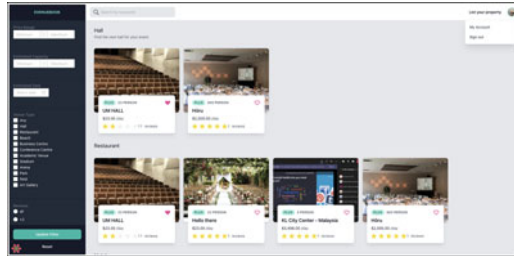


Fig. 6 Home page

check, and book a venue for the event. In addition, the online platform provides the user with the ability to keep an eye on the budget, which helps the user always be on track of expenses related to the event. From the feedback, this feature has a high positive attitude.

Based on data gathered, the proposed solution has been developed with three categories of users. The first type of user is the visitor. This type of user can search for the venue and view the details of the venue, as shown in Fig. 6.

Figure 6 presents the home page of the online application. This page represents all the venues, providing filters and a basic description of the venue, which helps the user find the venue based on user performance. The filters provided for the search allow the user to find the venue based on the price, capacity, date, type of venue, and reviews. Additionally, the user-provided search function is based on the name and location of the venue. When the user selects preferences for the venue from the filters or search, the list of venues are displayed according to the search. If the user clicks on the venue, all the detailed information of the venue are displayed, so that the user can have a better picture of the venue and be able to compare between other venues.

The second type is the authenticated user. This type of user has the same privilege as the visitor. Additionally, an authenticated user can book a venue after signing into the platform using google authentication. Moreover, this type of user is provided with various features like managing the booking, viewing the booking request, and receiving the bills when the payment is made. Moreover, the user is provided with tracking expenses feature.

Figure 7 illustrates the budget page. This page provides the user to create the event and set the budget for it. After the user finishes, the user needs to press on the event name, which will bring the user to the expenses page, as shown in Fig. 8.

Figure 8 demonstrates the budget tracking page. On this page, the user can view the budget, expenses, and total amount left, which is calculated from the expenses that the user has listed. This page allows the user to pen down all expenses for the event by citing the amount, the name, and description, which helps the user remember and track all the expenses.

The last type of user is the admin (vendor). This type of user lists all the venue details, manages all the requests, bookings, and payments for the clients.

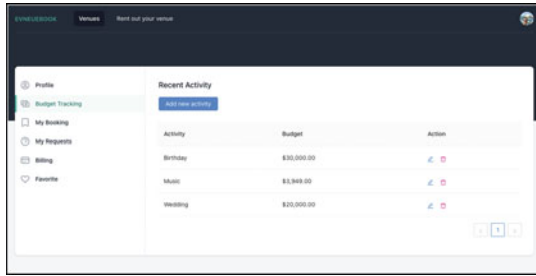


Fig. 7 Budget page

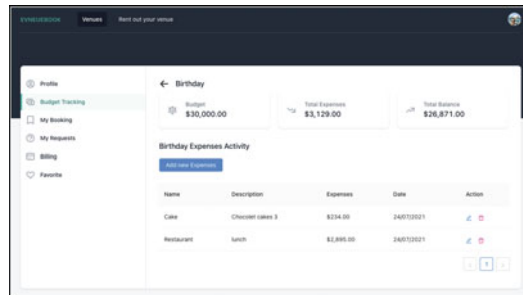


Fig. 8 Budget tracking page

The online venue booking and management system provides users with different functions that help organize the event more manageably, at reduced cost, and quicker. This system provides the small businesses, an opportunity to develop and earn a profit by attracting customers with their services.

5 Limitation and Future Work

This project has some limitations, such as the short duration of the study, lack of information from related works, and the inability to provide more venues. Nevertheless, these did not affect the objective of the project.

The future work is determined with the recommendation to improve by making the system more automated and providing online payment, chat, and map.

6 Conclusion and Limitation

With the advancement of technology, every business strives to increase profits by expanding their promotion through Internet service to attract more customers. One among the most common businesses globally, is the arrangement of venues for the event. Developing E-commerce for halls will increase revenues while also boosting clients by allowing them to book halls for events online on specific dates to save time. The present situation of E-commerce in the online booking sector has been investigated in this study. This project aims to study the possibility of establishing an e-commerce platform to solve the existing problems with the event organization. Customers can find a venue for the event, book the venue online, and manage the event's expenses in one place 'Event-Venues Booking and Management System'. The method carried out in this project is the 'mixed' method. From the result of this study, the online platform has been developed successfully by achieving the objectives of this project. In this way, people can enjoy searching for the venue and book the venue in a feasible way.

Moreover, there are two main contributions of e-venue booking, i.e., theoretical and practical. The theoretical contribution is that the e-venue booking is considered a new addition in the research space of venue bookings with its results and findings. These conclusions and outcomes will pave the way for future work in this domain. The practical contribution of this study is the actual usage of the system for both users and vendors, where it would benefit the industry commercially.

References

1. Mehrotra and J. Lobo, 'Technology Driving Event Management Industry to the Next Level', in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 2020, 1(1),436–441.
2. Bilgihan and M. Bujisic, 'The effect of website features in online relationship marketing: A case of online hotel booking', *Electron. Commer. Res. Appl.*, vol. 14, no. 4, pp. 222–232, Jul. 2015.
3. N. Serhiy, K. Yaroslav, and B. Kateryna, 'Online Reservation System Project', in *Proceedings of the 4th International Conference Computational Linguistics And Intelligent Systems*, Lviv, Ukraine, 2020, 2(1), 347–348.
4. F. Rodzi, E. Nasir, A. Azmi, D. Abdullah, A. Azmi, and S. Kamal, 'The Role of Compatibility, Information Quality and e-Service Quality in Predicting Mobile Hotel Booking Adoption: A Conceptual Framework', *Int. Acad. Res. J. Bus. Technol.*, 2(2), 123–128, 2016.
5. J. S. Madray, 'THE IMPACT OF COVID-19 ON EVENT MANAGEMENT INDUSTRY', *Int. J. Eng. Appl. Sci. Technol.*, 5(3), 533–535, 2020.
6. M Loke, PS JosephNg, AS Shibghatullah & HC Eaw (2020), Jomdrone: Data mining financial sense in the property agency, IEEE Symposium on Industrial Electronics and Application, Malaysia, 1-5
7. J. S. Zhang and Q. Lv, 'Understanding Event Organization at Scale in Event-Based Social Networks', *ACM Trans. Intell. Syst. Technol.*, 10(2), 16:1–16:23, 2019.

8. Emir, H. Halim, A. Hedre, D. Abdullah, A. Azmi, and S. Kamal, 'Factors Influencing Online Hotel Booking Intention: A Conceptual Framework from Stimulus-Organism-Response Perspective', *Int. Acad. Res. J. Bus. Technol.*, 2(2), 129–134, 2016.
9. M. Ahmed, S. H. Ahmed, and O. H. Ahmed, 'Dijkstra algorithm applied: Design and implementation of a framework to find nearest hotels and booking systems in Iraqi', in *2017 International Conference on Current Research in Computer Science and Information Technology (ICCCIT)*, Slemani - Iraq, Apr. 2017, 4(1), 126–132.
10. N. A. H. Zolkopli, S. S. Ramli, A. Azmi, and S. B. M. Kamal, 'Online Travel Shopping Intention', *Int. Acad. Res. J. Bus. Technol.*, 2(2), 140–144, 2016.
11. S. Kamal, D. Abdullah, N. Md Nor, A. Ngelambong, and K. Bahari, 'Hotel Booking Websites and their Impact on E-Satisfaction and E-Loyalty: Analysis on Utilitarian and Hedonic Features', *Int. Acad. Res. J. Bus. Technol.*, 8(15), 160–177, 2018.
12. KM Liow, PS JosephNg, YF Loh, JomDesignLab: Bringing Artwork Design Nearer, IEEE International Conference on Control Systems, Computing and Engineering, Malaysia, 1–6
13. N. Bikakis, V. Kalogeraki, and D. Gunopulos, 'Social Event Scheduling', *ArXiv180109973 Cs*, 4(2), 1272–1275, 2018.
14. P. Berners, *The Practical Guide to Organising Events*, 1st ed., 1(1). New York: Taylor & Francis, 2017.
15. H. Ujang, A. R. Omar, I. A. Rani, A. Azmi, S. B. M. Kamal, and D. Abdullah, 'Factors Influencing Consumers Intention to use Self Service Technology in Tourism and Hospitality Industry', *Int. Acad. Res. J. Bus. Technol.*, 2(2), 118–122, 2016.
16. M. Tiwari, T. Tiwari, S. Chaudhary, A. Marwah, and D. S. Bawa, 'Need for sustainable event management in the Indian context', *J. Inf. Optim. Sci.*, 41(5), 1291–1297, 2020.
17. T. C. Greenwell, L. A. Danzey-Bussell, and D. J. Shonk, *Managing Sport Events*, 1st ed., 1(1). USA: Human Kinetics, 2019.
18. M. A. Adu, 'VENUE MANAGERS AND MEETING PLANNERS: A COMBINED PERSPECTIVE OF THEIR ROLES, RELATIONSHIPS, AND ATTRIBUTES NECESSARY FOR HOSTING A SUCCESSFUL MEETING', *J. UKnowledge*, 15(4), 123–128, 2018.
19. Koh and H. Greene, 'Green Event Marketing: The Sustainable Community Event Portfolio', *J. Interdisciplinary Bus. Stud.*, 2(8), 1–14, 2013.
20. L. Brooks, E. I. Brooks, and D. Jonathan, *Interactivity and Game Creation: 9th EAI International Conference, ArtsIT 2020, Aalborg, Denmark, December 10–11, 2020, Proceedings*, 1st ed., 1(1), Hung Kong SAR: Springer Nature, 2021.
21. J. Ugoani, 'Budget Management and Organizational Effectiveness in Nigeria', *Bus. Manag. Econ. Res.*, 5(2), 33–39, 2019.
22. J. Ugoani, 'Imperatives of Career Management and its Effect on Employee Performance', *Int. J. Soc. Sci. Perspect.*, 5(2), 47–56, 2019.
23. Capriello, L. Altinay, and A. Monti, 'Exploring resource procurement for community-based event organization in social enterprises: evidence from Piedmont, Italy', *Curr. Issues Tour.*, 22(19), 2319–2322, 2019.
24. H. Singh and R. R. Shah, 'BOOKiiIT - Designing a Venue Booking System (Technical Demo)', in *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, New Delhi, India, Sep. 2020, 1(1), 287–291.
25. W. H. Guilford and R. H. Schmedlen, 'Perspectives on Successfully Implementing BME Design Courses Online: Notes from an ASEE Workshop', *Biomed. Eng. Educ.*, 1(1), 145–148, 2021.
26. K. Meshram, D. Mate, A. Tighare, D. Wangal, and A. Lanjewar, 'Celebrations - Online Venue Booking Management System', *Int. J. Res. Eng. Sci. Manag.*, 2(3), 319–321, 2019.
27. JosephNg, P. S., & Eaw, H. (2021). Making Financial Sense from EaaS for MSE during Economic Uncertainty. *Future of Information and Communication Conference* (pp. 976–989). Vancouver, Canada: Springer Advances in Intelligent Systems and Computing.
28. JosephNg, P. S. (2021). Economic Turbulence and EaaS Grid Computing. *Int. J. of Business Forecasting and Marketing Intelligence*, 7(1), 33–52
29. JosephNg, P. S., & Eaw, H. C. (2022). Still Technology Acceptance Model? Reborn with Exostructure as a Service. *International Journal of Business Information Systems*, forthcoming.

30. JosephNg, P. S. (2019). EaaS Infrastructure Disruptor for MSE. *International Journal of Business Information System*, 30(3), 373-385.
31. JosephNg, P. S. (2018). EaaS Optimization: Available yet hidden information technology infrastructure inside the medium-size enterprise. *Technological Forecasting and Social Change*, 132(July), 165 - 173.
32. JosephNg, P. S., & Kang, C. M. (2016). Beyond Barebones Cloud Infrastructure Services: Stumbling Competitiveness During Economic Turbulence. *Journal of Science & Technology*, 24(1), 101-121.
33. JosephNg Poh Soon, Kang Chon Moy, Ahmad Kamil Mahmood, Wong See Wan, Phan Koo Yuen, Saw Seow Hui, Lim Jit Theam (2016), EaaS: Available yet Hidden Infrastructure inside MSE, 5th International Conference on Network, Communication, and Computing, Kyoto, Japan, ACM International Conference Proceeding Series, 17–20.
34. PS, J. N., Kang, C. M., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. (2016). Exostructure Services for Infrastructure Resources Optimization. *Journal of Telecommunication, Electronic and Computer Engineering*, 8(4), 65-69.
35. JNP Soon, WS Wan, PK Yuean, LE Heng & LJ Theam (2015). Barebone Cloud IaaS: Revitalisation Disruptive Technology, *International Journal of Business Information Systems*, 18(1), 107-126
36. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. H. (2014). IaaS cloud optimisation during economic turbulence for Malaysia small and medium enterprises. *International Journal of Business Information Systems*, 16(2), 196-208
37. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W. Phan, K. Y., & Lim, E. H. (2013). Battles in volatile information and communication technology landscape: The Malaysia small and medium-size enterprise case, *International Journal of Business Information Systems*, 16(2), 196-208
38. C. Li, J. Li, H. Cao, and Z. Meng, 'Design and Implementation of Online Booking System of University Sports Venues', *MATEC Web Conf.*, 100(1), 20–24, 2017.
39. D. Alkhaldi, D. Alkhaldi, H. Aldossary, U. A. Badawi, M. Alshabanah, and D. Alrajhi, 'Developing and Implementing Web-based Online University Facilities Reservation System', *Int. J. Appl. Eng. Res.*, 13(9), 6700–6708, 2018.
40. I. T. on S. D. Innovation (ITSDI) and D. I. U. R. MM M. T. I., *Cloud Computing and its role in Information Technology*, 2nd ed., 1(1). Madinah Saudi Arabia: IAIC Transactions on Sustainable Digital Innovation (ITSDI), 2021.

Evaluating an Automated Analysis Using Machine Learning and Natural Language Processing Approaches to Classify Computer Science Students' Reflective Writing



Huda Alrashidi, Nouf Almujally, Methaq Kadhum,
Thomas Daniel Ullmann, and Mike Joy

Abstract Reflection writing is a common practice in higher education. However, manual analysis of written reflections is time-consuming. This study presents an automated analysis of reflective writing to analyze reflective writing in CS education based on conceptual Reflective Writing Framework (RWF) and application of natural language processing and machine learning algorithm. This paper investigates two groups of features extraction (n-grams and PoS n-grams) and random forest (RF) algorithm that utilize such features to detect the presence or absence of the seven indicators (description of an experience, understandings, feelings, reasoning, perspective, new learning, and future action). The automated analysis of reflective writing is evaluated based on 74 CS student essays (1113 sentences) that are from the final year project reports in CS's students. Results showed the seven indicators can be reliably distinguished by their features and these indicators can be used in an automated reflective writing analysis for determining the level of students' reflective writing. Finally, we consider the implications of how the conceptualization of providing individualized learning support to students in order to help them develop reflective skills.

H. Alrashidi (✉)
Ministry of Education, Kuwait, Al Farwaniyah, Kuwait
e-mail: H.Alrashidi01@gmail.com

N. Almujally
Princess Nourah Bint Abdul Rahman University, Riyadh, Saudi Arabia
e-mail: Naalmujally@pnu.edu.sa

M. Kadhum
Ashur University College, Bagdad, Iraq
e-mail: Methaq.kadhum@au.edu.iq

T. Daniel Ullmann
Institute of Educational Technology, The Open University, Milton Keynes, UK
e-mail: T.ullmann@open.ac.uk

M. Joy
University of Warwick, Coventry, UK
e-mail: M.s.joy@warwick.ac.uk

Keywords Reflection assessment · Machine learning · Natural language processing · Reflective writing · Reflection · Computer science education

1 Introduction

Reflection has been used in higher education to support students to become thoughtful practitioners by enabling them to extract knowledge from their experiences [1–3] and can support metacognition [1]. The disciplines of higher education have been including reflection in some of their programs, for instance, teachers' pre-service training [4], medicine [5], management [6], and Computer Science (CS) [7–9]. Although the broad using reflection in higher education, students are lack of giving personalised feedback to assist them in developing reflective skills [10]. This is because the writing assessment nature that the time-consuming and labour intensive.

The focus of this research is on reflective writing in CS education. In terms of this, Fekete [9] stated that 'reflection is worth encouraging, for its indirect effect on the technical skills and knowledge which are our ultimate purpose in teaching Computer Science' (p. 144). Reflection improves students' awareness of how to learn from situations, e.g., how to deal with a sequence of steps required to reach a certain goal or how to identify the roots of a problem rather than concentrate on their feelings about the problem [11].

Despite the widespread use of reflective writing approaches for assessment, leading to the support of personalised learning [12], assessing reflective writing remains a challenge [13–17]. Reflection assessment is labour-intensive when manual content analysis [18] is applied to the task. Such assessment is employed to understand how students reflect and to support their reflective practice [10]. Of course, the fact that such assessment is so labour-intensive has led to the idea that automated approaches might potentially have a role in achieving it.

Automated assessment methods which assess writing based on evaluating its reflective content have generally used natural language processing (NLP) [12, 13, 15, 17, 19] to automate the utilization of a reflection framework. This research aims to explore the automated assessment of reflective writing for CS education. This field represents a significant challenge, due to the limited research which has taken place with respect to automated methods for analysing reflective writing.

Little research has been undertaken on the automated assessment of reflective writing. This present research aims to evaluate a system that undertakes the automatic assessment of reflective writing for CS education using advanced methods of natural language processing (NLP). This paper aims to (a) determine empirically each indicator of reflective writing features by examining different linguistic groups (unigram (word), bigram (word), trigram (word), unigram (PoS), bigram (PoS) and trigram (PoS)) and (b) build and evaluate a machine learning approach for binary classification for reflective writing in CS. The findings shed light on the structure of CS students' reflections and the first attempt to develop an automated reflective writing assessment using advanced NLP and machine learning techniques to allow

personalized reflection. The research question for this study is as follows: What are the linguistic features which can be found in CS students' reflective writings which indicate the presence of each of the reflection indicators?

2 Literature Review

2.1 *The Importance of Reflection in CS Education*

Reflection is commonly described as evidence of understanding of one's experiences of the situation used to take action in the future [2, 20]. In CS education, various activities necessitate the application of variations on the common reflection processes, such as judgment, evaluation, reasoning, problem-solving, and memorizing [11, 21–23].

Chng study indicated that it is necessary to teach problem-solving and reasoning skills in the course of CS education to improve students' awareness of how to learn from a situation they are presented with [11]. Hazzan & Tomayko also showed the importance of reflection in CS to support the student in the complexity in the development of software systems, which requires the developer to improve their understanding of their mental processes [23], and as this can be achieved by applying a reflection approach, it teaches developers how to think effectively. For these reasons, reflective writing is important in CS education.

2.2 *Methods to Analyze Reflection*

The reflective writing frameworks can develop different indicators in different fields, such as medical education or teacher preparation [24–26], and along similar lines [27–35].

There is variation in the ranges of characteristics covered by each indicator of each framework, with some frameworks tending to combine two or more indicators into one [36] or to divide what is generally one indicator into multiple sub-indicators [25, 35]. For example, Moallem's framework focuses on the writers/students' perspectives by using indicators such as explore; imagine alternatives; and gain exposure (to a variety of interpretive considerations in dialogue with others). On the other hand, Babb, Hoda [36] use only one of their indicators to represent perspective.

In terms of CS, Alrashidi, Ullmann [37] proposed seven indicators (description of an experience, understandings, feelings, reasoning, perspective, new learning, and future action) of the conceptual RWF and the framework was empirically evaluated [38]. These indicators are described in the following section.

The manual analysis of reflection is time-consuming [10]. Of course, the fact that such assessment is so labour-intensive has led to an interest in using advanced methods of analyzing reflective writing [12, 13, 15, 17].

2.3 *Automatic Method to Analyze the Reflection*

The existing approaches to automatic reflection analysis can be classified into keyword-based and machine learning-based approaches [12, 13, 39]. The keyword-based methods depend on locating specific keywords in the input text as indicators of reflection, using a keyword matching process. The presence/absence or frequency of the keywords can be used to analyse input text using the keyword-based approach [16, 40, 41]. Further, machine learning-based frameworks use existing classification algorithms to find patterns associated with each indicator at the training stage and then classify ‘unseen’ input texts using these mined patterns [12, 13, 15, 17].

Ullmann proposed a data-driven keyword-based technique for automatic reflective writing classification [12, 13, 39]. The datasets used were constructed from the British Academic Writing English (BAWE) corpus [12, 13, 15, 17]—from the health, engineering, and business fields. This majority voting system raised the reliability yielded by Cohen’s κ from 62 to 92%. The approach used for determining the keywords was based on log-likelihood, as discussed in Ullmann [42].

Lin et al. used a Chinese version of Linguistic Inquiry and Word Count (LIWC) [41], a text processing tool that characterizes the words that are utilized in association with different psychological and cognitive processes [43]—to identify the functions of the words used in reflective writings. The results showed that the words designated as indicators were the ones most frequently in the analyzed narrative.

Cui et al. proposed a framework, based on the LIWC list, for identifying the most important words and phrases in terms of identifying each indicator in the proposed reflection framework [16]. A total of 27 dental students’ reflections (using six reflective statement types) over four years were employed to identify the required features.

The LIWC dictionaries are not specific to one word that means a word can to more than one group, such as the word ‘died’ appears in several categories in past tense words, verbs, and death. Chung & Pennebaker pointed out that ‘NLP approaches will outperform LIWC on many classification tasks’ in comparison to machine learning algorithms [43]. A recent study by Liu et al. stated that ‘the use of only LIWC emotional features is insufficient to detect the depth of the Feeling Factor’ (p. 12) [44].

3 Methodology

3.1 *Dataset*

The dataset consists of sample texts that were collected from the projects of CS undergraduate students in the UK university. The dataset used in this study—of 174 different reflective writing documents—was employed for use by coders for the annotation. The data were collected from 174 third- and fourth-year CS student

projects. These had been undertaken in the course of the academic years 2013 through 2016 at the author's university and anonymized before analysis.

The unit of analysis was taken sentence of reflection as indicators can occur across sentences as evidence of reflection [45]. The dataset was coded by three coders until the reliability was stable at an acceptable agreement for three coders; using Cohen's κ for each of the seven binary indicators (presence or absent); as suggested by Landis and Koch [46]: poor agreement results are indicated by a Cohen's κ below 0, slight agreement results are in the range [0–0.2], fair agreement results are in the range [0.21–0.4], moderate agreement results are in the range [0.41–0.6], substantial agreement results are in the range [0.61–0.8] and almost perfect results are represented by a Cohen's κ of above 0.8. The agreement between coders was calculated as follows: when all the coders agreed on the same sentence this was considered agreement, but when even just one coder did not agree on a particular sentence, this was classified as disagreement.

Applying the conceptual RWF of 1113 sentences for the CS dataset, the agreement between coders was calculated as the seven indicators achieved kappa statistic values of between 0.46 to 0.75, and this range means moderate to the substantial agreement [46].

3.2 Manually Analysis of Reflection' Indicators (Content Analysis)

Reflection's indicators are used to assess the presence of each reflection's indicator based on the framework of Alrashidi et al. [37, 38]. The conceptual RWF is described in detail with text examples.

The Description of an experience indicator occurs when the writer describes the experience with no interpretation.

The understanding indicator is encountered when attempts are made to reach an understanding of a concept or topic and/or an understanding related to personal experience.

The Feelings indicator occurs when the writer has identified their thoughts, feelings, and/or behaviours.

The Reasoning indicator emerges when in-depth analysis is made which leads to a significant conclusion—i.e., a deeper understanding of the experience. The reasoning indicator signifies that the writer has made an effort to explain the experience in question.

The Perspective indicator occurs when the writer shows awareness of alternative perspectives. For example, from the CS dataset, where the students show awareness of their and/or others' perspectives.

The New Learning indicator occurs when the writer describes what they have learned from experience.

The Future action indicator suggests that the writer would, given the same circumstances again, intentionally do something differently or that they would plan their actions based on the new understanding that has resulted from considering and reviewing the original experience.

Applying the conceptual RWF of 1113 sentences for the CS dataset, the agreement between coders was calculated as the seven indicators achieved kappa statistic values of between 0.46 to 0.75, and this range means moderate to the substantial agreement [46].

3.3 Proposed Framework

The proposed automated RWF applying, the n-gram is used as it can encompass the features that are commonly used in NLP which are particularly relevant for the classification of the seven indicators. According to Jurafsky and Martin [47], the n-gram model supports the processing of important kinds of features commonly encountered in speech and language processing in general [48]. However, one of the limitations of the n-gram model is the fact that the method is ignorant of the grammatical nature of the text. Recently, a Part of Speech (PoS) n-gram model was used extensively for text classification [49, 50].

In Fig. 1, to extract the relevant features from the input text, after the text has been preprocessed, the text is tagged with its part-of-speech. Then, a set of features are

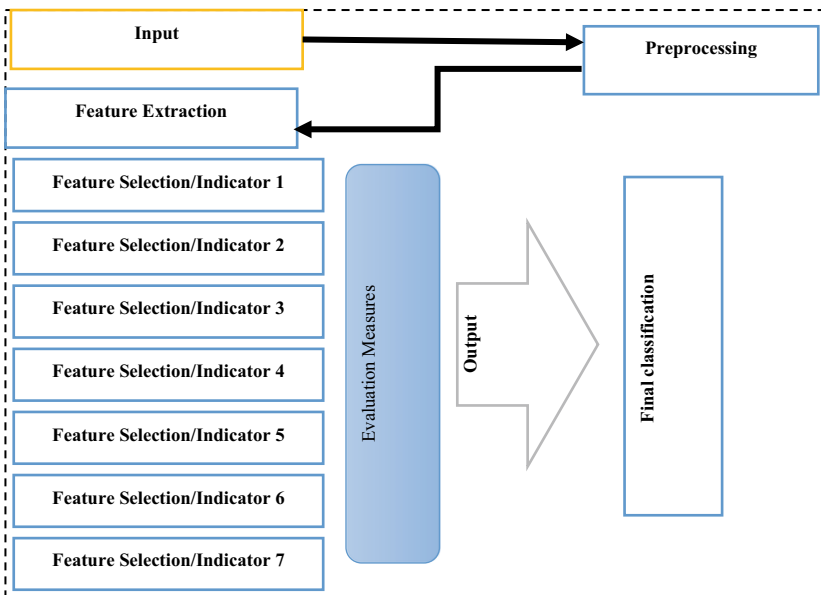


Fig. 1 The automated reflective writing framework

Table 1 The classification steps

| Component | Description |
|--------------|---|
| Inputs | Text, (coding the seven indicators for training-only) |
| Ground truth | Manual coding |
| Measurements | Accuracy and Cohen’s κ |
| Comparison | Performance of classifier (manual coding vs automated coding) |

extracted that allows for the analysis of the influence of these features on the indicator classification and hence the results of the automated RWF processing. Two groups of features are extracted, which are, n-grams (uni-gram, bi-gram and tri-gram) and PoS n-grams (uni-gram, bi-gram and tri-gram). In Table 1, illustrates the classification steps as input text until measuring the automated classification vs the manual coding.

Random forest is an ensemble classification technique that provides low-bias, low-variance performance. It also allows for feature inspection by building multiple decision trees using random subsets of features on bootstrapped samples [51]. Previous research has employed the RF algorithm as the best available for classification—as compared to the rest of the algorithms [12, 15, 17]. This study applies the RF algorithm.

Cohen’s κ is often used concerning the manual annotation in the educational area to measure IRR (as between human assessors), while the F-measure and accuracy are the most common measures used in automatic reflective writing assessment systems [17]—in order to assess the performance of the machine learning algorithms applied.

$$Accuracy = \frac{truePositive + trueNegative}{truePositive + trueNegative + falsePositive + falseNegative} \quad (1)$$

$$F - measure = \frac{2 * precision * recall}{precision + recall} \quad (2)$$

$$Cohen's \kappa = \frac{relativeObservedAgreement - hypotheticalProbabilityOfChanceAgreement}{1 - hypotheticalProbabilityOfChanceAgreement} \quad (3)$$

In this study the automated RWF, a binary classification process is implemented for each of the predetermined indicators. For each indicator, the inputs are classified into the classes, 0 or 1 (absence/presence of the indicator in the text) in the breadth-based form of the automated RWF task. For each input text, a feature vector based on the extracted and reduced features is created.

4 Results

Feature selection produces a set of features that can, potentially, be selected—one for each indicator. This selected set of features is sub-grouped based on the feature type. These are unigram (word), bigram (word), trigram (word), unigram (PoS), bigram (PoS) and trigram (PoS).

4.1 *Description of an Experience*

The description of an experience indicator is often positively associated with some PoS type features (unigram and bigrams); Using analysis of the dataset, it was discovered that the first person pronouns ‘I’, ‘my’, ‘me’, and the third-person pronoun ‘it’ have a positive effect in terms of this indicator. This means that the presence of first-person pronouns and third-person pronouns is associated with inputs that include the text of a purely descriptive nature. Similarly, Birney [52] and Ullmann [12] reported that the first and third-person pronouns can be found in the description of an experience category.

4.2 *Understanding*

In the dataset, the word ‘expect’ was appeared only in the non-understanding examples (i.e., text items that were as not complying with the understanding indicator by the expert), such as ‘this ended up being a lot of work which I was not expecting due to the number of modules’, ‘...would work as expected and performed the required role’. And ‘I learned how to deal with multiple deadlines in a way where you can achieve the best expected results’. Besides, the understanding indicator is characterized negatively by phrases such as ‘it would’, ‘for example’, and ‘gained further knowledge’. In the dataset, there are text items involving such phrases which can be characterized as non-understanding, such as, ‘I also gained further knowledge in people skills; due to the situation’, and ‘For example, instead of a plain jar file’.

4.3 *Feelings*

Using analysis of the dataset, it was discovered that the bigrams of words (there being 353 of these) form the majority of the complete list of features. In particular, there are the bigrams which include the sensing and thinking verbs that refer to mental processes, e.g., ‘I think’, ‘I feel’, ‘I believe’, ‘I find’, ‘I assume’ and ‘I realise’—in their present tense and past tense forms. There are many instances of words/bigrams

in the dataset which referred to negative feelings: such as ‘struggle’, ‘difficulty’, ‘negative’, ‘unfortunately’, ‘quite difficult’, and ‘conflict’. However, there were also quite a few words/phrases which referred to positive feelings: such as ‘satisfy’, ‘excellent’, ‘good’, ‘appreciation’, ‘happy’, and ‘proud’. Such intuitions, which are often encapsulated in expressions linked to feeling or thinking, can be a justification for considering that a particular text is reflecting on something in order to gain greater clarity.

Supporting evidence for the nature of these features as described above can be found in Birney [52], Ullmann [12], Ryan [53] and Cui et al. [16]. Feelings are linked to the use of pronouns: singular (I, my, me), plural (we) (as stated by Ullmann [12]) and sensing and thinking verbs (as stated by Birney [52] and Ryan [53]).

4.4 Reasoning

From the dataset, it was discovered that different bigrams of words also included in these selected features have different effects: for example, the bigram ‘because of’ has a negative effect in terms of identifying input with a reasoning label, while the bigram ‘role within’ has a positive impact. In the dataset, there are some items which include the phrase ‘role within’ that are classified positively concerning the reasoning label, such as ‘adopting a different role within the group would have led to a fundamentally different position for myself’. Other items (other than the one given above) which include the phrase ‘because of’ are classified as negative for the reasoning indicator, such as ‘...this sprint cycle was hard to adhere to because of time being focused.....’.

The best indicators were words unigram and bigrams specifying the presence of causal links, such as ‘hence’, ‘result’, ‘due to’, ‘the fact’, ‘as this’, ‘such as’, ‘as my’, ‘and hence’, ‘this cause’, and ‘result of’. These words and phrases evidenced that students were using causal links in their reflective writing, so adding explanation (Birney [52]; Ryan [53]).

4.5 Perspective

From the dataset, the word ‘question’ is a feature that was selected concerning the perspective indicator. It has a positive influence in terms of identifying input text items that comply with the perspective indicator. This is the case also with the word ‘would’ and some phrases involving this: ‘would also’, and ‘would have’, among others. Adverbs in general also have a positive influence. In the dataset, there are examples of the use of such constructs as, ‘If I could do anything differently, I would have chosen a clearer project title sooner,’, ‘....Given my intense interest in this topic, I really hope I have the opportunity to go back and answer those questions’,

and ‘Being able to discern helpful materials from the outset would have saved a considerable amount of time throughout the project’.

Expressions of ability may use modal verbs and phrases. The students are usually using the adjective ‘able’ to express their ability to do something.

Finally, this pattern of findings (as discussed above) is consistent, in technological terms, with Ullmann [12] results, and in theoretical terms, with Birney [52]; Ryan [53]. Ullmann [12], in particular, found that the automated detection of reflective writing concerning the perspective indicator showed evidence of the use of first- and third-person pronouns.

4.6 Future Action

From the dataset, auxiliary verbs such as ‘would’ can be used to talk about the past, and about the future in the past, or something desired but not actual at present, or about an imagined situation in general. Also, the word ‘will’ can be used to describe something that is to take place in the future, though it can refer to the past or future depending on the context.

In the dataset, the texts ‘In the future, I would definitely take a more proactive role in ensuring the health of the group as a whole’ and ‘If I were to complete this research project again’ can have a positive future action label assigned to them based on the words and phrases.

These findings relating to the future action indicator are aligned with those of Birney [52], Ullmann [12], Cui et al. [16] and Jung and Wise [17]. These all found that the use of the future tense and first-person pronouns evidence a consideration, by the student, of their future actions as moderated/modified by the lessons gained from the experience being described.

4.7 New Learning

From the dataset, The verb ‘learn’ can be used to refer to what the student has learned from an experience. However, the word was found in different tenses and parts of speech forms, such as ‘I have learned’, ‘the learning outcome’, ‘managed to learn’, ‘author has learned’, ‘lesson I learned’ ‘was learning through’, and ‘learned about java’. Additionally, the phrases, ‘teach me’ and ‘taught me’ can be used to refer to gain in knowledge.

The first-person singular pronouns, ‘I’, ‘my’, and ‘me’ had a positive influence on the detection of the learning indicator, which means that the presence of the first-person pronouns is associated with inputs that can be labelled with the learning indicator. Similarly, Birney [52], Ullmann [12] and Jung and Wise [17] found empirical evidence that the use of first-person pronouns can provide evidence of learning.

Table 2 Performance of random forest algorithm for each indicator

| Indicator | Accuracy | Kappa |
|------------------------------|----------|-------|
| Description of an experience | 0.80 | 0.43 |
| Understanding | 0.84 | 0.17 |
| Feelings | 0.75 | 0.51 |
| Reasoning | 0.75 | 0.51 |
| Perspective | 0.85 | 0.35 |
| Future action | 0.96 | 0.53 |
| New learning | 0.93 | 0.67 |

5 Evaluation Results

As given in Table 2, using random forest can improve the performance of n-grams and POS-n-grams features that showed straightforwardly to be effective in various text classification models. The classification model for reflective indicators showed slight to substantial performance ($\kappa = 0.17-0.67$). Future action and New learning classifiers had the highest accuracy (0.96 and 0.93) and the best performance ($\kappa = 0.67$ and 0.53), while Feelings and Reasoning classifiers performed moderate kappa by 0.51 aligned with the accuracy performance (≥ 0.75). The Perspective and Understanding classifiers had a high accuracy above 0.80, but the kappa had slight to fair (0.17 and 0.35). Accordingly, combining different types of features in the proposed framework improves the results. Different features need to experiment with a different field and with different inputs.

6 Discussion

RQ: What are the linguistic features which can be found in CS students' reflective writings which indicate the presence of each of the finding's reflection indicators?

The two types of linguistic features have been used for predicting/detecting the reflection indicators. These features were shown to be effective in predicting the presence or absence of the pre-determined indicators (i.e., description of an experience, understanding, feelings, reasoning, perspective, new learning, and future action).

Compared to the state-of-the-art, our automated RWF captured a wide range of features and obtained good results using features that had not previously been tested in relation to any similar task. Kovanović et al. [15] used three types of features, n-grams, LIWC, and Coh-matrix Ullmann [12] used uni-grams only while Jung and Wise [17] used features that were extracted based on LIWC only.

It is worth noting that our findings relating to the indicators are well aligned with the theoretical and technical literature concerning reflective writing. These findings have shown that many different linguistic features are useful for detecting indicators and that therefore analyzing these features may lead to a greater understanding of

the nature of the various levels of reflective writing and their characteristics. And these findings/features were indeed examined here in order to demonstrate the kind of value this kind of study can have in terms of highlighting potential areas for future investigation.

The features identified by the system for each indicator are generally different from those for any other indicator, but there is also some overlap of linguistic resources and terms with respect to the indicators. For example, the singular first-person pronoun is one of the top features for all the indicators. It was also found that verbs such as thinking and sensing were important for several indicators. The findings relating to first-person pronouns and the thinking and sensing words were investigated empirically as important features of reflection; this investigation was based on the theoretical literature of Birney [52], and the empirical evaluation incorporated in the reflective writing and technical literature analyses of Ullmann [12] and Jung and Wise [17]. These findings with respect to the thinking and sensing words and phrases suggest that such can be used to foster students' reflections which are conditional upon and require thinking [54]. Regarding the description of an experience indicator, the first-person pronoun and the third person pronoun are important for identifying this indicator. These pronouns can be used to describe who was involved in the experience.

Further, both the understanding and the feelings indicators exhibit similarities in terms of the use of linguistic features such as the subordinating conjunction, 'that', the auxiliary verbs (to be, have), adjectives, the past tense, and the thinking and sensing words (understand, think, feel, believe, relies on, etc.). These findings align theoretically with both Birney [52] and Ryan [53] who focused on manual analysis, and technically (in terms of features for machine learning) with Kovanović et al. [15], Ullmann [12], and Jung and Wise [17].

The use of n-grams and PoS (n-grams) showed great potential for making intuitive sense of reflective writings. These kinds of feature are linguistically based and so refer to well-established linguistics concepts. The features we mostly used were ones that have been widely seen in literature: the first-person pronoun, the sensing and thinking verbs and phrases (e.g., believe, see, feel), the causal links and phrases (i.e., because of, as a result, due to), and future and past words and phrases. The use of PoS tagging allowed differentiation of reflection indicators via the consideration of the syntactic relationships between the words and phrases in the sentences [14, 15, 17].

Random forest models showed a good performance in the majority of the indicator as it was outperformed other approaches of reflection classification tests, according to earlier research Kovanović et al. [15], Ullmann [12] and Cui et al. [16].

7 Conclusion

This paper makes two significant contributions. Firstly, we developed the automated reflective writing classification for students. The classification model for reflective indicators of random forest reached an accuracy of 0.96 and substantial performance ($\kappa = 0.67$), which is regarded as a moderate level of agreement. The application of

n-grams and PoS n-grams features shows considerable potential for understanding students' reflective writings, which are constructed using well-established linguistic for psychological processes.

Secondly, our study provides an evaluation of the linguistic features of the seven reflection indicators based on Alrashidi et al., [37, 38]. The features we mostly used were ones that have been widely seen in literature: the first-person pronoun, the sensing and thinking verbs and phrases (e.g., believe, see, feel), the causal links and phrases (i.e., because of, as a result, due to), and future and past words and phrases. The use of PoS tagging allowed differentiation of reflection indicators via the consideration of the syntactic relationships between the words and phrases in the sentences.

Lastly, our findings demonstrated some advantages of using the reflection in a specific context that captured a different kind of linguistic features for each indicator for CS students reflective writing. In the future, we will focus on examining more linguistic features in order to capture a specific semantic for each reflection indicator. We also will focus on developing our tool by using advanced techniques such as data mining.

References

1. Schön DA (1987) Educating the reflective practitioner: toward a new design for teaching and learning in the professions. Jossey-Bass
2. Boud D, Keogh R, Walker D (1985) Reflection: turning experience into learning. Routledge
3. Moon JA (1999) Learning journals: a handbook for academics, students and professional development. Routledge
4. Cohen-Sayag E, Fischl D (2012) Reflective writing in Pre-service teachers' teaching: what does it promote? *Aust J Teach Educ* 37(10):2
5. Wald HS et al (2012) Fostering and evaluating reflective capacity in medical education: developing the REFLECT rubric for assessing reflective writing. *Acad Med* 87(1):41–50
6. Betts J (2004) Theology, therapy or picket line? What's the 'good' of reflective practice in management education? *Reflective Pract* 5(2):239–251
7. George SE (2002) Learning and the reflective journal in computer science. *Austral Comput Sci Comm* 24(1):77–86
8. Demmans Epp C, Akcayir G, Phirangee K (2019) Think twice: exploring the effect of reflective practices with peer review on reflective writing and writing quality in computer-science education. *Reflective Pract* 20(4):533–547
9. Fekete A et al (2000) Supporting reflection in introductory computer science. *ACM SIGCSE Bull* 32(1):144–148
10. Koole S et al (2011) Factors confounding the assessment of reflection: a critical review. *BMC Med Educ* 11(1):104
11. Chng SI (2018) Incorporating reflection into computing classes: models and challenges. *Reflective Pract* 19(3):358–375
12. Ullmann TD (2019) Automated analysis of reflection in writing: validating machine learning approaches. *Int J Artif Intell Educ* 29(2):217–257
13. Ullmann TD (2015) Automated detection of reflection in texts. A machine learning based approach. The Open University
14. Gibson A et al (2017) Reflective writing analytics for actionable feedback. In: Proceedings of the seventh international learning analytics & knowledge conference. ACM

15. Kovanović V et al (2018) Understand students' self-reflections through learning analytics. In: Proceedings of the 8th international conference on learning analytics and knowledge
16. Cui Y, Wise AF, Allen KL (2019) Developing reflection analytics for health professions education: a multi-dimensional framework to align critical concepts with data features. *Comput Hum Behav* 100:305–324
17. Jung Y, Wise AF (2020) How and how well do students reflect? multi-dimensional automated reflection assessment in health professions education. In: Proceedings of the tenth international conference on learning analytics & knowledge
18. Krippendorff KH (2013) Content analysis—3rd edition: an introduction to its methodology. SAGE Publications Inc., Thousand Oaks
19. Liu M et al (2019) Evaluating machine learning approaches to classify pharmacy students' reflective statements. In: International conference on artificial intelligence in education. Springer
20. Dewey J (1933) A restatement of the relation of reflective thinking to the educative process. DC Heath
21. Schraw G, Dennison RS (1994) Assessing metacognitive awareness. *Contemp Educ Psychol* 19(4):460–475
22. Stone JA, Madigan EM (2007) Integrating reflective writing in CS/IS. *ACM SIGCSE Bull* 39(2):42–45
23. Hazzan O, Tomayko JE (2005) Reflection and abstraction in learning software engineering's human aspects. *Computer* 38(6):39–45
24. Surbeck E, Han EP, Moyer J (1991) Assessing reflective responses in journals. *Educ Leadersh* 48(6):25–27
25. Moallem M (1998) Reflection as a means of developing expertise in problem solving, decision making, and complex thinking of designers
26. Mamede S, Schmidt HG (2004) The structure of reflective practice in medicine. *Med Educ* 38(12):1302–1308
27. Plack MM, Greenberg L (2005) The reflective practitioner: reaching for excellence in practice. *Pediatrics* 116(6):1546–1552
28. Greiman B, Covington H (2007) Reflective thinking and journal writing: examining student teachers' perceptions of preferred reflective modality, journal writing outcomes, and journal structure. *Career Tech Educ Res* 32(2):115–139
29. Crawford PA, Roberts SK, Hickmann R (2010) Nurturing early childhood teachers as leaders: long-term professional development. *Dimens Early Child* 38(3):31
30. Black PE, Plowright D (2010) A multi-dimensional model of reflective learning for professional development. *Reflective Pract* 11(2):245–258
31. Kwon K, Jonassen DH (2011) The influence of reflective self-explanations on problem-solving performance. *J Educ Comput Res* 44(3):247–263
32. Ellmers G (2015) The graphic design project: employing structured and critical reflection to guide student learning. *Commun Des* 3(1):62–79
33. Amador JM et al (2019) Preparing preservice teachers to become self-reflective of their technology integration practices. In: Pre-service and In-service teacher education: concepts, methodologies, tools, and applications, pp 1298–1325. IGI Global
34. Antonio RP (2020) Developing students' reflective thinking skills in a metacognitive and argument-driven learning environment. *Int J Res Educ Sci* 6(3):467–483
35. Noer S, Gunowibowo P, Triana M (2020) Development of guided discovery learning to improve students reflective thinking ability and self learning. In: Journal of physics: conference series. IOP Publishing
36. Babb J, Hoda R, Nørbjerg J (2014) Embedding reflection and learning into agile software development. *IEEE Softw* 31(4):51–57
37. Alrashidi H et al (2020) A framework for assessing reflective writing produced within the context of computer science education. In: Proceedings of the 10th international learning analytics & knowledge conference (LAK20). Frankfurt, Germany

38. Alrashidi H et al (2020) Educators' validation on a reflective writing framework (RWF) for assessing reflective writing in computer science education. Springer International Publishing, Cham
39. Ullmann TD (2017) Reflective writing analytics: empirically determined keywords of written reflection. In: Proceedings of the seventh international learning analytics & knowledge conference. ACM
40. Ullmann TD (2011) An architecture for the automated detection of textual indicators of reflection. 1st European workshop on awareness and reflection in learning networks. Palermo, Italy, pp 138–151
41. Lin C-W et al (2016) A word-count approach to analyze linguistic patterns in the reflective writings of medical students. *Med Educ Online* 21(1):29522
42. Ullmann TD (2015) Keywords of written reflection—a comparison between reflective and descriptive datasets. In: CEUR workshop proceedings
43. Pennebaker JW, Chung CK (2011) Expressive writing: connections to physical and mental health. *Oxf Handb Health Psychol* 417–437
44. Liu M, Kitto K, Buckingham Shum S (2021) Combining factor analysis with writing analytics for the formative assessment of written reflection. *Comput Hum Behav* 120: 106733
45. Moon JA (2013) Reflection in learning and professional development: theory and practice. Routledge
46. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 159–174
47. Jurafsky D, Martin JH (2008) Speech and language processing: an introduction to speech recognition, computational linguistics and natural language processing. Prentice Hall, Upper Saddle River, NJ
48. Ogada K, Mwangi W, Cheruiyot W (2015) N-gram based text categorization method for improved data mining. *J Inf Eng Appl* 5(8):35–43
49. Schulman A, Barbosa S (2018) Text genre classification using only parts of speech. In: 2018 International conference on computational science and computational intelligence (CSCI). IEEE
50. Fang AC, Cao J (2015) Part-of-speech tags and ICE text classification. Text genres and registers: the computation of linguistic features. Springer, pp 71–82
51. Breiman L (2017) Classification and regression trees. Routledge
52. Birney R (2012) Reflective writing: quantitative assessment and identification of linguistic features. Waterford Institute of Technology
53. Ryan M (2011) Improving reflective writing in higher education: a social semiotic perspective. *Teach High Educ* 16(1):99–111
54. Boud D, Keogh R, Walker D (2013) Reflection: turning experience into learning. Routledge

Secured and Lightweight Key Distribution Mechanism for Wireless Sensor Networks



P. Ezhil Roja and D. S. Misbha

Abstract Because of the enormous technological revolution, the most difficult tasks that humans cannot perform are now easily addressed by small digital gadgets. An example of this technological revolution is Wireless Sensor Networks (WSN). In the WSN, numerous heterogeneous types of sensors work together to sense, monitor, capture, process, and control the physical condition of a specific region or object. The WSN has been used extensively in a number of key areas, particularly in the medical, agricultural, military, and traffic management sectors. The little electronic gadget that forms the WSN is composed of three major components: a sensing component, a power storage component, and a wireless transmission component. Despite the numerous benefits of this modern technology, security remains an unresolved concern. This work aims to develop an efficient Lightweight Secure Key Distribution Protocol (LSKDP) for enhancing the process's safety in heterogeneous medical WSNs. Additionally, security and performance analyses were conducted to demonstrate the suggested methodology's efficacy. The security analysis is performed to ascertain the proposed method's vulnerability to network threats. Additionally, performance analysis is used to determine the CPU utilisation, memory consumption, processing power, and computational overhead of the suggested methodology. The experimental results indicate that the suggested strategy detects attacks more accurately and with a lower computational overhead.

Keywords WSN · Key distribution · Key agreement · Security · Authentication · Attack detection

P. Ezhil Roja (✉)

Department of Computer Science, Nesamony Memorial Christian College, Manonmaniam Sundaranar University, Marthandam, India
e-mail: roja_z@yahoo.com

D. S. Misbha

Department of Computer Application, Nesamony Memorial Christian College, Marthandam, India
e-mail: misbhasatheesh4@gmail.com

1 Introduction

The largest network design interconnected by small sensors is the wireless sensor network. Due to its adaptability and efficiency, WSN is employed in a wide variety of real-time applications, including healthcare, environment monitoring, military and other applications [1–3]. WSN is divided into three layers: sensors, clusters or signs, and monitoring and control. The power, compute, storage, and communication capabilities of WSN sensor nodes are limited. Limitation of resources the wireless communication medium poses various obstacles, the most critical of which is security [4–8]. Hence, wireless sensor networks security concerns must be addressed. The purpose of this research is to design an efficient chronographic key distribution module capable of analysing and resolving the security and privacy issues that arise when resource-constrained sensor nodes are used. The overall design of the WSN is depicted in Fig. 1.

As illustrated in Fig. 1, cryptographic measures should be in place to ensure the secure communication of connected parties within a WSN.

The primary objective of this research is to create a significantly lighter security solution for these small digital sensors, hence improving WSN overall security.

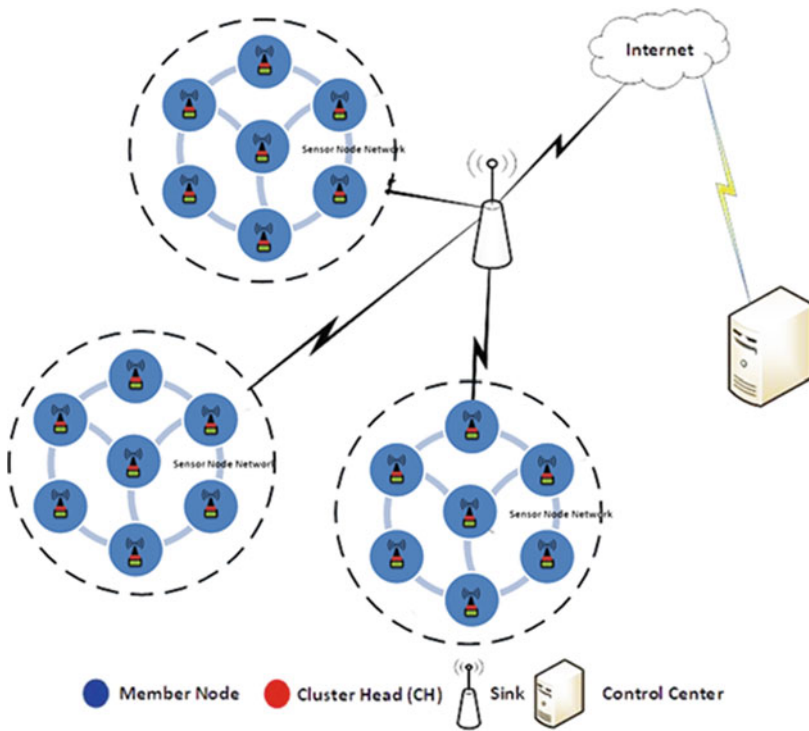


Fig. 1 The secure data transmission model for WSN

Thus, an effective WSN clustering strategy is presented to reduce the sensor nodes' communication and key storage overhead. Additionally, the proposed protocol incorporates additional security variables and key distribution algorithms to guard against network attacks. Finally, a comparison to existing approaches has been performed to demonstrate this key distribution protocol's computational efficiency.

Section 2 discusses the benefits and drawbacks of recently developed key distribution systems. Section 3 details the suggested key distribution protocols' core modules. Section 4 compares the newly developed key distribution protocols' to the proposed protocol using experimental analysis. Finally, the proposed approach is concluded.

2 Literature Review

This section summarises related research on secure key distribution and communication techniques in wireless sensor networks. Unfortunately, only a few studies have addressed critical key distribution concerns in WSNs, and the remaining research area is still in its early stages.

Huang et al. [9], recommended a method by which keys can be distributed over the wireless sensor networks. In the proposed method key is sent through handshaking authentication. In order to minimize the communication cost, physical unclonable functions are stored in the sensor nodes. In this method, the concurrent key is generated on every sensor node by running the PUF module, in the meantime, an optimal path to transfer the factors for producing a group key. Moreover, the method is resistant towards cloning and eavesdropping, thus providing security in the process of transfer of group keys.

Harn et al. [10], presented a method for initiating a class of keys in wireless sensor networks (WSNs) and its transfer. The method uses a multivariate polynomial and some sensors for building the polynomial. Usage of multi-variant polynomial minimizes the storage cost to be linearly proportional to the size of group communications. The method has no communication overhead. Concerning security, the method is efficient toward computational complexity.

Camtepe et al. [11], proposed a hybrid method for specifying the number and the keys to be assigned to each key-chain as to communicate over wireless sensor networks requires pairwise keys to the sensor nodes. To perform key distribution efficiently, mapping of Balanced Incomplete Block Designs (BIBD) and Generalized Quadrangles (GQ) is done. The method is tested in terms of its performance and security, analytically and computationally.

Bechkit et al. [12], recommended a better way for managing keys in wireless sensor networks by means of unital design theory. Wireless sensor networks provide better and protected connectivity coverage. Mapping from units to pre key distribution provides high network accessibility. The method improves network accessibility while providing better key sharing techniques. The work is tested over with network scalability, network connectivity, and storage and provides better results.

Li et al. [13], recommended a secured random key distribution scheme that protects unauthorized access. As the demand for networks increases, the need for security and privacy is also in almost demand. The proposed scheme SRKD identifies and provides revocation of the unauthorized attacks by combining the localized algorithm with a voting mechanism. Then to eliminate the duplication attack, the grammatical meaning of the attributes are changed. If the number of shared starting keys is larger than the maximum number of starting keys, then the connection is not established.

Shahzad et al. [14], proposed a method for increasing the network's life based on the pre-deterministic key-distribution based CCEF (PKCCEF). The method uses the security protocol commutative cypher en-route filtering (CCEF), for saving energy by performing filtering at the design stage. The security protocols used determines the shortest path for routing. In case of more attacks, a protective path is selected to identify more attacks. Furthermore, for the smaller number of FTR energy-efficient path is selected. The method does not contain a buffer.

McCusker et al. [15], recommended a method for sending a key over the wireless sensor network based on Identity Based Cryptography (IBC). The method enhances the resistivity of the node in the wireless sensor network by maintaining a list of the devices in the radio range. The Tate pairing component is shifted to hardware due to the consumption of a large amount of energy. Parameters such as area, timing and energy are calculated using an accelerator, Complementary Metal Oxide Silicon (CMOS).

The majority of these existing studies have several drawbacks, including probabilistic key distribution between heterogeneous sensors nodes and simple sensor nodes, non-scalability after implementation, increasing vulnerability to node compromise, memory and storage constraints on resource-constrained LSNs, and a high bandwidth consumption.

3 Proposed Methodology

This section discusses the essential security components of the proposed WSN key distribution mechanism in detail.

3.1 Network Architecture

Typically, a WSN infrastructure comprises three types of sensor nodes: low-resource-capability sensors nodes (LSNs), high-resource-capability sensors nodes (HSN's), and base stations (BS). The base station is a supercomputer with extremely high computing capacity (GPU processors, high bandwidth, high processing power, and vast memory) that is capable of simultaneously executing any complex security application.

3.2 Cluster Formation

Sensor nodes clustering is a vital part of WSN. By properly clustering the sensor nodes in the WSN, security and energy can be appropriately managed. In this proposed research, the clustering of sensor nodes is first divided according to the types of sensors. For example, temperature, humidity, and rain sensors are separated in the weather monitoring system, and clustering is done according to the type of sensors. The cluster formation model of the proposed research is depicted in Fig. 2. This protocol is designed to communicate CH directly to each LSN in order to minimise the computational overhead associated with key distribution. Therefore, each cluster is star-shaped, the CH is located in the cluster’s centre, and the LSN communicates directly with the CH without the assistance of other LSNs in the cluster.

Each cluster has an n number of LSNs. These LSNs are intended for use in communication with the cluster’s head (CH). The network is constructed with one CH for every twenty LSNs to avoid the LSNs’ computational overhead and malicious attacks. The following formula 1 determines the number of CHs in a cluster.

$$T_{CH} = \frac{T_{LSN}}{20} \tag{1}$$

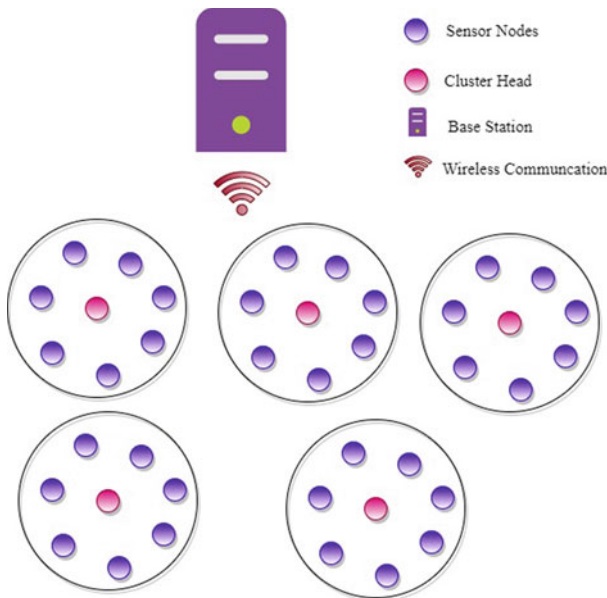


Fig. 2 Cluster formation of LSKDP

Formula 1 the variable T_{CH} refers to the total number of CHs in a cluster, whereas T_{LSN} refers to the total number of LSNs in a cluster. Data acquired by LSNs is wirelessly transferred to BS via CHS.

3.3 Key Distribution

The proposed key distribution protocol is divided into five stages: key generation and registration, key pre-distribution, post key distribution, secure communication LSN to BS and Secure communication BS to LSN. Table 1 shows the notations used to represent the proposed key distribution methodology.

3.3.1 Key Generation and Registration

This research employed symmetrical key payers. For that, the most widely used public-key cryptosystem, RSA, is utilised for key generation. Rivest, Shamir, and Adleman created the RSA cryptosystem in 1977 to facilitate secure data transmission. RSA's security is predicated on the difficulties associated with factoring the product of two independent and large prime numbers. The process of key generation is based on formulas 2, 3, and 4.

$$\{P_K, S_K\} \leftarrow RSA_{Kgen} \quad (2)$$

Table 1 The notations used in this research

| Notations | Descriptions |
|-----------|---------------------------------------|
| LSN | Low-resource-capability sensors nodes |
| BS | Base station |
| CH | Cluster head |
| P_K | Primary key |
| S_K | Secret key |
| Dis | Key distribution |
| E | Encryption |
| D | Decryption |
| S_{Id} | Sensor identification number |
| CH_{Id} | Cluster identification number |
| T | Timestamp |
| τ | Threshold limit |
| P_T | Plain text |
| C_T | Chipper text |

$$P_K \leftarrow RSA_{K_{gen}}\{0, 1\}^{128}, S_K \leftarrow RSA_{K_{gen}}\{0, 1\}^{128} \tag{3}$$

$$P\{P_K, S_K\} = \frac{1}{|\{0, 1\}^{128}|} \tag{4}$$

where P_K denotes the public key and S_K denotes the private key. Two 128 bit secret key payers have been generated in this research utilising formula 2. Formula 3 is used to determine its probability values. Before LSNs can be deployed in a WSN, BS must generate these key pairs. Additionally, BS generates S_{Id} , a unique identifier for all LSNs, to facilitate LSN identification. Finally, the number of CHs is calculated by formula 1 based on the total number of LSNs, and BS generate the unique identification number for each CH. The key payers and sensor ids produced by BS are securely maintained in the BS database to facilitate secure communication.

3.3.2 Key Pre-distribution

In key pre-distribution, private key and sensor id randomly selected from key payers generated by BS is deployed in LSN. This is accomplished by the use of formula 5. Second, the cluster head stores the randomly generated private key and CH id generated by BS for secure communication. This is accomplished with the help of formula 6. All of this occurs prior to the placement of LSN and CH on the WSN network. The pre-distribution of security keys and registration information greatly minimises the communication overhead of the WSN.

$$LSN_i \underline{\text{def}}_{BS} [Dis[S_{Id}, P_K]] \tag{5}$$

$$CH_i \underline{\text{def}}_{BS} [Dis[CH_{Id}, P_K]] \tag{6}$$

3.3.3 Post Key Distribution

In post key distribution, the sensors send a secret key request to the base station to decrypt the messages $\Rightarrow^{Send_{S_K KeyReq}} E[CH_{Id}, S_{Id}] \rightarrow$. As soon as a request from the sensors is received, BS verifies the authenticity of the received message. If the authentication is successful, the sensor identification number is used to locate the secret key stored in the BS database ($S_{Id}(S_K)$). To strengthen the security, the hashed random number and time stamp are encapsulated in a secret key $Tag \leftarrow H(R_{SN}), T \leftarrow Comp(Stamp)$. Finally, security information is securely delivered to the specified sensor using the sensor id and cluster id $\leftarrow E_{P_K}[CH_{Id}, S_{Id}, T, Tag, S_K] \Rightarrow^{Sendsensornode}$. The process of LSKDP's post-key distribution is detailed in Table 2.

Table 2 Post key distribution

| Sensor nodes (LSN) | Base station BS |
|---|---|
| $\Rightarrow^{SendS_K KeyReq} E[CH_{Id}, S_{Id}]$ | |
| | $\Leftarrow^{RecvS_K KeyRequest} E[CH_{Id}, S_{Id}]$ $D[CH_{Id}, S_{Id}]$ $F(S_{Id}(S_K)), Tag \leftarrow H(R_{SN}), T \leftarrow Comp(Timestamp)$ $E_{P_K}[CH_{Id}, S_{Id}, T, Tag, S_K]$ $\leftarrow E_{P_K}[CH_{Id}, S_{Id}, T, Tag, S_K]$ $\Rightarrow^{Sendsensornode}$ |
| $\Leftarrow^{RecvS_K Key} E_{P_K}[CH_{Id}, S_{Id}, T, Tag, S_K]$ if($T > \tau$) { $D_{P_K}[CH_{Id}, S_{Id}, T, Tag, S_K]$ } | |

Table 3 Secure communication LSN to BS

| Sensor nodes (LSN) | Base station BS |
|---|---|
| $CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]]$ $\Rightarrow^{SendtoBS} CP_T \rightarrow$ | |
| | $\Leftarrow^{RecvCP_T fromLSN} CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]]$ if($T > \tau$) { $F(S_{Id}(S_K))$ $PT \leftarrow D_{S_K}[CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]]]$ } |

3.3.4 Secure Communication LSN to BS

This section discusses how to safely share the sensor nodes’ chipper text to the base station. The generated text is encrypted using the primacy key and the sensor node’s hashed random number and sent to the base station; the encrypted message includes the sensor node’s sensor id S_{Id} and cluster head id CH_{Id} details $E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]] \Rightarrow^{SendtoBS} CP_T \rightarrow$. When the base station receives the chipper text, the message is instantly rejected if the threshold is exceeded. Otherwise, it is decrypted into plain text using the sensor node’s secret key and hashed value. Table 3 details the secure connection between the LSN and the BS of LSKDP.

3.3.5 Secure Communication BS to LSN

To deliver chipper text securely to the sensor node, the base station encrypts plain text using the primary key and hashed value. Further to prevent reply attacks, the time stamp is computed and encapsulated in encrypted chipper text $CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]]$. This message is securely transmitted to

Table 4 Secure communication BS to LSN

| Base station BS | Sensor nodes (LSN) |
|---|---|
| $CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P_T]]]$ $\Rightarrow^{SendtoCS} CP_T \rightarrow$ | $\Leftarrow^{RecvCP_T\ from\ BS} CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P]]]$ <i>if</i> ($T > \tau$) { $PT \leftarrow D_{S_k}[CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P]]]]$ } |

the sensor using the sensor and cluster identification numbers $\Rightarrow^{SendtoCS} CP_T \rightarrow$. When a message is received from the base station, the sensor verifies the time stamp, and if the verification is successful, the encrypted text is converted to plaintext with the help of the secret key received from the base station $PT \leftarrow D_{S_k}[CP_T \leftarrow E[P_K[Tag[CH_{Id}, S_{Id}, T, P]]]]$. Table 4 details the secure connection between the BS and the LSN of LSKDP.

4 Experimental Analysis

4.1 System Configuration and Software Tools

The proposed key distribution mechanism is developed by using.net with C # framework on the computer with the configuration Intel Core i7 processor, 16 GB RAM, 1000 GB HDD, NVIDIA GEFORCE GTS 1600, and Windows 10 operating system. The communication time, computational overhead, and security analysis are compared with the existing key distribution protocols, including Random Key Distribution Scheme (RKDS), Group Key Distribution Scheme (GKDS), and Key Pre-Distribution Scheme (KPDS).

4.2 Energy Consumption

Energy consumption refers to the amount of energy required for the sensor node to transmit data and perform cryptography. Equation 7 can be used to determine the overall energy consumption of sensor nodes.

$$SE_C = TE - (SE_T + SE_R) \tag{7}$$

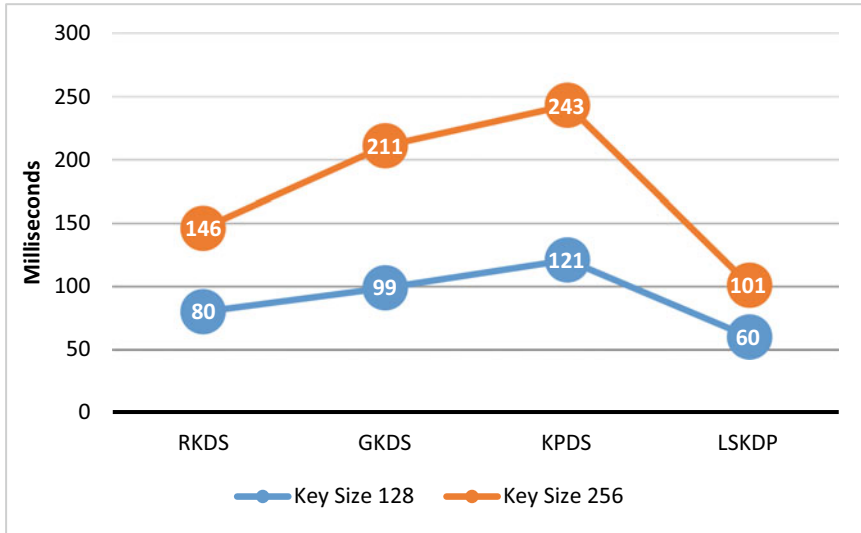


Fig. 3 Energy consumption comparison of proposed key distribution protocol with the existing methods

In the above equation, SE_C denotes energy consumption, TE denotes the total energy consumed by the sensor nodes, and SE_T denotes the energy consumed by general data transfer and the variable SE_R denotes the energy usage of sensor nodes when they receive data.

Figure 3 illustrates the energy consumption necessary for the existing key distribution protocol and the proposed key distribution protocol when data is transferred through the sensor nodes. With current key distribution methods, the sensor node consumes the maximum energy when the key size rises, which is indicated in Fig. 3. In the proposed method, the dominant cryptographic operations are executed by a powerful base station, which minimize the energy consumption of sensor nodes when the key size increases.

4.3 Communication Overhead

The WSN's primary concern is communication overhead. This is because typically keep a greater number of keys on the same LSN. During the critical key distribution and key updation phases, the majority of the communication overhead occurs. The major cause of communication overhead is the improper clustering of LSNs. For instance, in a network model with 50 serialised LSNs, the first LSN must store 50 keys. As a result, WSNs are vulnerable to sensor failure and a variety of attacks. This proposed protocol totally avoids LSN to LSN communication. Thus, the communication overhead is significantly decreased when compared to existing methods, as

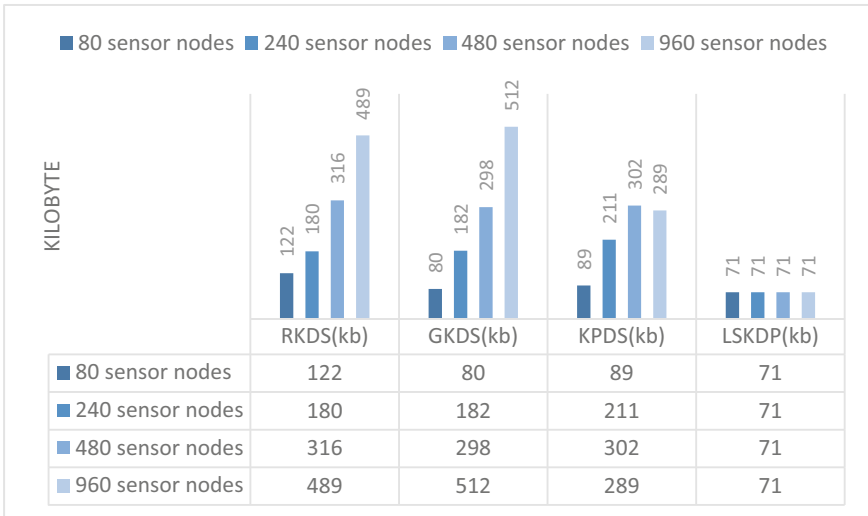


Fig. 4 The proposed key distribution protocol’s communication overhead is compared to the existing approaches

illustrated in Fig. 4. The proposed method requires only 71 kb for key distribution and update. Among the currently available approaches, initial LSN has the highest communication overhead.

4.4 Memory Storage

The next big challenge for WSN is memory utilisation. The suggested technique supports the storage of up to 256-bit key payers. This proposed system completely avoids LSN-to-LSN communication, requiring the LSN to maintain only a single key payer for encryption and decryption. Thus, LSKDP consumes significantly less memory than existing approaches for key processing. The comparison of memory consumption when storing a 128-bit key payer in LSN is shown in Fig. 5. Figure 6 shows the memory usage comparison findings while storing a 256-bit key payer in an LSN.

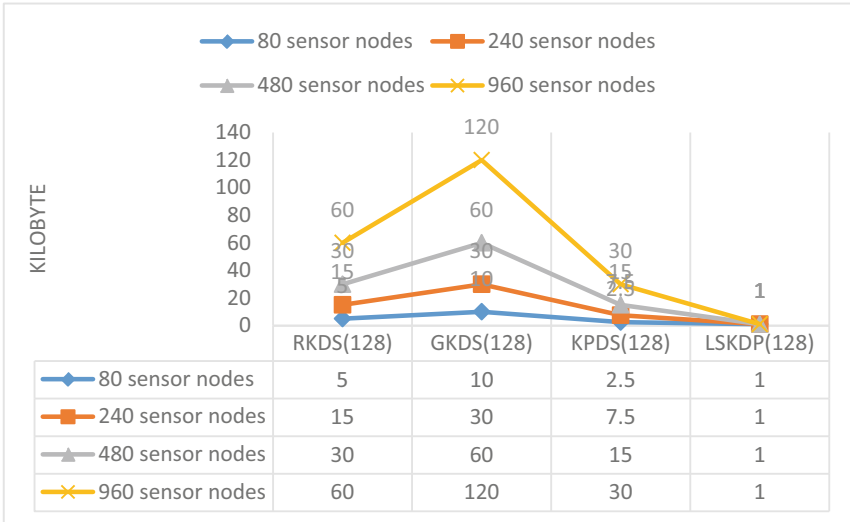


Fig. 5 Memory consumption comparison when storing a 128-bit key payer in LSN

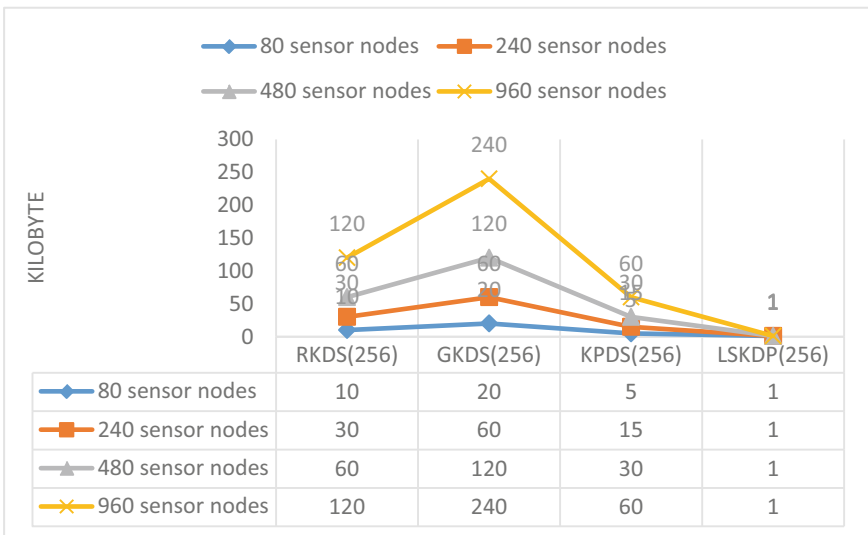


Fig. 6 The comparison of memory utilisation for storing a 256-bit key payer in an LSN

4.5 Security Analysis

This section discusses how the proposed LSKDP protocol safeguards against malicious attacks on the WSN. Attackers perform a variety of malicious attacks against the

WSN, stealing sensitive information, most notably spoofing attacks, replay attacks, and man-in-the-middle attacks.

Spoofing attack

Spoofing attack is the most often used security threat by adversaries on the WSN. When it comes to network traffic, attackers capture key information incorrectly [16]. This typically occurs during key distribution and upgrading. To avoid this, the LSKDP protocol's private key securely stores all sensors prior to their deployment in the WSN. This significantly reduces the likelihood of key pair theft.

Replay attack

In a replay attack, attackers intercept previously delivered communications and send them to recipients wrongly [17]. The primary objective of adversaries is to provide recipients with misleading data. The timestamp is utilised by the LSKDP protocol to efficiently prevent replay attacks.

Man-in-the-Middle Attack

In Man-in-the-Middle Attack the adversary secretly executes malicious communication between senders and receivers [18]. This happens without the knowledge of the authorized sender and receivers. To avoid Man-in-the-Middle attacks, the LSKDP protocol generates and stores key payers at the base station and LSNs prior to deploying the LSNs.

5 Conclusion

The Lightweight Secure Key Distribution Protocol (LSKDP) was developed as a result of this study to strengthen the security of the WSN. The protocol is composed of two critical modules: the formation of LSN clusters and the distribution of keys. LSN cluster formation significantly reduces sensor nodes' energy usage, communication overhead, and memory storage requirements. The key distribution module is used to protect the network from malicious network attacks such as spoofing, replay, and man-in-the-middle. All of this has been proven by experiment and security analysis.

References

1. Liu Y, Dong M, Ota K, Liu A (2016) ActiveTrust: secure and trustable routing in wireless sensor networks. *IEEE Trans Inf Forensics Secur* 11(9):2013–2027. <https://doi.org/10.1109/TIFS.2016.2570740>
2. Zhu J, Zou Y, Zheng B (2017) Physical-layer security and reliability challenges for industrial wireless sensor networks. *IEEE Access* 5:5313–5320. <https://doi.org/10.1109/ACCESS.2017.2691003>

3. Lee CC (2020) Security and privacy in wireless sensor networks: advances and challenges. *Sens (Basel)* 20(3):744. <https://doi.org/10.3390/s20030744>. (Published 20 Jan 29)
4. Li P, Sun L, Fu X, Ning L (2013) Security in wireless sensor networks. in: *wireless network security*. Springer, Berlin. https://doi.org/10.1007/978-3-642-36511-9_8
5. Yasin A, Sabaneh K (2016) Enhancing wireless sensor network security using artificial neural network based trust model. *Int J Adv Comput Sci Appl (IJACSA)* 7(9). <https://doi.org/10.14569/IJACSA.2016.070932>
6. Alotaibi M (2019) Security to wireless sensor networks against malicious attacks using Hamming residue method. *J Wirel Com Netw* 2019:8. <https://doi.org/10.1186/s13638-018-1337-5>
7. Bhalaji N (2021) Cluster formation using fuzzy logic in wireless sensor networks. *IRO J Sustain Wirel Syst* 3(1):31–39
8. Sathesh M, Deepika M (2020) Implementation of multifactor authentication using optimistic fair exchange. *J Ubiquitous Comput Commun Technol (UCCT)* 2(02):70–78
9. Huang M, Yu B, Li S (2018) PUF-assisted group key distribution scheme for software-defined wireless sensor networks. *IEEE Commun Lett* 22(2):404–407. <https://doi.org/10.1109/LCOMM.2017.2778725>
10. Harn L, Hsu C (2015) Predistribution scheme for establishing group keys in wireless sensor networks. *IEEE Sens J* 15(9):5103–5108. <https://doi.org/10.1109/JSEN.2015.2429582>
11. Camtepe SA, Yener B (2007) Combinatorial design of key distribution mechanisms for wireless sensor networks. *IEEE/ACM Trans Netw* 15(2):346–358. <https://doi.org/10.1109/TNET.2007.892879>
12. Bechkit W, Challal Y, Bouabdallah A, Tarokh V (2013) A highly scalable key pre-distribution scheme for wireless sensor networks. *IEEE Trans Wirel Commun* 12(2):948–959. <https://doi.org/10.1109/TWC.2012.010413.120732>
13. Li L et al (2020) A secure random key distribution scheme against node replication attacks in industrial wireless sensor systems. *IEEE Trans Industr Inf* 16(3):2091–2101. <https://doi.org/10.1109/TII.2019.2927296>
14. Shahzad MK, Cho TH (2015) Extending the network lifetime by pre-deterministic key distribution in CCEF in wireless sensor networks. *Wirel Netw* 21:2799–2809. <https://doi.org/10.1007/s11276-015-0941-0>
15. McCusker K, O'Connor NE (2011) Low-energy symmetric key distribution in wireless sensor networks. *IEEE Trans Dependable Secur Comput* 8(3):363–376. <https://doi.org/10.1109/TDSC.2010.73>
16. Jiang Z, Zhao K, Li R et al (2020) PHYAlert: identity spoofing attack detection and prevention for a wireless edge network. *J Cloud Comp* 9:5. <https://doi.org/10.1186/s13677-020-0154-7>
17. Singh AK, Misra AK (2012) Analysis of cryptographically replay attacks and its mitigation mechanism. In: Satapathy SC, Avadhani PS, Abraham A (eds) *Proceedings of the international conference on information systems design and intelligent applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012. Advances in intelligent and soft computing*, vol 132. Springer, Berlin. https://doi.org/10.1007/978-3-642-27443-5_90
18. Kügler D (2003) “Man in the middle” attacks on bluetooth. In: Wright RN (ed) *Financial cryptography. FC 2003. Lecture notes in computer science*, vol 2742. Springer, Berlin. https://doi.org/10.1007/978-3-540-45126-6_11

Detection of Exercise and Cooking Scene for Assistance of Visually Impaired People



Shripad Bhatlawande, Swati Shilaskar, Anant Abhyankar, Mahesh Ahire, Ankush Chadgal, and Jyoti Madake

Abstract This paper focuses on scene understanding for a visually impaired person. It classifies an image into two possible scenes namely, (i) A scene involving some sort of exercise such as yoga. (ii) A scene involving cooking. This system is developed with the aim of assisting the visually impaired people. It will help them to gain knowledge about their surroundings. The existing scene detection systems focus majorly on outdoor activities. This system may be the first system developed to identify indoor activities for visually impaired people. There are considerable challenges in terms of accessibility and usability in the existing systems in terms of lack of useful functions, intuitive feedback, cost, and operation methods. We have developed a system to understand if the detected scene is exercise or cooking. This system uses Scale Invariant Feature Transform to extract the features from the acquired scene. It uses Principal Component Analysis to optimize the dimensions of the feature vector. This optimization is based on maximum explained variance. This optimized feature vector was used to train 5 different classifiers viz- Decision Tree, SVM, K-NN, Random Forest, and Logistic Regression. These classifiers provided recognition accuracy as 75.91%, 80.17%, 79.31%, 80.83%, 79.68% respectively.

Keywords Scene detection · Scene understanding · Image recognition · Blind · Visual impairment · SIFT · Random forest

S. Bhatlawande (✉) · S. Shilaskar · A. Abhyankar · M. Ahire · A. Chadgal · J. Madake
Vishwakarma Institute of Technology, Pune, India
e-mail: Shripad.bhatlawande@vit.edu

S. Shilaskar
e-mail: Swati.shilaskar@vit.edu

A. Abhyankar
e-mail: anant.abhyankar19@vit.edu

M. Ahire
e-mail: mahesh.ahire19@vit.edu

A. Chadgal
e-mail: ankush.chadgal19@vit.edu

J. Madake
e-mail: Jyoti.madake@vit.edu

1 Introduction

Visual Impairment is a medical term used for a person who has any type of vision problem. Whether a person is not able to see anything or just has a problem with not being able to see properly. Many people who are blind cannot see anything but on the other hand, there are people who can see partially and those are called as legal blind.

If we talk about the worldwide statistics of visual impairment people are, In India over an estimated 2.19% of cases of visual impairment are from moderate-severe. And if we talk about the prevalence of pinhole blindness over 0.32% of people are suffering. In short, approximately 18.7 million people are blind in India. One out of every three blind people in the world is Indian.

Age-wise statistics of the people suffering from visual impairment are as 11.6% of people are above the age of 80, 4.1% of people are in between the age of 70–80, 1.6% are from 60–70, and from the age of 50–60 over 0.5% of people are suffering.

Their major problems related to mobility are having problems seeing, so they are not able to move freely because they might collide with obstacles on the road or surface.

The visually impairment people have many problems identifying what is in front of them due to which they deal with many problems while moving from one place to another. There are many devices, or we can say Aids are available in the market for the help of visually defected persons some of them are A radio frequency identification (RFID) smart cane project. RFID is used to detect an impediment [1]. Kines-thesis is a high-tech video game designed specifically for visually impaired people. It provides feedback on the surrounding environment using an in-built camera, buzzers, and vibration motors [2]. Device design and functioning—this device is equipped with sensors that detect impediments such as staircases [4] and potholes [5]. Voice commands are also used to transmit information about them. The blind stick navigator is a technology that delivers an SMS notification and position to the appropriate guardian [6]. “Smart gloves for the blind” are wearable devices in the shape of a glove that costs near about INR 4000/-. Its purpose is to detect obstructions so that blind people can navigate independently. It is less expensive since it requires fewer components. Because of its small size, it is flexible, lightweight, and portable. It delivers real-time translation, but it requires extensive training to utilize. Arduino Uno, IShield, Ultrasonic Sensors, and a Smartphone were among the hardware components used in the construction. Shield Library and Arduino IDE were used as software components [7]. Be My Third Eye [8] is a type of Smart Electronic Blind Stick with Goggles is a hand-held stick featuring an Ultrasonic and IR sensor, goggles, GSM, GPS module motion sensor, PIR sensor, Arduino Uno, and a price tag of INR 1000/-. A stairwell, a wall, and other things are detected as obstacles. By employing ultrasonic waves to detect nearby barriers and notifying them with a buzzer sound or vibration, the device allows blind people to navigate with speed and confidence. However, accurate obstacle detection [9] is impossible due to limited hardware devices and insufficient data [10] 12-inch braille ruler which is specially designed for binds to draw straight

lines and this costs same as the normal ruler [11] E-Z Grabber with Twist Shaft is a tool which costs near about INR 1500/- and is used to reach certain places where it is impossible to reach and is capable of holding a book and has non-slip rubber grip support [12]. Auditory-based Devices for the Blind Perceptual compensation is caused by blindness and manifests as an over-performance of hearing terms of physiology. In truth, a blind person's hearing is significantly more sensitive than a regular person. A highly skilled blind pedestrian can approach an intersection, end to the traffic, and judge the spatial layout of intersecting streets, the width of the street, the number of lanes of traffic in each direction, and the presence of pedestrian islands or medians solely based on auditory information, according to related physiologic research [13]. The vOICE: Meijer, a Dutch physicist, presented a real-time device named voice, which stands for "Oh, I can see." The voice system is a blind navigating system that scans and digitizes an image acquired by a video camera before converting the digital image to sound. It's a wearable gadget. The voice has shown to be highly valuable to the VI (Visually Impaired). The only drawback is that mastering and learning it takes a long time. It's almost as if it's your first language [14]. Sonic Torch and KASPA: Leslie Sy's contributions to solar mobility as a blind aid for the visually handicapped are significant. The majority of the study is focused on binaural, sensor, and Sonic Torch sonar systems. The above uses FM signals to calculate/generate the object's distance based on the pitch of the created sound [15]. Navi (Navigation Assistant for Visually Impaired): R. Nagarajan, Sazali Yaacob, and G. Sainarayanan designed this system. A significant improvement over VOICE. The information in front of the visually impaired was recorded by the vision sensor. The photograph was then analysed to determine what was in front of the person. Object recognition was being done in real-time image processing using fuzzy algorithms at the same time. It was a wearable gadget with much-enhanced image processing algorithms [16]. Blind People's Haptic Devices The haptic or touch senses of VI's are heightened in the same way as the auditory senses of the blind are. Various haptic feedback is provided on the screens to offer the visually impaired an idea about the image, similar to reading braille, which is the language for the blind [17]. Prosthetic eye/artificial eye: The artificial eye provides a visual overview of the surroundings in the shortest period of time possible, allowing a vision-impaired individual to interact with normal people. This work proposes an all-optical technique for greatly improving man-made low-light imaging performance. The concept of superposition and the elephant nose, fisheye were the catalysts for this discovery. It boosted scotopic vision with an innovative photosensitivity enhancer, allowing our device to produce high-resolution images in low-light situations. Many applications require this capability [18]. Haptic smartphone: This study suggests the use of smartphone vibration to deliver turn-by-turn wall directions that are given ahead of time. It compares two feedback modalities, Wand and Screen Edge, with a third, Pattern. The prototype is created by performing a poll on eight random people and then walking them on a pre-programmed path using three different types of input, including vibration. The average error rate was found to be 4%. Pattern approach is selected after the experiment, followed by Screen Edge and Wand method [19].

A patient preference management system is explored in [20] is an enhanced Internet of Things (IoT) module with an artificial intelligence structure for efficiently moving data to the cloud and application layer. This system varies from typical IoT systems in that it adds a primary AI screening layer between the network layer and the cloud layer, and a secondary AI screening layer between the application layer and the cloud layer. Design of Deep Learning Algorithm for IoT Application by Image-based Recognition is explored in [21] by using image modification, the Principal Component Analysis (PCA) technique has extracted the different image features, resulting in excellent experimental findings. The high degree of dispersion, or scatter, occurred after projection aided image identification on the IoT. Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes is explored in [22] each object component can be separated, and a robust scene model can be trained. The geometrical features that concatenate extreme point features, orientation, and polygon displacement values are then extracted from each component. Object detection is aided by these features, while scene understanding is aided by Gaussian Naïve Bayes. The goal of scene understanding is to make machines seem like humans and to grasp visual scenes completely. Cognitive vision influences scene understanding, involving major areas such as computer vision, software engineering, and cognitive engineering [23]. Voice-directed autonomous navigation: This research proposes a voice-directed indoor navigation system for wheelchairs that includes sensors, laser scanners, microphones, and a ring of sonars. It uses the mouse navigation paradigm in an interior setting and can also navigate using voice commands delivered by the user. This greatly aids users with physical and visual impairments in navigating the site [24]. Sightless Literary Devices: There is a paucity of interaction with modern technologies among the world's blind population. This is due to two factors: (1) the restriction of such devices in certain places, and (2) the availability of such devices. (2) Owing to a lack of the necessary software Due to the diversity of the human population, each country has its own set of languages. It is vital to provide effective solutions to such difficulties as a result of the aforementioned elements [25]. Webel Mediatronics Ltd. is the industrial partner for the Bharati Braille Information System, which is a national program managed by IT Kheradpir. For blind individuals, the Braille coding scheme [1–7] is the most common means of reading and writing printed information. Due to differences in script systems, the sightless community in India and the developing globe suffers significant challenges in (a) obtaining source printed reading material and (b) connecting with the sighted community in writing [26]. Laser Wheelchair is a special wheelchair developed for the physically challenged visually impaired people by the Luca University in Sweden. This wheelchair is expected to be available in the markets as of 2022. Lulea University's laser wheelchair is structurally similar to that of traditional chairs. It comes with sturdy wheels and a durable leather seat which would provide comfort for those visually impaired users who need assistance for mobility [27]. The Finger Reader is a finger reader that uses an ultrasonic holography sensor to read the user's fingerprints. This project has produced excellent results; nonetheless, a study on it is currently ongoing in order to improve it [28]. OrCam My Eye: It is a miniature, portable, lightweight, and portable device the size of a finger. It can be held in the

user's hand even magnetically attached to any pair of glasses. This allows users to use both hands for something while wearing the device. Moreover, it does not require an internet connection so that users can use it anywhere and anytime [29]. A system for detecting human posture was introduced. Three-dimensional angles of human arm movement could be determined using parameter data acquired from inertial sensors using a quaternion algorithm for data fusion [30]. 'Human posture recognition' [31] and 'Facial expression recognition have emerged as a key study subject in both the computer-based intelligent video surveillance system and pattern recognition fields. The study assesses the use of human body posture mechanisms for computer interaction, offering numerous ways for accurate identification and facial expression. In facial recognition, the goal is to recognize faces in any image, extract facial characteristics (eyes and lips), and categorize them into six different moods sadness, fear, neutral, disgust, anger, happy. The training data is subjected to a number of filters and procedures before being classified using a Support Vector Machine (SVM) [32].

All the aids which are mentioned above have many limitations but one of the major limitations which we noticed, and which have to be taken care of is cost. All the products which are produced for blinds are made from high technology and hence they are very costly which make these products to be the out of the range. Due to this most people are not able to take this opportunity. And the second one is that most of the tools are highly advanced and need extra care while operating but due to lack of education most of the people are not able to use them and if somehow, they manage to use them they feel very uncomfortable operating the tools without complete knowledge, which also sometimes leads to accidents.

2 Methodology

This paper presents a scene detection system for visually impaired people. It can detect two scenes namely, (i) Exercise and (ii) Cooking. The block diagram of the system is shown in Fig. 1.

It consists of a camera, a processor-based system, and an earphone. The camera captures the scene of the surrounding environment and provides it as an input to a processor-based system. An algorithm has been implemented to detect and interpret exercise and cooking scenes. The system converts the detected scene into audio feedback and conveys it to a visually impaired user via an earphone.

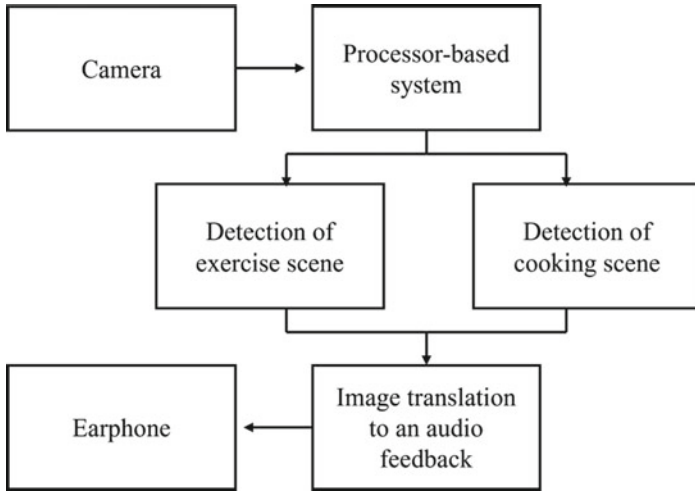


Fig. 1 Block diagram of a complete system

2.1 Dataset and Preprocessing

Total 8220 images were used for implementing this application. The distribution of images in the dataset is shown in Tables 1 and 2.

The exercise scene was defined as a blend of postures such as stretching of arms, legs, bending position, and yoga. The cooking scene was a typical kitchen scene having objects such as vessels, gas stoves, vegetables, and a pan as shown in Fig. 2.

The entire dataset was obtained from the internet [33, 34]. All the images were resized to 200×200 pixels and then converted to grayscale. The dataset preprocessing operations are presented in Algorithm 1.

Table 1 Details of the images in the dataset

| Priority scenes | No. of positive images |
|-----------------|------------------------|
| Exercise | 3550 |
| Cooking | 1120 |

Table 2 Details of the negative images and total images in the dataset

| | |
|------------------------|------|
| No. of negative images | 3550 |
| No. of total images | 8220 |



Fig. 2 Sample images of exercise and cooking dataset

Algorithm 1: Pre-processing of Images

```

Input: Input image
Output: Pre-processed images
    Initialization:
    LOOP Process
1: for every image in the directory do
2: Resizing the images to 200 x 200 pixels
3: Gray scaling the image
4: end for
5: return pre-processed image
    
```

2.2 Feature Vector Compilation

Scale-Invariant Feature Transform (SIFT) was used to extract features from all images in the dataset as shown in Fig. 3. SIFT is a computer vision algorithm that recognizes and matches local features in images. It finds critical locations and then provides descriptors that are utilized for object recognition, SIFT has the capability of uniform scaling and orientation. SIFT is invariant to scale, illumination changes, rotations, occlusions whereas other algorithms fail if they have scale variance, rotation, non-uniform illumination. Therefore, SIFT is superior to other machine learning feature extraction to get better accuracy. Features of every image were individually extracted by SIFT and appended into a CSV file. An $M \times 128$ feature vector array is generated for every image. Where ‘M’ represents a number of key points extracted from an image. The size of the SIFT feature vector for the entire dataset was 2019256×128 .

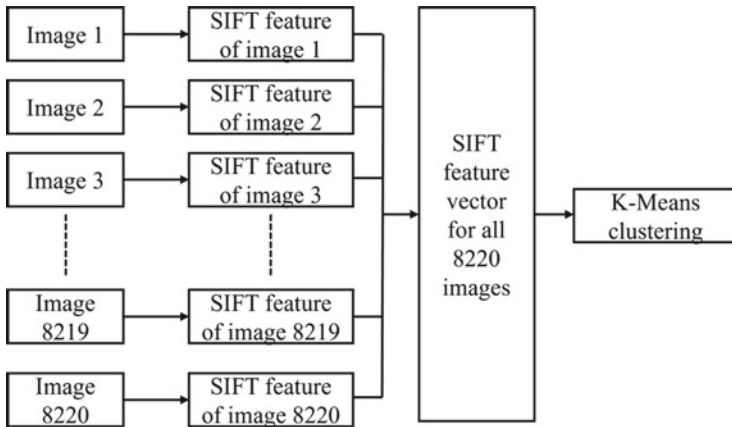


Fig. 3 Training of large feature vector

The large size feature vector was divided into 5 clusters by using the K-Means clustering algorithm. The number of clusters was decided based on the elbow method. SIFT features of all images were predicted with the pre-trained K-Means model. Then the output histogram of each image was normalized and appended into a CSV file. The feature vector was reduced to 8220×5 after applying the K-Means clustering algorithm. The principal vector component analysis (PCA) was used to further reduce the dimensions of the feature vector to 8220×4 . We measured the information contained in each bin, by comparing the maximum explained variance that was accommodated in the first 4 principal components. The information table in each principal component is shown in Table 3. After adding the first 4 principal components 99% of information was obtained. Therefore, 4 PCA components were selected based on maximum explained variance.

The overall procedure for dimensionality reduction and feature vector compilation is shown in Fig. 4 and presented in Algorithm 2. The optimized feature vector was then provided as an input to an array of classifiers.

Table 3 The percentage of variance explained by principal components (PC) generated by PCA

| PC1 | PC2 | PC3 | PC4 | PC5 |
|--------|--------|-------|-------|-------|
| 78.16% | 13.14% | 4.77% | 2.93% | 0.98% |

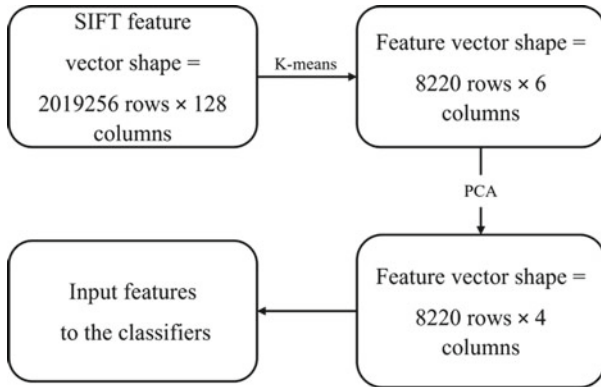


Fig. 4 SIFT feature compilation and dimensionality reduction

Algorithm 2: Dimensionality Reduction

Input: Feature vector (2019256 rows × 128 columns)

Output: Optimized feature vector (8220 rows × 4 columns)

Initialization:

LOOP Process

- 1: **for** every image in a specified path **do**
 - 2: Perform prediction on pre-trained k-means clustering
 - 3: Normalized the data
 - 4: Append it to feature vector
 - 5: **end for**
 - 6: Standardization using standard scalar
 - 7: Perform PCA with n_components=4
 - 8: Perform PCA transform
 - 9: **return** optimized feature vector (8220 rows × 4 columns)
-

2.3 Classification and Detection of Scene

This system uses an array of five classification algorithms namely, (i) Decision Tree (DT) (ii) Support Vector Machine (SVM) (iii) K-Nearest Neighbour (K-NN) (iv) Random Forest (RF) and (v) Logistic Regression (LR) for classification and detection of desired scenes. Performance parameters such as training accuracy, testing accuracy, precision score, recall score, F1 score, were measured for the assessment of the detection.

The first classifier used was DT. It generates rules to forecast any unseen data.

$$E(s) = \sum_{i=1}^c -p_i \log_2 p_i \quad (1)$$

Here p_i , stands for the probability of class i .

The second classifier used was SVM. It describes the best decision boundary for categorizing n -dimensional space into classes. In SVM for multiclass classification, we used two methods namely (i) One-vs-rest (OVR), (ii) One-vs-one (OVO). The equation for SVM is.

$$L(w) = \sum_{i=1} \max(0, 1 - b_i(w^s m_i + z)) + \lambda \|w\|_2^2 \quad (2)$$

where w is the input and m_i , is the support vector. Coefficients b_i , z and λ are estimated from training data.

The K-NN was the third classifier used. It predicts the values of new data points based on feature similarity. The new data point will be assigned a value depending on how closely it matches the points in the training set. It is based on Euclidean distance measure and its equation is,

$$D = \sqrt{(m_1 - m_2)^2 + (n_1 - n_2)^2} \quad (3)$$

where D is the calculated distance, m_1 and n_1 , are x coordinates and y coordinates of the feature vector.

The Random Forest algorithm is the fourth classifier used. The dataset is divided into subsets by a random forest classifier. Every decision tree receives these subsets, and each one generates its own output.

$$MSE = \frac{1}{N} \sum_{i=1}^N (f_i - b_i)^2 \quad (4)$$

where N is the number of data points, f_i is the value returned by the model, and b_i is the actual value for data point i .

The last classifier used was LR. It is the supervised machine learning algorithm.

$$p = \frac{e^{a+bx}}{1 + e^{a+bx}} \quad (5)$$

The probabilistic value ranges in the form 0 and 1. The outcomes are categorized in the form of 1 and 0.

The overall process for classification and detection of the exercise and cooking scene is described in Algorithm 3. The overall flow diagram of the proposed system is shown in Fig. 5. The system classifies the detected scene and label it 0 if it is exercise scene. It assigns label 1, if the detected scene is cooking scene. The system generates spoken message “Exercise” or “Cooking” based on label 0 and 1 respectively.

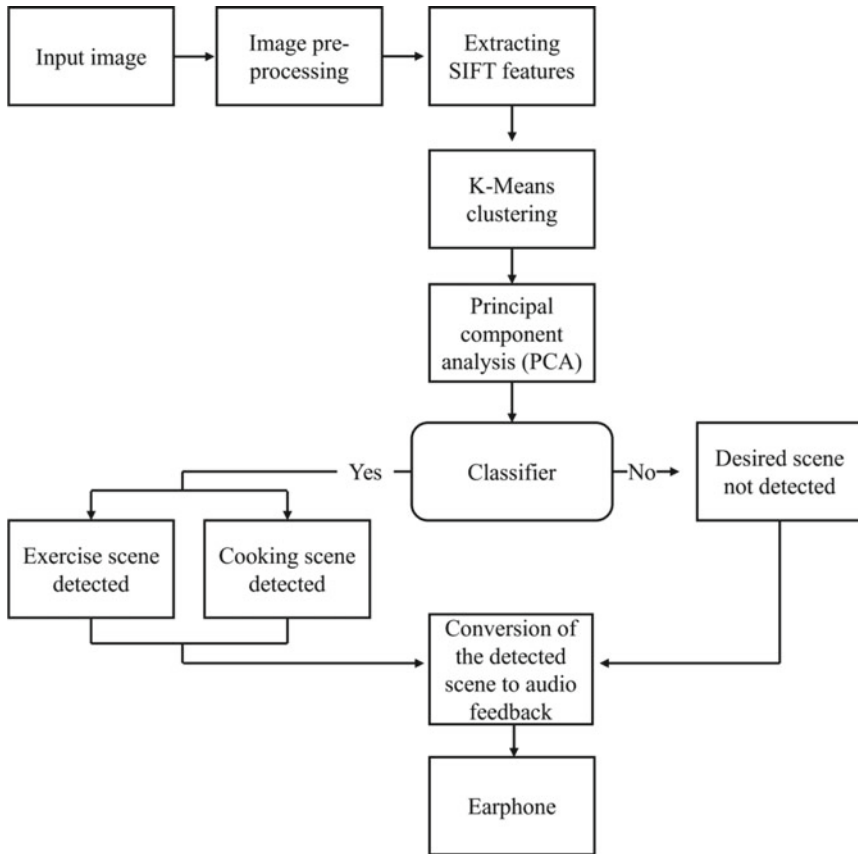


Fig. 5 Overall flow diagram for scene detection and communication

Algorithm 3: Classification and Detection of the Scene

Input: Optimized feature vector(8220 rows \times 4 columns)

Output: Classified Exercise and Cooking Scene

Initialize: min_accuracy = 75.91%

Procedure:

- 1: Fit the model with training data for the Random Forest classifier
 - 2: Predict the test data using this pre-trained model
 - 3: **if** (predicted class label ==0 && accuracy \geq min_accuracy)
 - 4: **return** Exercise Scene Detected
 - 5: **else if** (predicted class label ==1 && accuracy \geq min_accuracy)
 - 6: **return** Cooking Scene Detected
 - 7: **else if** (predicted class label ==2 && accuracy \geq min_accuracy)
 - 8: **return** Neither Exercise nor Cooking Scene
 - 9: **else**
 - 10: **return** Not able to classify the scene
-

3 Result

An array of five classifiers was used for the assessment of scene detection. The detailed performance analysis of the five classifiers is shown in Tables 4 and 5. and cooking scene detection.

Total 1000 test images (exercise and cooking) were used for the performance evaluation of classifiers. All classifiers accurately classified majority images. Among all classifiers, the Random Forest classifier provided the highest training and testing accuracy of 92.92% and 80.83% respectively. This model uses a Random Forest classifier for predicting the exercise and cooking scene.

Table 4 Training accuracy and testing accuracy analysis for detection of exercise and cooking scene

| Priority classifiers | Training accuracy (%) | Testing accuracy (%) |
|----------------------|-----------------------|----------------------|
| Decision tree | 75.71 | 75.91 |
| SVM—OVR linear | 78.68 | 79.50 |
| SVM—OVR polynomial | 76.58 | 77.98 |
| SVM—OVR RBF | 79.80 | 80.17 |
| SVM—OVO linear | 78.68 | 79.50 |
| SVM—OVO polynomial | 76.58 | 77.98 |
| SVM—OVO RBF | 79.80 | 80.17 |
| K-NN | 84.48 | 79.31 |
| Random forest | 92.92 | 80.83 |
| Logistic regression | 78.90 | 79.68 |

Table 5 Analysis of precision, recall, and F1-score

| Priority classifiers | Precision score (%) | Recall score (%) | F1 score (%) |
|----------------------|---------------------|------------------|--------------|
| Decision tree | 50.71 | 58.29 | 54.22 |
| SVM—OVR linear | 74.64 | 67.13 | 68.22 |
| SVM—OVR polynomial | 76.09 | 63.85 | 64.25 |
| SVM—OVR RBF | 75.25 | 68.18% | 69.40 |
| SVM—OVO linear | 74.50 | 67.13 | 68.22 |
| SVM—OVO Polynomial | 76.09 | 63.85 | 64.25 |
| SVM—OVO RBF | 75.25 | 68.18 | 69.14 |
| K-NN | 72.43 | 70.09 | 70.88 |
| Random forest | 75.36 | 71.58 | 72.80 |
| Logistic regression | 73.86 | 68.34 | 69.50 |

4 Conclusion

This system helped in classifying an image into two possible scenes. The two distinct scenes are (i) A scene involving some sort of exercise such as yoga, ii) A scene involving cooking. It is very important for everyone to recognize surroundings while walking or moving from one place to another. This project will help visually disabled people to sense what's happening near them.

The system accuracy reduces if there is more than one person at a time in the frame. The system performance is reduced when there is non-uniform illumination in the captured image.

This system can be extended to include activities such as swimming, playing cricket, and many other things if trained accordingly. This can be further improved with a Graphical user interface (GUI) for better understanding and can also be deployed in a kind of hardware such as a blind stick that will alert the user that there are people in front of you performing the following tasks.

Acknowledgements We express our sincere gratitude to the visually impaired participants in this study, orientation and mobility (O&M) experts and authorities at The Poona Blind Men's Association, Pune. The authors thank the La Fondation Dassault Systemes, India and Vishwakarma Institute of Technology Pune for providing support (IN-2019-056) to carry out this research work.

References

1. Dey N, Paul A, Ghosh P, Mukherjee C, De R, Dey S (2018) 2018 International conference on current trends towards converging technologies (ICCTCT)—ultrasonic sensor based smart blind stick (IEEE 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, India (2018.3.1–2018.3.3))
2. Mennens J, Van Tichelen L, Francois G, Engelen JJ (1994) Optical recognition of Braille writing using standard equipment. *IEEE Trans Rehabil Eng* 2(4):207–212
3. Yokokohji Y, Yoshikawa T (1994) Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment. *IEEE Trans Robot Autom* 10(5):605–620
4. Munoz R, Rong X, Tian Y (2016) 2016 IEEE international conference on multimedia & expo workshops (ICMEW)—depth-aware indoor staircase detection and recognition for the visually impaired (IEEE 2016 IEEE international conference on multimedia & expo workshops (ICMEW), Seattle, WA, USA (2016.7.11–2016.7.15))
5. Reddy EJ, Reddy PN, Maithreyi G, Balaji MBC, Dash SK, Kumari KA (2020) Development and analysis of pothole detection and alert based on NodeMCU. In: 2020 International conference on emerging trends in information technology and engineering (ic-ETITE)
6. Saaid MF, Ismail I, Noor MZ (2009) Radio frequency identification walking stick (RFIWS): a device for the blind. In: 2009 5th international colloquium on signal processing & its applications
7. Jain S, Varsha SD, Bhat VN, Alamelu JV (2019) 2019 International conference on communication and signal processing (ICCSP)—design and implementation of the smart glove to aid the visually impaired (IEEE 2019 international conference on communication and signal processing (ICCSP), Chennai, India (2019.4.4–2019.4.6))

8. Arvitha MVS, Biradar AG, Chandana M (2021) Third eye for visually challenged using echolocation technology. In: 2021 International conference on emerging smart computing and informatics (ESCI).
9. Chen L, Guo B-L, Sun W (2010) 2010 Fourth international conference on genetic and evolutionary computing—obstacle detection system for visually impaired people based on stereo vision (IEEE 2010 fourth international conference on genetic and evolutionary computing (ICGEC 2010), Shenzhen (2010.12.13–2010.12.15))
10. Paul IJ, Sasirekha S, Mohanavalli S, Jayashree C, Priya PM, Monika K (2019) 2019 International conference on computational intelligence in data science (ICCIDS)—smart eye for visually impaired—an aid to help the blind people (IEEE 2019 international conference on computational intelligence in data science (ICCIDS), Chennai, India (2019.2.21–2019.2.23))
11. Mennens J, Van Tichelen L, Francois G, Engelen JJ (1994) Optical recognition of Braille writing using standard equipment. *IEEE Trans Rehabil Eng*
12. Vaccarella A, Comparetti MD, Enquobahrie A, Ferrigno G, De Momi E (2011) 2011 Annual international conference of the IEEE engineering in medicine and biology society—sensors management in robotic neurosurgery: the ROBOCAST project (IEEE 2011 33rd annual international conference of the IEEE engineering in medicine and biology society, Boston, MA (2011.08.30–2011.09.3))
13. Leporini B, Rosellini M, Forgione N (2020) Designing assistive technology for getting more independence for blind people when performing everyday tasks: an auditory-based tool as a case study. *J Ambient Intell Humlzed Comput*
14. Karthik A, Raja VK, Prabakaran S (2018) 2018 International conference on communication, computing and Internet of Things (IC3IoT)—voice assistance for visually impaired people (IEEE 2018 international conference on communication, computing and Internet of Things (IC3IoT), Chennai, India (2018.2.15–2018.2.17))
15. Ng SS (1991) Proceedings of the annual international conference of the IEEE engineering in medicine and biology society—sonic electronic guide for the blind, vol 13 (IEEE Annual International Conference of the IEEE engineering in medicine and biology society, Orlando, FL, USA (31 Oct–3 Nov 1991))
16. Khan S, Nazir S, Khan HU (2021) Analysis of navigation assistants for blind and visually impaired people: a systematic review. *IEEE Access* 9:26712–26734
17. Menikdiwela MP, Dharmasena KMIS, Abeykoon AM, Harsha S (2013) 2013 International conference on circuits, controls and communications (CCUBE)—haptic based walking stick for visually impaired people (IEEE 2013 international conference on circuits, controls and communications (CCUBE), Bengaluru, India (2013.12.27–2013.12.28))
18. Ponnavaillio M, and Kumar VP (2000) The artificial eye. *IEEE Potentials* 18(5):33–35 (Dec 1999)
19. Akhter S, Mirsalahuddin J, Marquina FB, Islam S, Sareen S (2011) 2011 IEEE 37th annual northeast bioengineering conference (NEBEC)—a smartphone-based haptic vision Substitution system for the blind (IEEE 2011 37th annual northeast bioengineering conference (NEBEC), Troy, NY, USA (2011.04.1–2011.04.3))
20. Sathesh A (2020) Computer vision on IoT based patient preference management system. *J Trends Comput Sci Smart Technol* 2(2):68–77
21. Jeena IJ, Darney PE (2021) Design of deep learning algorithm for IoT application by image based recognition. *J ISMAC* 3(03):276–290
22. Rafique AA, Jalal A, Ahmed A (2019) Scene understanding and recognition: statistical segmented model using geometrical features and gaussian naïve bayes. *Int Conf Appl Eng Math (ICAEM)* 2019:225–230
23. Aarthi S, Chitrakala S (2017) Scene understanding—a survey. 2017 International conference on computer, communication and signal processing (ICCCSP), pp 1–4. <https://doi.org/10.1109/ICCCSP.2017.7944094>

24. Megalingam RK, Nair RN, Prakhya SM (2011) 2011 2nd International conference on wireless communication, vehicular technology, information theory and aerospace & electronic systems technology (Wireless VITAE)—automated voice based home navigation system for the elderly and the physically challenged (IEEE electronic systems technology (Wireless VITAE), Chennai, India (2011.02.28–2011.03.3))
25. Basu A, Dutta P, Roy S, Banerjee S (1998) A PC-based Braille library system for the sightless, 6(1):60–65. <https://doi.org/10.1109/86.662621>
26. Chaves DR, Peixoto IL, Lima ACO, Vieira MF, de Araujo CJ (2009) 2009 IEEE international workshop on medical measurements and applications—microactuators of SMA for Braille display system (IEEE 2009 IEEE international workshop on medical measurements and applications (MeMeA), Cetraro, Italy (2009.05.29–2009.05.30))
27. Trieu HT, Nguyen HT, Willey K (2008) 2008 30th Annual international conference of the IEEE Engineering in medicine and biology society—shared control strategies for obstacle avoidance tasks in an intelligent wheelchair, pp 4254–4257. <https://doi.org/10.1109/iembs.2008.4650149>. (IEEE 2008 30th Annual international conference of the IEEE engineering in medicine and biology society, Vancouver, BC (2008.08.20–2008.08.25))
28. Kowshik S, Gautam VR, Suganthi K (2019) 2019 11th International conference on advanced computing (ICoAC)—assistance for visually impaired using finger-tip text reader using machine learning (IEEE 2019 11th International conference on advanced computing (ICoAC), Chennai, India (2019.12.18–2019.12.20))
29. Younis O, Al-Nuaimy W, Rowe F, Alomari MH (2018) Real-time detection of wearable camera motion using optical flow. *IEEE Congr Evol Comput (CEC)* 2018:1–6
30. Ni W et al (2016) Human posture detection based on human body communication with multi-carriers modulation. In: 2016 39th International convention on information and communication technology, electronics and microelectronics (MIPRO), pp 273–276
31. Adochiei F-C, Adochiei IR, Ciucu R, Pietroiu-Andruseac G, Argatu FC, Jula N (2019) Design and implementation of a body posture detection system. In: 2019 E-health and bioengineering conference (EHB), pp 1–4
32. Jain C, Sawant K, Rehman M, Kumar R (2018) Emotion detection and characterization using facial features. In: 2018 3rd International conference and workshops on recent advances and innovations in engineering (ICRAIE), pp 1–6
33. Saxena S (2021) Yoga pose image classification dataset. <https://www.kaggle.com/shrutisaxena/yoga-pose-image-classification-dataset>. Last Accessed 03 Apr 2021, 25 Nov 2021
34. Cooking using pan. <https://www.google.co.in/search?q=cooking+using+pan&hl=en&tbm=isch&source>. Last Accessed 20 Oct 2021, 25 Nov 2021

Electronic System for Navigation of Visually Impaired People



Jyoti Madake, Bhumika Bijane, Geetai Charde, Abhijit Chine, Shripad Bhatlawande, and Swati Shilaskar

Abstract People who are blind or visually challenged require frequent support when moving from one location to another. A smart device can help assist a blind person in the best possible way. In this research paper, the proposed system gives a path to the user to reach one of the 10 destinations from one of the source positions. The system is designed to provide guided instructions on 100 chosen paths. The interface for path selection is user-friendly considering the limitation of the blind person. System mapped an appropriate route for their selections. This system has voice help and is user-interactive.

Keywords Path planning · Visually impaired · Voice assistance system · And user friendly

1 Introduction

The most crucial thing is vision, which allows you to view the world around you. Several components within your eye and brain work together to give you vision. Vision loss can occur for a variety of reasons as stated by [1, 2]. These could be caused by eye damage, the brain's inability to accept and interpret visual cues supplied by

J. Madake (✉) · B. Bijane · G. Charde · A. Chine · S. Bhatlawande · S. Shilaskar
Vishwakarma Institute of Technology, Pune, India
e-mail: jyoti.madake@vit.edu

B. Bijane
e-mail: bhumika.bijane19@vit.edu

G. Charde
e-mail: geetai.charde19@vit.edu

A. Chine
e-mail: abhijit.chine19@vit.edu

S. Bhatlawande
e-mail: shripad.bhatlawande@vit.edu

S. Shilaskar
e-mail: Swati.shilaskar@vit.edu

the eyes, and so on. A functional limitation of the eye or eyes, or the vision system, is referred to as visual impairment. As a result [1] the person loses visual acuity and is unable to perceive objects as clearly as a healthy person. Loss of the visual field refers to an individual's inability to see as far as the ordinary person without moving his or her eyes or rotating his or her head. There were an estimated 253 million people with vision impairment globally, according to Worldwide Statistics for Blindness and Visual Impairment [3, 4]. 36 million were blind, with another 217 million suffering from moderate to severe visual impairment (MSVI). People with range visual impairment account for 3.44% of the population, with 0.49% being blind and 2.95% having MSVI. One further 1.1 billion people are thought to be affected by functional presbyopia. Most eye diseases become more common as people get older; as a result, the prevalence of blindness and MSVI is substantially higher in older age groups. 80% of the 253 million visually impaired people on the planet are above the age of 50. Women account for 55% (139 million) of the world's 253 million visually impaired people, while males account for 45%. Blind folks have their own way of doing things and lead normal lives. They do, however, confront difficulties as a result of inaccessible infrastructure and social issues. Navigating around locations is the most difficult task for a blind person, especially one who has lost all eyesight. The majority of people with visual impairments prefer to use a long walking cane as an aid in their daily mobility activities. The most basic, adaptable, and low-maintenance assistive device is a white cane [5, 6].

A guide dog is useful to detect and avoid barriers in the course of travel, and also extreme circumstances [7]. The Medico Stick allows people who are blind to confidently walk down the road [8], both the white cane and the guide dog, give a brief description and are unable to identify overhead objects. Since then, a descriptive and analytical study has been performed. Many mobility aids have been created in recent decades to address such problems, namely obstacle avoidance [9] understanding maps [10, 11], and navigating in unfamiliar places [12, 13]. There are numerous existing systems that are based on sonar [14–16]. There are several types of navigation systems, including Ultrasonic sensor-based navigation [17], GPS-based system [18], Laser-based navigation i.e. Laser cane [19], and Teletact [20]. A GPS-based system was implemented for navigation. With the help of a secure and effective navigation device namely GPS [21] this system covers unusual routes as well as unexpected destinations. Another approach for the goal of guiding visually challenged people, GPS and voice recognition, as well as obstacle avoidance [22]. An Android GPS-Based Navigation Application for Blind, the application is proposed for traveling from one place to another [23]. RFID and GPS Integrated Navigation System, the system uses RFID and GPS based localization while operating indoor and outdoor respectively [24]. GPS and GSM methods are used to create a navigation system for blind persons. A walking stick with a microcontroller, infrared sensors, a GPS receiver, label surface detection, a buzzer, haptic feedback [11], and a vibrating motor have been constructed to assist a blind person in detecting obstacles and navigating to their destination [25]. In this case, navigation refers to a process that involves moving from one location to another. Visually impaired people require the assistance of others in order to travel and are dependent on others to transport them from one location to

another. This work proposed one such system that builds a path from current places to their destinations using input from a keypad with pre-assigned keys. Considering a person’s decision of where they want to travel using pre-assigned keys from a keypad and system mapped an appropriate route for their selections. This system has voice help and is user-interactive. This is a very safe and reliable navigation system that allows many blind people to travel without worrying about their position or course guidance while on the go. So that blind people can travel and move around independently even in uncomfortable situations.

2 Methodology

This paper presents an electronic system for the navigation of visually impaired people. Figure 1 shows a block diagram. It consists of a numerical keypad, processor-based system, internet dongle, and earphones. It accepts the source and destination of the known locations such as school, tea stall, clinic, market, bus stop, temple etc. from user and compiles a path. It provides details of the identified path to the user through earphones and generates a map for functional verification.

Two datasets are created for the system. The first dataset is of source and destination and the second is for the waypoint. The first dataset contains 4 columns and 10 rows in the excel file. Columns include Loc_id (key number) to which key specific destination is assigned. The second column is about the locations. The third and fourth columns contain the respective latitude and longitude of a location is mentioned in Table 1. Location dataset.

The second dataset is of waypoints between the source and the destination. It contains 6 columns. First column is waypt_start_id that indicates key number of source location. Second column is waypt_stop_id that indicates key number of destination location. Third column is waypt_seq indicating number of waypoints between source and destination in sequential manner. Fourth and fifth columns are waypt_lat and waypt_long contain the latitude and longitude of waypoints respectively. And last column is waypt_nm contains the name of the waypoints between the source and destinations.

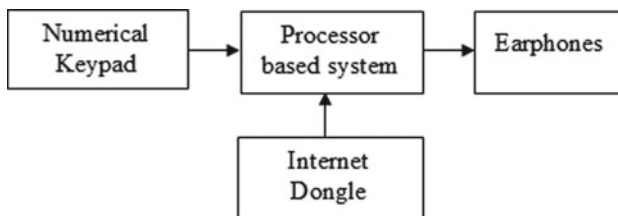


Fig. 1 System schematic

Table 1 Location dataset

| Loc_id | Location | Latitude | Longitude |
|--------|--------------------------|----------|-----------|
| 0 | VIT | 18.46362 | 73.86819 |
| 1 | Dolphin chowk | 18.46135 | 73.86631 |
| 2 | Sukhsagar nagar | 18.45653 | 73.87011 |
| 3 | Lake town | 18.45743 | 73.86403 |
| 4 | Manapa bus station | 18.52319 | 73.85383 |
| 5 | Chintamani nagar | 18.46631 | 73.86618 |
| 6 | Mahesh society | 18.46905 | 73.86505 |
| 7 | Bibwewadi police station | 18.46733 | 73.86414 |
| 8 | K K market | 18.46987 | 73.85843 |
| 9 | Swargate bus stand | 18.49715 | 73.85806 |

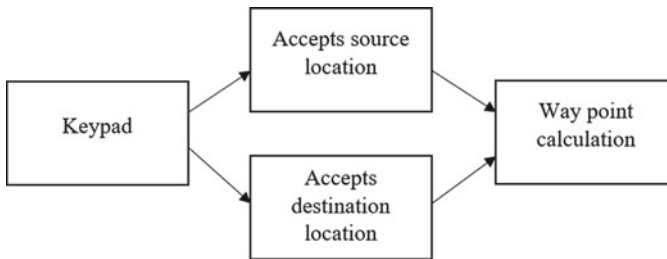


Fig. 2 Initial phase (accepting input)

System asks the user to enter the key of the source location among the 10 locations from Table 1 and then to enter the key for destination location as mentioned in Fig. 2 the data is used to calculate waypoints.

Figure 3 shows the calculation of waypoints between source and destination. If the selected source ID is less than the destination ID then the loop is executed in the increasing order of the waypoint sequence. Otherwise, loop is executed in the decreasing order of the waypoint sequence.

Algorithm 1 shows the calculation of the waypoints. The calculated waypoint is then given to the next part of the model i.e. path planning.

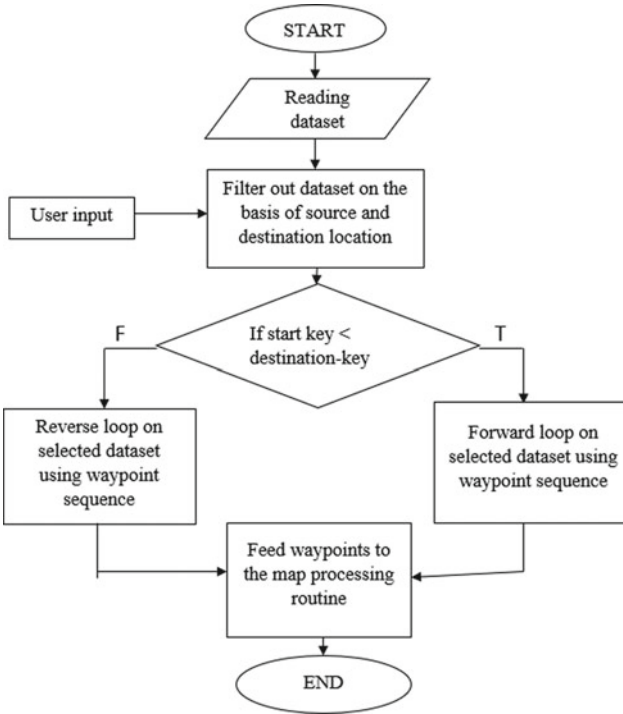


Fig. 3 Calculation of waypoints

Algorithm 1 Calculating waypoints

Input Waypoints

Output Path

1. Reading waypoints dataset and taking source and destination keys from the user.
 2. Filter out dataset on the basis of source and destination key
 3. **If start_id < destination_id:**
 Execute forward loop on dataset using waypoints sequence
else:
 Execute reverse loop on dataset using waypoints sequence.
 4. Pass the calculated dataset list to the map processing function of algorithm 2.
 5. Display map and plot all the locations / waypoints on a map using algorithm 2.
 6. **End**
-

Figure 4 shows integration of computed path with Google Map and Audio output. As the system displays the path on the google map at the same time it will also give the audio output of the track between entered source location and the destination location to inform the blind person.

If the input is less than or equal to 10, then the second condition of user conformations is executed. Otherwise, system will utter invalid input as shown in Fig. 5. The required output is displayed as well as given as a voice message.

Algorithm 2 Path planning

Input Source and destination

Output Map showing shortest path from source to destination

1. Importing required Libraries and Google map API key.
 2. Define functions for voice message and text to speech conversion
 3. Reading predefined location dataset.
 4. Taking source and destination location ids from the user and give a voice message
 5. Validating users input:
 - If Loc_id <= 10:**
Proceed further
 - Else:**
Back to ask user's input (step 4).
 6. Confirm the user's input (Yes / No):
 - If input == Yes:**
Proceed further
 - Else:**
Back to ask user's input (step 4).
 7. Select location details (location name, longitude, latitude) from the dataset on the basis of user's input.
 8. Pass coordinates waypoints list (from algorithm 1) and configurations to the map processing function.
 9. Give voice message for the planned path
 10. **End**
-

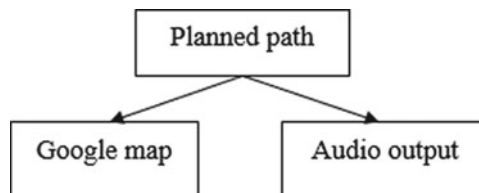


Fig. 4 Integration of path with Google map and audio output

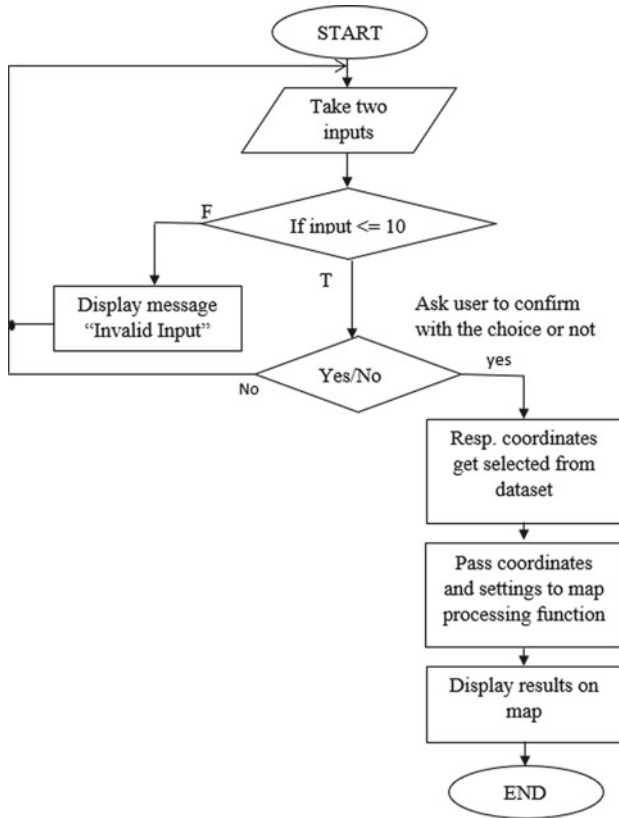


Fig. 5 Flow diagram for path planning

Algorithm 2 shows that the waypoints are given and then it completes the calculations of path planning and as a result of it output is shown on the map and a voice message.

The system is only designed to provide the path between sources and destination mentioned in the dataset. If the user presses any invalid key then the system will utter the message “Invalid key” and again ask the user to press valid key for source and destination. The Google Directions API is used to design the path. This API plots the result on the map. Google Map is used for the functional verification of routes

and waypoints. The blind person will get to know the correct path before starting his journey so that the user can start his plan. The proposed can be either deployed on low cost microcontroller based device or can be ported on android based mobile phone. The cost associated with its development is low and the navigation assistance provided by this device is efficient.

3 Result

The entire path planning trial was run using a blind person. The user has entered ‘1’ for the source as Dolphin Chowk and ‘0’ for the destination as VIT and requested for path planning. There were total 3 waypoints in between them which are Supriya cafe, upper market road, and upper Indira Nagar corner. System processed on waypoint calculation and path planning. System generated audio output of the route between source and destination with the waypoints in sequential manner as A, B, C, etc. and displays the shortest path between source and destination which is highlighted with blue color line along with the waypoints as shown in Fig. 6.

The route is also verified using satellite view. The output of the system was examined and compared with the actual route, which is depicted on the map correctly indicates that the system is efficiently working as shown in Fig. 7.



Fig. 6 Map showing path from dolphin chowk to VIT

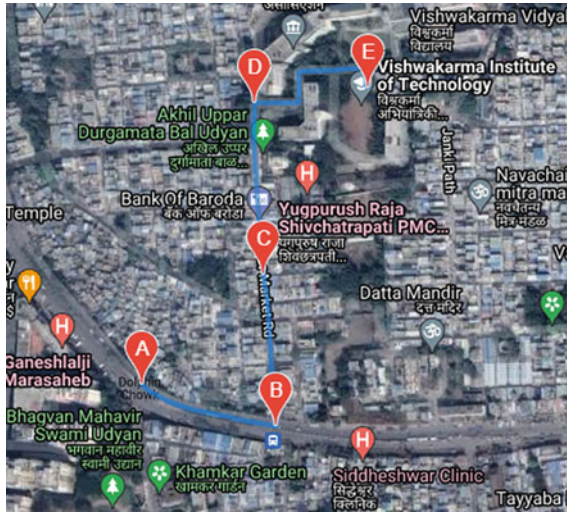


Fig. 7 Map showing path from dolphin chowk to V I T in satellite view

4 Conclusion

In this work, a solution for the navigation of blind people is proposed. The system is designed to guide blind person on his 100 frequently traveled path. A collection of frequently preferred pathways taken by individuals in their daily lives are considered for mapping. It is user friendly interface as user has to just press the key which is assigned with respect to source and destination and system will automatically displays the route between them. This method is both cost-effective and efficient. This gives the blind individual a portable, dependable, and long-lasting navigational choice. Also, the waypoint added in the path on the map gives more confidence to the user while using the system. With every new update, a new feature appears. In the future scope of this work through cognition, a preview of the path can be made available to the visually impaired person. This system can be extended to dynamically take updates from the existing navigation system and suggest alternative routes. A GPS can be added to track the user. Also, this will be useful to their relatives to track them. The limitation of this system is that it is limited to one specific area. This system makes blind people feel confident and independent.

Acknowledgements We express our sincere gratitude to the visually impaired participants in this study, orientation and mobility (O&M) experts and authorities at The Poona Blind Men’s Association, Pune. The authors thank the La Foundation Dassault Systèmes, India and Vishwakarma Institute of Technology Pune for providing support (IN-2019-056) to carry out this research work.

References

1. Manduchi R, Kurniawan S (2011) Mobility-related accidents experienced by people with visual impairment. *AER J Res Pract Vis Impair Blind* 4(2):44–54
2. Zeng L (2015) A survey: outdoor mobility experiences by the visually impaired. In: Mensch und computer 2015–Workshopband. De Gruyter, pp 391–398
3. Ackland P et al (2017) World blindness and visual impairment: despite many successes, the problem is growing. *Community Eye Health* 30(100):71–73
4. Blindness and visual impairments: global facts (2012). <http://www.vision2020.org>
5. Faria J, Lopes S, Fernandes H, Martins P, Barroso J (2010) Electronic white cane for blind people navigation assistance. *World Autom Congr* 2010:1–7
6. Pyun R, Kim Y, Wespe P, Gassert R, Schneller S (2013) Advanced augmented white cane with obstacle height and distance feedback. In: 2013 IEEE 13th international conference on rehabilitation robotics (ICORR), pp 1–6. <https://doi.org/10.1109/ICORR.2013.6650358>
7. Iwatsuka K, Yamamoto K, Kato K (2004) Development of a guide dog system for the blind people with character recognition ability. In: Proceedings of the 17th international conference on pattern recognition, ICPR 2004, vol 1. IEEE, pp 453–456. <https://doi.org/10.1109/ICPR.2004.1334161>
8. Shrivastava P, Anand P, Singh A, Sagar V (2015) Medico stick: an ease to blind & deaf. In: 2015 2nd International conference on electronics and communication systems (ICECS). IEEE, pp 1448–1452
9. Zeng L, Prescher D, Weber G (2012) Exploration and avoidance of surrounding obstacles for the visually impaired. In: Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility, pp 111–118
10. Wang Z, Li N, Li B (2012) Fast and independent access to map directions for people who are blind. *Interact Comput* 24:91–106
11. Zeng L, Miao M, Weber G (2015) Interactive audio-haptic map explorer on a tactile display. *Interact Comput* 27(4):413–429
12. Katz B, Kammoon S, Parsehian G, Gutierrez O, Brilhault A, Auvray M, Trulliet P, Denis M, Thorpe S, Jouffrais C (2012) NAVIG: augmented reality guidance system for the visually impaired. *Virtual Reality* 16(4):253–269
13. Yatani K, Banovic N, Truong KN (2012) SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback. In: Proceedings of ACM CHI 2012, pp 415–424
14. Ikbal MA, Rahman F, Ali MR, Kabir MH, Furukawa H (2017) Smart walking stick for blind people: an application of 3D printer. In: Proceedings of the SPIE 10167, nanosensors, biosensors, info-tech sensors and 3D systems, vol 101670T (24 April 2017)
15. Patel S, Kumar A, Yadav P, Desai J, Patil D (2017) Smartphone-based obstacle detection for visually impaired people. In: 2017 International conference on innovations in information, embedded and communication systems (ICIECS). IEEE, pp 1–3
16. Dhod R, Singh G, Singh G, Kaur M (2017) Low cost GPS and GSM based navigational aid for visually impaired people. *Wirel Pers Commun* 92(4):1575–1589
17. Noorithaya A, Kumar MK, Sreedevi A (2014) Voice assisted navigation system for the blind. In: International conference on circuits, communication, control and computing, pp 177–181. <https://doi.org/10.1109/CIMCA.2014.7057785>
18. Mohapatra BN, Mohapatra RK, Panda P (2019) Path guidance system for blind people. *Int J Open Inf Technol* 7(5):29–32
19. Borenstein J, Ulrich I (1997) The GuideCane—a computerized travel aid for the active guidance of blind pedestrians. In: Proceedings of international conference on robotics and automation, vol 2, pp 1283–1288. <https://doi.org/10.1109/ROBOT.1997.614314>
20. Farcy R, Bellik Y (2002) Locomotion assistance for the blind. In: Keates S, Langdon P, Clarkson PJ, Robinson P (eds) Universal access and assistive technology. Springer, London. https://doi.org/10.1007/978-1-4471-3719-1_27

21. Zhan FB, Noon CE (1998) Shortest path algorithms: an evaluation using real road networks. *Transp Sci* 32(1):65–73
22. Loomis JM, Golledge RG, Klatzky RL (2001) GPS-based navigation systems for the visually impaired. *Fundam Wearable Comput Augment Reality* 429:46
23. Gawari H, Bakuli M (2014) Voice and GPS based navigation system for visually impaired. *Int J Eng Res Appl* 4(4):2248–9622
24. Nisha KK, Pruthvi HR, Hadimani SN, Reddy GR, Ashwin TS, Domanal S (2014) An android gps-based navigation application for blind. In: *Proceedings of the 7th international symposium on visual information communication and interaction*, pp 240–241
25. Yelamarthi K, Haas D, Nielsen D, Mothersell S (2010) RFID and GPS integrated navigation system for the visually impaired. In: *2010 53rd IEEE international midwest symposium on circuits and systems*. IEEE, pp 1149–1152

Proposal of Criteria for Selection of Oil Tank Maintenance Companies at Transpetro Through Multimethodological Approaches



Eric Bremm De Carvalho, Miguel Ângelo Lellis Moreira, Adilson Vilarinho Terra, Carlos Francisco Simões Gomes, and Marcos dos Santos

Abstract The present study consists of an analysis based on the Multicriteria Gaussian AHP method on the definition of better potential companies for maintenance and rehabilitation works in atmospheric oil tanks and other static equipment in specific cases. To define the analysis criteria for the method and to determine the fundamental objectives of this work and the relationships between them, we implemented the approach to decision-making known as VFT, or Value Focused Thinking. This technique brought us a more holistic view of the problem and assisted in developing a systematized process for decision-making that can be based on criteria other than the economic one.

Keywords Maintenance of oil tanks · VFT approach · Gaussian AHP method · Multicriteria methods

1 Introduction

Significant engineering works are commonly carried out by companies specializing in this type of service, especially when it comes to the maintenance and rehabilitation of tanks and other static equipment that store oil and its derivatives [1].

Given the specificity, complexity and dangerousness of the services involved, not even a large part of the companies that claim to be able and operating in the market currently end up satisfactorily executing such ventures [2].

E. Bremm De Carvalho (✉) · M. Ângelo Lellis Moreira (✉) · A. Vilarinho Terra (✉) · C. Francisco Simões Gomes
Fluminense Federal University, Niterói, RJ 24210-346, Brazil
e-mail: adilsonvterra@gmail.com

M. Ângelo Lellis Moreira · A. Vilarinho Terra · M. dos Santos
Naval Systems Analysis Center, Rio de Janeiro, RJ 20091-000, Brazil

M. dos Santos
Military Institute of Engineering, Urca, RJ 22290-270, Brazil

Despite knowing the maintenance process, Brazilian public companies in this field, such as TRANSPETRO and Petrobras [3], end up resorting to such services because they do not have specialized labor in their technical staff primary end activity of the companies [4].

It is also true that there are limitations arising from law 13.303/2016, which provides for the contracting of services by public companies and mixed-economy companies. However, certain types of hiring allow the selection of bidders based on technical and non-purely financial criteria [5].

In this sense, the present study aims to provide, through the application of a decision-making approach known as VFT, a selection of criteria considered fundamental by technical staff to evaluate candidate companies for a given maintenance service in a more holistic way. To choose a specific company, we applied the multicriteria analysis method known as Gaussian AHP [6].

2 Basic Theoretical Framework

This section aims to list and explain the theoretical elements on which this work could be constructed. Initially, we addressed an overview of the various concepts related to the maintenance of atmospheric oil tanks and derivatives to explain the more general objective of this analysis. With such elements brought to light, this framework defines the Value Focused Thinking (VFT) [7] decision aid tools, as well as the multicriteria analysis methodology known as Gaussian AHP [6] to be used to determine the best company analyzed.

2.1 *Maintenance and Contracting of Tasks in Transpetro*

The maintenance and rehabilitation of tanks and other static equipment is a common and necessary practice in oil companies such as Petrobras and TRANSPETRO, the largest companies in the field within Brazil [8]. The latter and object of study of this article currently has a tanking park of 543 tanks spread over 20 waterway terminals and 27 land terminals, constantly needing to rehabilitate its equipment throughout Brazil.

Although the time elapsed between tank maintenance is at least ten years and may receive additional deadlines according to the deterioration control measures implemented, the amount of tanks makes the demand for maintenance practically constant in the company's reality [9].

Considering that such maintenance is not performed by the company itself, except for more straightforward cases, the company has to seek each new bidding process engineering company specialized in this type of work [10]. Despite the complexity

and barriers of entry to the Brazilian scenario, after Operation Car Wash, large companies that had irregular contracts with TRANSPETRO were prevented from trading with the company, thus hindering choosing notified companies [11].

Although the criterion of judging proposals by cost is mostly chosen, law 13.303/2016 provides for other forms of contracting such as direct contracts in emergency cases or even through other judgment criteria in bids such as the best combination of technique and price [12]. In addition, it is also worth noting that the company has been seeking selection and qualification methods for these companies to avoid unqualified companies.

2.2 *The VFT Approach*

VFT, or Value Focused Thinking, is a tool that serves as an auxiliary way to make decisions in general. Keeney [7] states that such an approach has a value-oriented bias, unlike other approaches that may focus on pre-existing alternatives. In other words, the implementation of this process aims to identify which values a decision-maker should take as parameters for this decision-making, not limited to the analysis of these alternatives [13].

In a statement, the VFT can be considered composed of two activities, the first referring to the determination of what is desired and the second referring to the discovery of how to obtain it [7].

The most significant benefits of using this approach are creating better alternatives to any decision-making problems and understanding decision-making problems in the form of opportunities to decide [14].

According to Keeney [15], concerning its fundamental steps, this method has four auxiliary procedures for its use: identifying the objectives, classifying the objectives, creating alternatives, and examining objectives to identify decision-making opportunities. It is worth mentioning that these objectives should be the expression of what one wishes to achieve and, therefore, must be characterized by their contexts and objects and be separated according to their definitions, that is, whether they are fundamental, strategic objectives or means.

Finally, to stimulate creativity at this stage of the process and generate a list with all the objectives of the study, question techniques can be applied to specialists that evidence what the relevant factors are, their redundancies and their interrelationships. Such a procedure is generally explicit in Table 1.

Table 1 Techniques for identifying objectives

| |
|---|
| 1. <i>A wish list.</i> What do you want? What do you value? What should you want? |
| 2. <i>Alternatives.</i> What is a perfect alternative, a terrible alternative, some reasonable alternative? What is good or bad about each? |
| 3. <i>Problems and shortcomings.</i> What is wrong or right with your organization? What needs fixing? |
| 4. <i>Consequences.</i> What has occurred that was good or bad? What might occur that you care about? |
| 5. <i>Goals, constraints, and guidelines.</i> What are your aspirations? What limitations are placed upon you? |
| 6. <i>Different perspectives.</i> What would your competitor or your constituency be concerned about? At some time in the future, what would concern you? |
| 7. <i>Strategic objectives.</i> What are your ultimate objectives? What are your absolutely fundamental values? |
| 8. <i>Generic objectives.</i> What objectives do you have for your customers, employees, shareholders, and yourself? What are essential environmental, social, economic, or health and safety objectives? |
| 9. <i>Structuring objectives.</i> Follow means-ends relationships: why is that objective important, how can you achieve it? Use specification: what do you mean by this objective? |
| 10. <i>Quantifying objectives.</i> How would you measure the achievement of this objective? Why is objective A three times as important as objective B? |

2.3 The Gaussian AHP Method

Multicriteria methods have been used in several complex real problems, such as in [16–18]. Among the main methods, the AHP [19], or Analytic Hierarchy Process, is a multicriteria approach developed by the American Tomas Saaty to serve as an instrument for decision-making. Such a method is widely known and used in Academia for having great application flexibility, including highly complex problems [20].

In general terms, this method serves as a measurement tool by staggering the analysis of paired comparisons based on expert decision-makers' judgments. Thus, through a hierarchization of the global system, the method analyzes the relative importance of the established criteria and provides, based on this importance, a ranking for the outlined alternatives [21].

The Gaussian AHP method, the object of this study, was proposed by Santos, Costa and Gomes [6] and presented a new approach concerning the original method by presenting a sensitivity analysis based on the Gaussian factor. In other words, with such an approach, the weighting of relative importance between the criteria is given by the existing values for each of these criteria and not the pairwise comparison done in the traditional AHP method. This approach has the primary advantage of reducing the decision maker's cognitive effort to compare the listed criteria.

Table 2 below presents the steps listed for the application of the Gaussian AHP method in 7 macro steps:

Table 2 Steps of the Gaussian AHP method

| |
|---|
| Step 1—Establish the decision matrix |
| Step 2—Calculate the average of the alternatives |
| Step 3—Calculate the standard deviation of the alternatives on criteria |
| Step 4—Calculate de Gaussian factor for each criterion |
| Step 5—Product of Gaussian factor by the decision matrix |
| Step 6—Normalization of results |
| Step 7—Obtaining the new ranking |

3 Methodology

At the end of the theoretical framework, to organize the method of analysis to be proposed, we first established a flowchart with the steps of the actions of this research to serve as a guide throughout its development. Additionally, Fig. 1, presenting this flowchart of activities, also shows how we ordered and structured this development. In other things, we first interviewed a specialist working in Tank Maintenance at TRANSPETRO. Through predetermined questions, we then understood the elements and values of the decision-maker for his decision [22].

With these elements, it was then possible to establish the fundamental objectives of this study and deduce the network of half-end objectives. After this stage, these objectives could then be transformed into decision criteria for the final comparison between alternative companies through the AHP method.

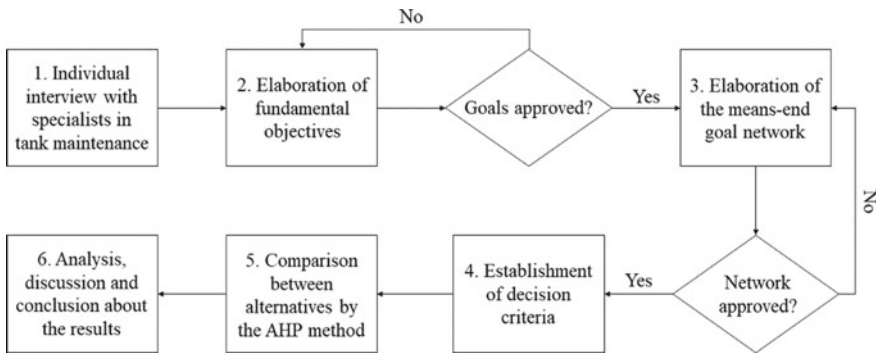


Fig. 1 Workflow of study actions

Table 3 Questionnaire for structuring values and objectives

| |
|---|
| 1. What do you aim for in terms of tank maintenance? |
| 2. What ideally should a maintenance company provide if there is no restrictions? |
| 3. Which are the weak point of the companies? |
| 4. What can be improved? |
| 5. What are the limitations of hiring these companies? |
| 6. What is expected at the end of the construction? |
| 7. How could that be achieved? |
| 8. What is the relative importance of selected goals? |

4 Application and Discussion

In this chapter, we applied the tools described in the theoretical framework and those in the methodology.

4.1 Application of VFT Method

With the data collected in the interview process, questions derived from the VFT approach and which are detailed in Table 3 were asked, the values were collected, and the objectives were set.

With the answers to the above questions and the debate itself on the main points sought in hiring companies, we elaborated the hierarchy of fundamental objectives, highlighted in Table 4, and the network of mid-end objectives (Fig. 2). It is worth mentioning that the experts validated these analysis instruments as compliant and appropriate for the study in question.

Figure 2 illustrates the conclusion of the relationship process between the different types of objectives and degrees to meet what was considered as a strategic objective of the present study, that is, “To maximize efficiency and effectiveness in reforms of oil tanks and other static equipment”.

Table 4 Hierarchy of objectives

| |
|---|
| Maximize efficiency and effectiveness in the oil tank and other static equipment retrofits (strategic goal) |
| 1. Reduce construction time |
| 1.1. Increase amount of HM |
| 1.2. Reduce mobilization time |
| 1.3. Maximize service time |
| 2. Minimize construction costs |
| 2.1. Use sources with efficiency |
| 2.2. Plan purchase of supplies in advance |
| 3. Delivery services with quality and lawfulness |
| 3.1. Obtain qualified labour |
| 3.2. Proceeding field actions |
| 3.3. Carry out effective planning |
| 3.4. Act according to the law |
| 4. Reduce recordable accidents |
| 4.1.. Invest in safety equipment |
| 4.2 Increase number of security inspectors |
| 4.3. Invest in safety culture |

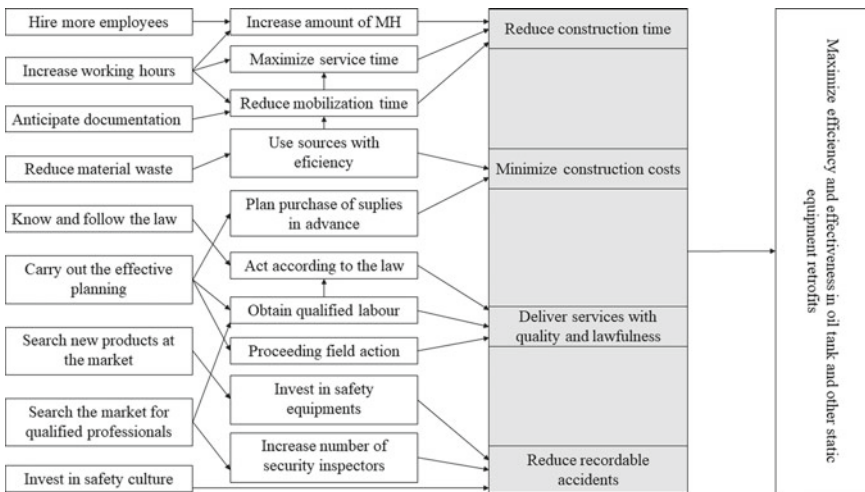


Fig. 2 Means-end-goals

It is also worth mentioning that, through the values and objectives established with the decision-making company, the fundamental objectives of this work could also be defined. Thus, it is understood that reducing the work time, reducing recordable accidents, delivering the service with quality, and minimizing the cost of reforms are the basis for achieving the strategic objective.

Finally, we established the criteria to be used for decision-making. Such relationship is explicit in Table 5:

Table 5 Definition of criteria based on fundamental objectives

| | |
|---|---------------------------------|
| Maximize efficiency and effectiveness in the oil tank and other static equipment retrofits (strategic goal) | Criteria |
| 1. Reduce construction time 1.1. Increase amount of HM 1.2. Reduce mobilization time 1.3. Maximize service time | C1. Scope runtime |
| 2. Minimize construction costs 2.1. Use sources with efficiency 2.2. Plan purchase of supplies in advance | C2. Total costs of construction |
| 3. Delivery services with quality and lawfulness 3.1. Obtain qualified labour 3.2. Proceeding field actions 3.3. Carry out effective planning 3.4. Act according to the law | C3. Integrity risk degree |
| 4. Reduce recordable accidents 4.1. Invest in safety equipment 4.2. Increase number of security inspectors 4.3. Invest in safety culture | C4. Recorded accident rate |

4.2 Application of the Gaussian AHP Method

It is then possible to define alternative companies to apply the method with the established criteria. It should be noted that to maintain the confidentiality of information restricted to the company, fictitious values and names were created for the companies involved to exemplify the method’s applicability to possible real cases.

At first, we applied values to alternative companies, considered here as “Careless”, “Accelerated,” and “Slow”. Values are shown in Table 6 below. It should also be noted that, according to the nature of the variables, the matrix used for the calculation itself had to use the inverse of the assigned values, since the higher the cost, the accident rate, the construction time and the degree of risk, the worse would be the result of the company in question. Finally, it is also highlighted that such values need to be normalized, that is, placed under the same quantitative basis of comparison, to homogenize the different metrics of each of the criteria as expressed in Table 7.

Next, expressed in Table 8 and as established in the subsequent steps proposed by the method, each criterion’s means, standard deviations, and Gaussian factor, namely the ratio between standard deviation and mean, are then calculated.

Table 6 Matrix of alternatives by criteria

| | C1 | C2 | C3 | C4 |
|-------------|-----|--------|----|-----|
| Careless | 300 | 10.000 | 60 | 1.4 |
| Accelerated | 290 | 17.000 | 20 | 0.9 |
| Slow | 400 | 12.000 | 15 | 0.7 |

Table 7 Normalized matrix of alternatives by criteria

| | C1 | C2 | C3 | C4 |
|-------------|-------|-------|-------|-------|
| Careless | 0.359 | 0.413 | 0.125 | 0.220 |
| Accelerated | 0.372 | 0.243 | 0.375 | 0.341 |
| Slow | 0.269 | 0.344 | 0.500 | 0.439 |

Table 8 Mean, standard deviations and Gaussian factor of the criteria

| | C1 | C2 | C3 | C4 |
|--------------------|--------|--------|--------|--------|
| Mean | 0.3330 | 0.3330 | 0.3330 | 0.3330 |
| Standard deviation | 0.0557 | 0.0855 | 0.1909 | 0.1099 |
| Gaussian factor | 0.1672 | 0.2565 | 0.5728 | 0.3299 |
| Norma. G factor | 0.1260 | 0.1934 | 0.4317 | 0.2487 |

Table 9 Results of the Gaussian AHP Method

| AHP-Gaussian | Company | Rank (°) |
|--------------|-------------|----------|
| 0.234 | Careless | 3 |
| 0.341 | Accelerated | 2 |
| 0.426 | Slow | 1 |

As a conclusion of the Gaussian AHP method, the normalized matrix (Table 7) is then multiplied by the normalized Gaussian factor in each of the listed criteria, resulting in the relative ranking values among the companies highlighted in Table 9.

5 Conclusion

The present study could show new ways to approach a contract based on historical data of companies already consolidated in the market and beyond the financial criterion, already widely used in public and mixed company companies, which presents flaws inherent to a complex decision-making process.

Although there are legal modalities and limitations for hiring companies according to pre-established criteria, some other hiring modalities allow the creation of performance metrics such as those created here. Similarly, the use of the system proposed here may serve to contract itself and classify existing companies and aid in the qualification and qualification stages.

In this sense, the implementation of the VFT and Gaussian AHP approaches proved to be possibly valuable tools for TRANSPETRO to realize specific contracts. According to the objectives established by experts in the study in question, it was possible to create a metric that could clearly distinguish the most important values for the maintenance of tanks within the company. In the illustrative case, the

company “Slow”, although with longer construction time, proved to be the best of the options since it meets the criteria of safety and integrity, no longer more valuable to TRANSPETRO.

References

1. Zio E (2018) The future of risk assessment. *Reliab Eng Syst Saf* 177:176–190
2. Chesbrough H, Tucci CL (2020) The interplay between open innovation and lean startup, or, why large companies are not large versions of startups. *Strat Manag Rev* 1:277–303
3. Duarte L, Nabarro W (2021) Financial circles and oil circuit: financial instruments for investment in Petrobras’ activities and suppliers. *GEOUSP* 25
4. Massi E, Nem Singh J (2018) Industrial policy and state-making: Brazil’s attempt at oil-based industrial development. *Third World Q* 39:1133–1150
5. Goulielmos AM (2018) The unresolved issues in maritime economics. *Mod Econ* 9:1687–1715
6. dos Santos M, de Araújo Costa IP, Gomes CFS (2021) Multicriteria decision-making in the selection of warships: a new approach to the AHP method. *Int J Anal Hierarchy Process* 13. <https://doi.org/10.13033/ijahp.v13i1.833>
7. Keeney RL (1992) *Value-focused thinking: a path to creative decision-making*. Harvard University Press, London
8. Giotitsas C, Nardelli PHJ, Williamson S, Roos A, Pournaras E, Kostakis V (2022) Energy governance as a commons: engineering alternative socio-technical configurations. *Energy Res Soc Sci* 84:102354
9. de Almeida AG, Vinnem JE (2020) Major accident prevention illustrated by hydrocarbon leak case studies: a comparison between Brazilian and Norwegian offshore functional petroleum safety regulatory approaches. *Saf Sci* 121:652–665. <https://doi.org/10.1016/j.ssci.2019.08.028>
10. Harris F, McCaffer R, Baldwin A, Edum-Fotwe F (2021) *Modern construction management*. Wiley & Sons
11. Evans P (2018) State structures, Government-business relations, and economic transformation. In: *Business and the state in developing countries*, pp 63–87. Cornell University Press
12. Clough RH, Sears GA, Sears SK, Segner RO, Rounds JL (2015) *Construction contracting: a practical guide to company management*. Wiley & Sons
13. de Araújo Costa IP, Sanseverino AM, dos Santos Barcelos MR, Belderrain MCN, Gomes CFS, dos Santos M (2021) Choosing flying hospitals in the fight against the COVID-19 pandemic: structuring and modeling a complex problem using the VFT and ELECTRE-MO_r methods. *IEEE Lat Am Trans* 19:1099–1106. <https://doi.org/10.1109/TLA.2021.9451257>
14. Morais DC, Alencar LH, Costa APCS, Keeney RL (2013) Using value-focused thinking in Brazil. *Pesquisa Operacional* 33:73–88. <https://doi.org/10.1590/S0101-74382013000100005>
15. Keeney RL (1996) Value-focused thinking: identifying decision opportunities and creating alternatives. *Eur J Oper Res* 92:537–549. [https://doi.org/10.1016/0377-2217\(96\)00004-5](https://doi.org/10.1016/0377-2217(96)00004-5)
16. De Almeida IDP, de Pina Corriça JV, de Araújo Costa AP, de Araújo Costa IP, do Nascimento Maêda SM, Gomes CFS et al (2021) Study of the location of a second fleet for the Brazilian Navy: structuring and mathematical modeling using SAPEVO-M and VIKOR methods. *ICPR-Am 2020 Commun Comput Inf Sci [Internet]* 1408:113–24. <https://link.springer.com/>, https://doi.org/10.1007/978-3-030-76310-7_9
17. Gomes CFS, dos Santos M, de Souza de Barros Teixeira LF, Sanseverino AM, dos Santos Barcelos MR (2020) SAPEVO-M: a group multicriteria ordinal ranking method. *Pesqui Operacional [Internet]* 40. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-74382020000100212&tlng=en
18. Moreira MÂL, de Araújo Costa IP, Pereira MT, dos Santos M, Gomes CFS, Muradas FM (2021) PROMETHEE-SAPEVO-M1 a hybrid approach based on ordinal and cardinal inputs: multicriteria evaluation of helicopters to support Brazilian Navy operations. *Algorithms [Internet]* 14(5):140. <https://www.mdpi.com/1999-4893/14/5/140>

19. Saaty TL (1977) A scaling method for priorities in hierarchical structures. *J Math Psychol* 15:234–281
20. de Araújo Costa IP, Basílio MP, do Nascimento Maêda SM, Rodrigues MVG, Moreira MÂL, Gomes CFS, dos Santos M (2021) Bibliometric studies on multi-criteria decision analysis (MCDA) applied in personnel selection. *Front Artif Intell Appl* 341. <https://doi.org/10.3233/faia210239>
21. do Nascimento Maêda SM, de Araújo Costa IP, Castro Junior MAP, Fávero LP, de Costa AP, de Corriça JVP, Gomes CFS, dos Santos M (2021) Multi-criteria analysis applied to aircraft selection by Brazilian Navy. *Production* 31:1–13. <https://doi.org/10.1590/0103-6513.20210011>
22. de Araújo Costa IP, Moreira MÂL, de Araújo Costa AP, de Souza de Barros Teixeira LFH, Gomes CFS, dos Santos M (2021) Strategic study for managing the portfolio of IT courses offered by a corporate training company: an approach in the light of the ELECTRE-MOr multicriteria hybrid method. *Int J Inf Technol Decis Mak* 1–29. <https://doi.org/10.1142/S0219622021500565>

A Deep Learning-Based Approach with Semi-supervised Level Set Loss for Infant Brain MRI Segmentation



Minh-Nhat Trinh, Van-Truong Pham, and Thi-Thao Tran

Abstract Nowadays, numerous deep learning techniques for segmenting medical images have been proposed, with excellent outcomes and the valuable success of machine learning. However, most models use supervised methods while others use unsupervised methods, with less effective outcomes than supervised learning methods. Therefore, this paper introduces a novel level set loss function for unsupervised tasks and incorporates it with the Active Contour loss for supervised tasks. Besides, since previously introduced deep learning models generate less accurate results due to the intensity inhomogeneity issues and the often presence of low-intensity contrast tissues in infant brain segmentation, we propose a new convolutional neural network model to alleviate this problem. Instead of binary segmentation, our proposed loss entitles our model to segment multiple classes with promising outcomes. The proposed technique is utilized to segment neonatal brain magnetic resonance images into four non-overlap regions. The iSeg-2017 challenge, which offers a collection of neonatal brain magnetic resonance images from different sites, is used to evaluate our proposed process. The experiment demonstrates that our new loss function achieves promising results among the 21 participating teams. This illustrates the effectiveness of our technique in multiclass medical image segmentation.

Keywords Infant brain segmentation · Semi-supervised learning · Unsupervised learning

1 Introduction

In computer vision field, image segmentation is immensely crucial [1]. It divides an image into areas based on pixel features to find boundaries or objects to analyze the image more effortless. Computer vision in general and image segmentation tasks in particular, have significantly been developed thanks to the advanced achievements of deep learning. Therefore, numerous evolved segmentation methods are more

M.-N. Trinh · V.-T. Pham · T.-T. Tran (✉)

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology,
No.1 Dai Co Viet, Hanoi, Vietnam
e-mail: thao.tranthi@hust.edu.vn

powerful with deep learning models, notably convolution neural networks (CNNs), with outstanding outcomes [2].

In the medical field, image segmentation is an essential stage in the diagnostic method and various applications. While the manual segmentation process always delivers precise and dependable results, it squanders lots of time and demands professional understanding from skilled experts [3]. Therefore, automatic segmentation techniques have become a demanding topic [4, 5], guaranteeing accurate results, and diminishing inference time. In addition, they were developed in a number of areas, which also attracted many researchers to develop medical imaging modalities, such as CT [6], dermoscopic images [7], X-rays, and magnetic resonance imaging (MRI). Although many breakthrough approaches to image segmentation have been introduced like many other subjects, the deep learning methodologies have shown promising progress in achieving state-of-the-art outcomes [8–11]. Generally, the effectiveness of these methods derives from the enhancement of loss functions or the model construction on the basis of U-net architecture. Especially with balanced data sets, the loss of Cross-Entropy frequently works better, while the Dice and Focal loss tackle highly-imbalanced data problems effectively [12].

Many convolutional neural networks (CNNs) [13–15] are often used in deep learning-based approaches and have shown highly effectiveness in a variety of applications [16]. For example, the CNN architecture in the brain tissue segmentation was introduced in [17] on the multimodal MRI, 2D U-Net with excellent performance for the segmentation of medical images was developed in [9], and a CNN utilizing multiscale features was presented in [18] for both infant and adult MR images. The effectiveness of deep learning-based techniques in medical image segmentation is undeniable. However, these studies mainly concentrate on processing 2D image segmentation problems where each slice was individually segmented. Therefore, developing a CNN for 3D image segmentation of volumetric data, such as brain MR images, is essential where the model can investigate extra spatial information of neighboring slices.

Another issue is that it is often trapped in local minima as a result of the non-convexity feature of loss functions and anomalies of the training database. Unluckily, medical images frequently contain anomalies such as occlusion and inhomogeneity of intensity [19]. Especially for infant brain MR images, the low contrasts between tissue, excessive noise, and ongoing white matter' myelination [20] often lead to misclassification and reduced segmentation algorithms' accuracy. Hence, many segmentation approaches [19, 21–23] have been developed to mitigate the effects of intensity inhomogeneity problems. One of them proposed a region-based model using multiple level set functions to raise the number of segmentation classes, which extends from the two-phase Chan-Vese model [24]. Besides, these region-based methods utilize unique region characteristics to control the motion of the active contour. However, these methods have additional limits depending on the initial conditions and frequently experience boundary outflow problems [19].

To tackle the issues mentioned above and inspired by Boah Kim et al. [25], level set method [26] and active contour loss in [27], we propose a novel semi-supervised level set loss function to handle the intensity inhomogeneity issues for newborn

brain MR images and incorporate it into our CNN model. Hence, this enables us to effortlessly train our model in an end-to-end disparity manner and efficiently segment multiclass at one and the same time. Our proposed approach is tested on the iSeg-2017 challenge [28], where a community of 6-month children from various locations, including several inhomogeneity images, have different protocols/scanners. Finally, we also show our experimental results in comparison with other state-of-the-arts.

In this work, our main contributions are:

- Proposing a loss function based on the Mumford-Shah loss function, active contour and level set method in image processing techniques to handle intensity inhomogeneity problems for the infant brain segmentation task.
- Building a learning model on the basis of U-net structural design can extract multi-scale contextual information with a small computational cost.
- Achieving promising results on the iSeg-2017 test dataset.

The following are the remaining parts of this paper: The suggested technique is described in detail in Sect. 2. Then, Sect. 3 represents experimental information, including comparisons with state-of-the-art approaches. Eventually, we summarize this study and investigate the potential applications in Sect. 4.

2 Materials and Method

2.1 Materials

We use TensorFlow 2.3.0 [29] and train the deep learning model with Tesla P100 16 GB GPU. Besides, we also utilize OpenCV 3.4 and NumPy [30] Library mainly to process the dataset. For multiclass segmentation, the ground truth is encoded using a one-hot vector.

2.2 Proposed Coefficient of Variation Active Contour Loss Function

$\Omega \subset \mathbb{R}^2$ denotes the image I domain. In [25], the softmax layer in the neural network may be utilized as a differentiable approximation to the characteristic function, allowing the Mumford-Shah loss to minimize directly. To be specific, the n_{th} feature map softmax output from a CNN, which has N classes, is denoted as follows:

$$P_n(\mathbf{u}; \theta) = \frac{e^{b_n(\mathbf{u}; \theta)}}{\sum_{i=1}^N e^{b_i(\mathbf{u}; \theta)}}, \quad n = \overline{1, N} \quad (1)$$

where, $\mathbf{u} \in \Omega$, $b_i(\mathbf{u}; \theta)$ represents the prediction at pixel \mathbf{u} of the last layer before applying the softmax activation, and θ refers to the trainable network parameters. From the formula (1), the value of $P_n(\mathbf{u}; \theta)$ is probability of the voxel \mathbf{u} belongs to the n_{th} class. In [25], authors proposed the Mumford-Shah functional for unsupervised task:

$$L_{MScnn}(I; \theta) = \sum_{n=1}^N \int_{\Omega} |I(\mathbf{u}) - c_n|^2 P_n(\mathbf{u}; \theta) d\mathbf{u} + \beta \sum_{n=1}^N \int_{\Omega} |\nabla P_n(\mathbf{u}; \theta)| d\mathbf{u} \quad (2)$$

where β is a hyperparameter and $c_n = \frac{\int_{\Omega} I(\mathbf{u}) P_n(\mathbf{u}; \theta) d\mathbf{u}}{\int_{\Omega} P_n(\mathbf{u}; \theta) d\mathbf{u}}$ denotes as the average voxel value of the n_{th} class.

Equation (2) is differentiable by θ . Consequently, backpropagation during the training can minimize the loss function $L_{MScnn}(\theta; \mathbf{u})$. This loss function is also used for unsupervised image segmentation tasks, while the supervised loss function is not accessible for these problems. Moreover, the loss function enables our model to segment based on voxel value distribution, which can be boosted when ground truths are provided. Nevertheless, the Mumford-Shah loss requires each component to have a comparable value with a contour length regularization to separate different regions commonly occur if voxel values differ considerably from each area due to the intensity of inhomogeneity.

This aspect is alleviated by adding information about variation on a little small estimate [19] discussed above. Using the ideal of coefficient of variation (CoV) [31, 32] and inspired in [26], we proposed a coefficient of variation level set loss function to solve the intensity inhomogeneity issues:

$$L_{CoV}(I; \theta) = \sum_{n=1}^N \int_{\Omega} \left| \frac{I(\mathbf{u}) - CoV_n}{CoV_n} \right|^2 P_n(\mathbf{u}; \theta) d\mathbf{u} + \beta \sum_{n=1}^N \int_{\Omega} |\nabla P_n(\mathbf{u}; \theta)| d\mathbf{u} \quad (3)$$

where $CoV_n = \frac{\int_{\Omega} I(\mathbf{u})^2 P_n(\mathbf{u}; \theta) d\mathbf{u} + \varepsilon}{\int_{\Omega} I(\mathbf{u}) P_n(\mathbf{u}; \theta) d\mathbf{u} + \varepsilon}$, and ε is a small positive parameter preventing division by zero. The CoV value is greater in regions with edges than in areas with uniformity [32, 33]. Therefore, a larger value implies that voxels belong to the edges, while a low value shows that the voxels belong to a uniform area. Undoubtedly, the characteristics of CoV [32, 33] imply that it can be utilized as both an area detector and a fitting term.

To improve the performance of the training phase, we leverage the active contour loss function in [27] for segmenting multiclass simultaneously as below:

$$L_{acms}(G; \theta) = -\frac{1}{|\Omega|} \sum_{n=1}^N \int_{\Omega} \left((a_{1n} - G_n(\mathbf{u}))^2 \log(P_n(\mathbf{u}; \theta) + \varepsilon) + (a_{2n} - G_n(\mathbf{u}))^2 \log(1 - P_n(\mathbf{u}; \theta) + \varepsilon) \right) d\mathbf{u} \quad (4)$$

where $\log(\cdot)$ is natural logarithm with base value is e , $G_n(\mathbf{u})$ is one-hot vector of the semantic label located at \mathbf{u} for n_{th} class, a_{2n} is average voxel value of area n_{th} in ground truth and a_{1n} is the average voxel of the background. Besides, total number of voxels of input is defined as $|\Omega|$.

To summarize, we introduce the semi-supervised level set loss for medical image segmentation:

$$L_{semi_CoV}(I, G; \theta) = L_{acms}(G; \theta) + \alpha L_{CoV}(I; \theta) \quad (5)$$

where α is a hyperparameter to adjust the weight of the unsupervised loss.

2.3 Network Architecture

We customize the U-net [9] to evaluate our proposed loss function as our infrastructure segmentation system in this part. The general construction of our model is demonstrated in Fig. 1.

Our model construction includes two parts: encoder blocks and decoder blocks. The task of an encoder block is to extract helpful feature maps from the image and transfer the skip connections required to the corresponding decoder block. The input image is initially standardized by a min-max scaler then passed into the encoder block 1. Each encoder block contains four discrete downsample blocks, which is inspired from [34]. The j th encoder block comprises $R = 5 - j$ small downsample blocks and upsample blocks. For example, as demonstrated in Fig. 2, in 1st encoder block, the feature maps are fed into 4 small downsample blocks before going through a ConvBlock. After going through 4 upsample block, the output feature maps would be added with the input feature maps. Finally, the resolution of the output would be decrease by 2 times through the downsample block.

With a U-Net-like structure, our proposed encoder block learns to encode and extract multi-scale contextual information from the input feature maps. The deeper the encoder block (the larger R), the more downsample operations, the richer global and local features and the more range of receptive fields are increased. Multi-scale features could be extracted from different input spatial resolutions then be encoded into high-resolution feature maps by consecutive upsample blocks. Subsequently, they are fused with input feature maps by summation in order to mitigate the gradient vanishing phenomenon. Furthermore, it is noticeable that because almost operations are employed on the downsampled feature maps, the computational cost of these encoder blocks is tiny. Hence, this design enables the network to grow deeper and achieve higher resolution without dramatically increasing memory and processing costs.

Each ConvBlock comprises 3D convolution layers (because the input is 3D MR images), followed by a batch normalization layer and Swish activation [35]. After that, a Convolutional Block Attention Module (CBAM) [36] block is concatenated

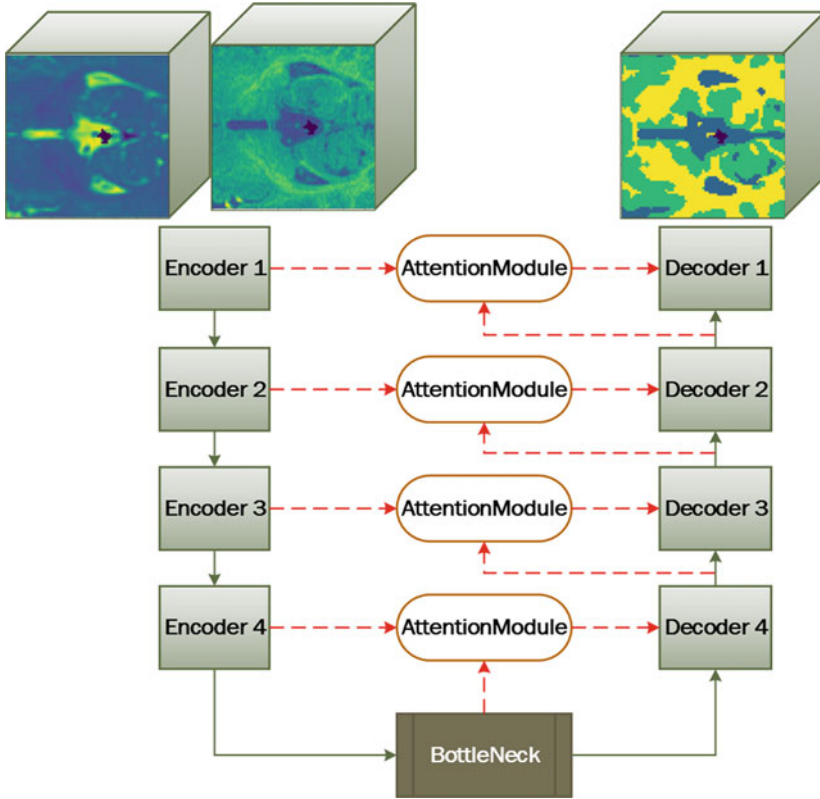


Fig. 1 Overall diagram of our model architecture

to the block above in order to reduce the number of parameters of the model. This CBAM block simply adds to each channel parameters for a convolutional block so that the network can change the weighting with a small cost calculation of each feature map [37]. Eventually, a modified DropBlock for 3D image segmentation in [38] with a keep probability rate of 0.95 is the final procedure of the block above. This DropBlock in [38] is proved that be very effective to avoid overfitting problems in CNNs.

Regards the skip connection path, the outputs of four encoder blocks are symbolized as $B_1, B_2, B_3, B_4, B_i \in R^{H_i \times W_i \times D_i \times C_{si}}$ with $i = \overline{1, 4}$ and H_i, W_i, D_i are height, width, and depth respectively. Subsequently, B_i with $i = \overline{1, 4}$ go through an attention module to aggregate more precise contextual information before attaching with the output of the decoder block i . The general architecture of the attention module is described in Fig. 3. Before using this concatenation, an attention module entitles the network to put more weight on the valuable features of the skip layers and reduce noise, especially in brain images. The direct connection will focus on some specific parts of the input feature map instead of feeding into all features. The

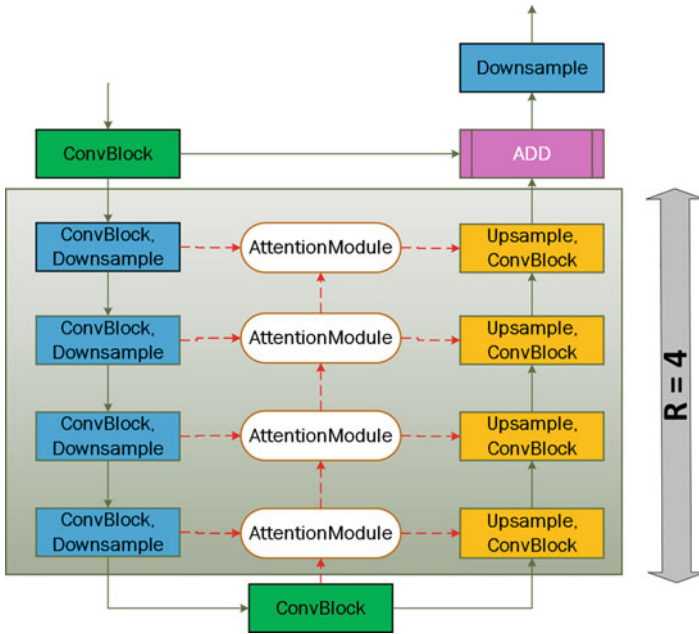


Fig. 2 The first encoder block

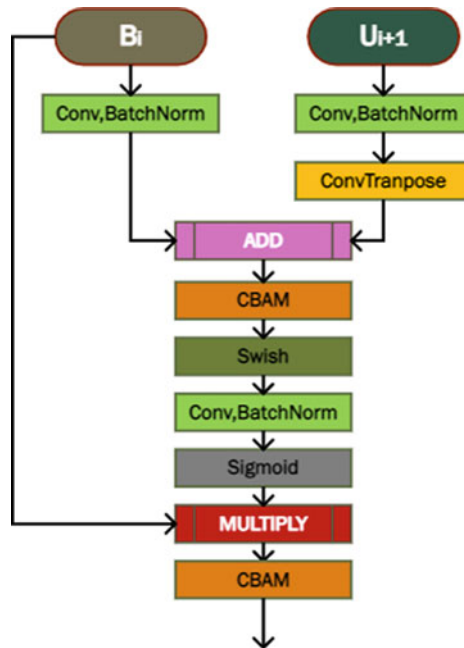


Fig. 3 Attention module

attention distribution is thus enhanced to only hold the important parts by the skip connection feature map.

About the decoder path, outputs of four upsample blocks and the bottle neck are defined as U_1, U_2, U_3, U_4, U_5 $U_i \in R^{H_i \times W_i \times D_i \times C_{di}}$, with $i = \overline{1, 5}$. Initially, U_{i+1} and B_i are passed into the attention module, which returns B'_i , which has the same resolution as feature map B_i . Subsequently, B'_i and U_{i+1} are concatenated into $[U_{i+1}, B'_i]$, with $i = \overline{1, 4}$, then are passed into each upsample block to yield U_i .

3 Experimental Results

3.1 Dataset

We utilize the iSeg-2017 dataset [28] to assess the efficacy of our proposed loss function. In the iSeg-2017 challenge, MR images of 6-month infants are randomly selected from the research of the Baby Connectome Project. The iSeg-2017 dataset contains of ten (1–10) subjects with manual labels for training and 13 subjects (11–23) for testing without ground truth labeling. For training process, we use 10 subjects for the training and validation. Each subject contains two MR image modalities which are T1-weighted with the resolution of $1 \times 1 \times 1 \text{ mm}^3$ and T2-weighted with the resolution of $1.25 \times 1.25 \times 1.95 \text{ mm}^3$, which have different sizes ranging from $144 \times 192 \times 256$ to $256 \times 256 \times 256$. The challenge requirement is the segmentation of the infant brain into four areas that are not overlapping: the cerebrospinal fluid (CSF), gray matter (GM), white matter (WM) and the background. In contrast with the manual segmentation, each participant team performed using three performance measures which have been measured using the public assessment tool [39]. Especially, the participants are not given annotations of testing subjects. Eventually, our method is evaluated on 13 subjects of the testing dataset by the organizer. The testing dataset segmentation results may be submitted up to two times for review, with only the most recent/best results being reported.

3.2 Evaluation Metrics

Dice coefficient (DICE)

To determine the overlap between the segmentation prediction and the ground truth, the DICE is denoted as below:

$$DICE = \frac{2 \times TP}{FP + FN + 2 \times TP} \quad (6)$$

where FP , FN and TP denote the number of false positives, false negatives, and true positives, respectively.

Modified Hausdorff Distance (MHD) [28]

We use the Hausdorff distance (HD) to determine the distance from ground truth to segmentation prediction. Let E and Y be the sets of vertices included inside ground truth and prediction boundaries, respectively. Formally, the Hausdorff distance is defined as:

$$HD(E, F) = \max \left\{ \max_{x \in E} \min_{y \in F} \|x - y\|, \max_{y \in F} \min_{x \in E} \|y - x\| \right\} \quad (7)$$

We utilize the 95th-percentile HD as the MHD to remove the effect of a minor subset of the outliers.

Average Surface Distance (ASD) [28]

ASD error is the third measuring metric, which is defined as:

$$ASD(E, F) = \frac{1}{2} \left(\frac{\sum_{V_i \in S_E} \min_{V_j \in S_F} d(V_i, V_j)}{\sum_{V_i \in S_E} 1} + \frac{\sum_{V_j \in S_F} \min_{V_i \in S_E} d(V_j, V_i)}{\sum_{V_j \in S_F} 1} \right) \quad (8)$$

where S_E , S_F is the surface of the ground truth and automatically segmented (the outputs of neural network), respectively. The Euclidean distance from vertex V_i to vertex V_j is demonstrated by $d(V_i, V_j)$.

3.3 Training

Initially, the intensities of the T1w and T2w images are clipped by a min–max scaler to a range of $[0, 1]$. Due to the limit usage of the Tesla P100 GPU, we randomly crop $32 \times 32 \times 32$ sub-volume samples to reduce the computational complexity before feeding them as the input to the network. Besides, we also augment data by applying histogram equalization to increase the high contrast of each input image and adding Gaussian noise to alleviate adversarial attacks from noisy images. Weights initialization approach that introduced by He et al. [40] is utilized. Then, the weights of the model are optimized by the Nadam optimization [41] with a learning rate of 10^{-3} . The total number of epochs is 500. The learning rate is multiplied a factor of 0.5 if the DICE of validation set did not increase for 7 epochs. For obtaining the above hyperparameters, the network of 9 subjects with the ground-truth mask was first trained, and validation was carried out using a single subject. Subsequently, the final model was trained using the whole training dataset, which includes 10 subjects after obtaining the necessary hyperparameters. The ultimate prediction is reached by

averaging all outputs from overlapping patches of the 8-step overlap. The training process took about two weeks, and each subject took nearly one minute for the inference process. We trained our neural network by loss function L_{semi_CoV} . Furthermore, the hyperparameters α , β , ε are set as 10^{-7} , 10^{-3} and 10^{-7} respectively.

3.4 Results

The achieved results of our final segmentation were assessed by the contest organizer. To qualitatively assess the effectiveness of our technique, we compare our results with the top five teams in Leaderboard the iSeg-2017 testing dataset. The details are demonstrated in Table 1.

By observation, our method obtains the best performance in mostly evaluation metrics on the iSeg-2017 testing dataset compared to other methods. From the visual analysis, the MHD value for all CSF, WM, and GM has a high variance, while the ASD and the DICE have a lower variance. Besides, the top 5 teams attain good results with DICE scores of all categories (all greater than 0.880). However, our approach still achieved the highest accuracy of six out of nine measures in terms of GM and WM in this dataset in comparison with other techniques. Although the MSL_SKKU team obtain the highest DICE score of CSF (0.958), our method has better results in terms of GM (0.921 DICE vs. 0.919 DICE) and WM (0.903 DICE vs. 0.901 DICE).

Figure 4 illustrates the prediction of our proposed technique on three cases (with ID 11, 15, 19 from above to below respectively) in the iSeg-2017 database. From the evaluation of the organizer, these cases have achieved promising results that their Dice score in three classes are higher than 0.91. However, we still have not looked into any previous domain awareness of infant brain MR images, such as whether the cortical thickness is within a certain range [42, 43]. Furthermore, because of the maturation, tissue contrast between GM and WM is tremendously low. Although our method obtains more accurate segmentation results in GM and WM, we ignore the comparatively high contrast between CSF and GM, which may be the reason for reducing accuracy on CSF segmentation. Hence, prior knowledge is site/scanner

Table 1 The DICE, MHD and ASD scores in testing dataset of top five teams in the iSeg-2017 challenge and our proposed approach (MHD: mm, ASD: mm)

| Method | CSF | | | GM | | | WM | | |
|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | DICE | MHD | ASD | DICE | MHD | ASD | DICE | MHD | ASD |
| nic_vicorob | 0.951 | 9.178 | 0.137 | 0.910 | 7.647 | 0.367 | 0.885 | 7.154 | 0.430 |
| TU/e IMAG/e | 0.947 | 9.426 | 0.150 | 0.904 | 6.856 | 0.375 | 0.890 | 6.908 | 0.433 |
| Bern_IPMI | 0.954 | 9.616 | 0.127 | 0.916 | 6.455 | 0.341 | 0.896 | 6.782 | 0.398 |
| LIVIA | 0.957 | 9.029 | 0.138 | 0.919 | 6.415 | 0.338 | 0.897 | 6.975 | 0.376 |
| MSL_SKKU | 0.958 | 9.072 | 0.116 | 0.919 | 5.980 | 0.330 | 0.901 | 6.444 | 0.391 |
| Our method | 0.954 | 9.252 | 0.127 | 0.921 | 5.203 | 0.326 | 0.903 | 6.291 | 0.368 |

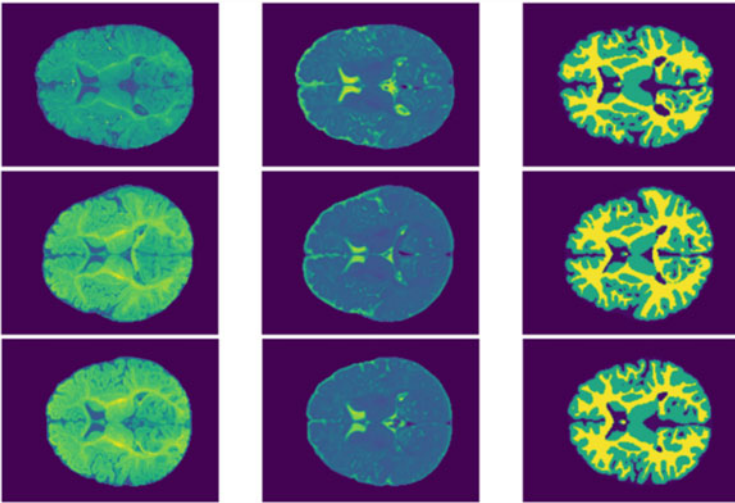


Fig. 4 Segmentation results on different slices of testing sample ID 11, 15, 19 respectively. The T1 image, T2 image and our segmentation result are presented from left to right

agnostic, which may be a key approach to take in order to better deal with this problem challenge in the future.

4 Conclusion

We have introduced a deep learning-based approach with the semi-supervised level set loss for multiclass segmentation. Our experimental results on the iSeg-2017 dataset illustrate decent improvement for 3D magnetic resonance image segmentation in comparison with state-of-the-art techniques. As an overall framework, our introduced loss function is not restricted to medical image segmentation tasks. In the future, we will focus our efforts to develop new loss functions and explore the potential of our technique in order to boost the overall accuracy of semantic segmentation tasks.

Acknowledgements This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2021-PC-005. Minh-Nhat Trinh was funded by Vingroup JSC and supported by the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Institute of Big Data, code VINIF. 2021.ThS.33.

References

1. Jahne B, Haubsecker H (2000) *Computer vision and applications: a guide for students and practitioners*. Elsevier
2. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2019) Unet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans Med Imaging* 39(6):1856–1867
3. Zhou T, Ruan S, Canu S (2019) A review: deep learning for medical image segmentation using multi-modality fusion. *Array* 3:100004
4. Bashar A (2019) Survey on evolving deep learning neural network architectures 1(2):73–82
5. Manoharan JS (2021) Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm 3(2), 83–95
6. Manoharan S (2020) Early diagnosis of lung cancer with probability of malignancy calculation and automatic segmentation of lung CT scan images 2(4), 175–186 (2020)
7. Balasubramaniam V (2021) Artificial intelligence algorithm with SVM classification using dermoscopic images for melanoma diagnosis 3(1):34–42 (2021)
8. Pham V-T, Tran T-T, Wang P-C, Lo M-T (2020) Tympanic membrane segmentation in otoscopic images based on fully convolutional network with active contour loss. *Signal Image Video Process*, 1–9
9. 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*. Springer, pp 234–241
10. Tran T-T, Tran T-T, Ninh Q-C, Bui M-D, Pham V-T (2020) Segmentation of left ventricle in short-axis MR images based on fully convolutional network and active contour model. In: *5th international conference on green technology and sustainable development*, vol 1284, pp 49–59
11. Koresh HJD, Chacko S, Periyanyagi MJPR (2021) A modified capsule network algorithm for oct corneal image segmentation 143:104–112
12. Jadon S (2020) A survey of loss functions for semantic segmentation, arXiv preprint [arXiv:2006.14822](https://arxiv.org/abs/2006.14822)
13. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
14. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 25:1097–1105
15. 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
16. Pham V-T, Tran T-T, Wang P-C, Chen P-Y, Lo M-T (2021) EAR-UNet: a deep learning-based approach for segmentation of tympanic membranes from otoscopic images. *Artif Intell Med* 115:1–12
17. Zhang W et al (2015) Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *Neuroimage* 108:214–224
18. Moeskops P, Viergever MA, Mendrik AM, De Vries LS, Benders MJ, Išgum I (2016) Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans Med Imaging* 35(5):1252–1261
19. Li C, Huang R, Ding Z, Gatenby JC, Metaxas DN, Gore JC (2011) A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans Image Process* 20(7):2007–2016
20. Wang L et al (2014) Segmentation of neonatal brain MR images using patch-driven level sets. *Neuroimage* 84:141–158
21. Chan TF, Vese LA (2000) Image segmentation using level sets and the piecewise-constant Mumford-Shah model, in *Tech. Rep. 0014*, Computational Applied Math Group, 2000: Citeseer
22. Shyu K-K, Pham V-T, Tran T-T, Lee P-L (2012) Global and local fuzzy energy based-active contours for image segmentation. *Nonlinear Dyn* 67(2), 1559–1578
23. Tran T-T, Pham V-T, Shyu K-K (2014) Zernike moment and local distribution fitting fuzzy energy-based active contours for image segmentation. *SIViP* 8(1):11–25

24. Chan TF, Vese LA (2001) Active contours without edges. *IEEE Trans Image Process* 10(2):266–277
25. Kim B, Ye JC (2019) Mumford–Shah loss functional for image segmentation with deep learning. *IEEE Trans Image Process* 29:1856–1866
26. Badshah N, Chen K, Ali H, Murtaza G (2012) Coefficient of variation based image selective segmentation model using active contours. *East Asian J Appl Math* 2(2):150–169
27. Trinh M-N, Nguyen N-T, Tran T-T, Pham V-T (2021) A Semi-supervised deep learning-based approach with multiphase active contour loss for left ventricle segmentation from CMR images. In: *Proceedings of Third International Conference on Sustainable Computing, 2021*: Springer
28. Wang L et al (2019) Benchmark on automatic six-month-old infant brain segmentation algorithms: the iSeg-2017 challenge. *IEEE Trans Med Imaging* 38(9):2219–2230
29. Abadi M et al (2016) *Tensorflow: Large-scale machine learning on heterogeneous distributed systems*
30. Harris CR et al (2020) Array programming with NumPy 585(7825):357–362
31. Ahmed S (1995) A pooling methodology for coefficient of variation. *Sankhyā: Indian J Stat Ser B*, 57–75
32. Mora M, Tauber C, Batatia H (2005) Robust level set for heart cavities detection in ultrasound images. In: *Computers in cardiology. IEEE*, pp 235–238
33. Schulze MA, Wu QX (1995) Nonlinear edge-preserving smoothing of synthetic aperture radar images. In: *Proceedings of the New Zealand image and vision computing'95 Workshop*, pp 28–29
34. Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M (2020) U2-Net: going deeper with nested U-structure for salient object detection. *Pattern Recogn* 106:107404
35. Ramachandran P, Zoph B., Le QV (2017) Searching for activation functions. *arXiv preprint [arXiv:1710.05941](https://arxiv.org/abs/1710.05941)*
36. Woo S, Park J, Lee J-Y, Kweon IS (2018) Cbam: convolutional block attention module. In: *Proceedings of the European conference on computer vision (ECCV)*, pp 3–19
37. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 7132–7141
38. Ghiasi G, Lin T-Y, Le QV (2018) Dropblock: a regularization method for convolutional networks. *arXiv preprint [arXiv:1810.12890](https://arxiv.org/abs/1810.12890)*
39. Taha AA, Hanbury A (2015) Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 15(1):1–28
40. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision*, p. 1026–1034
41. Dozat T (2016) Incorporating nesterov momentum into adam
42. Wang L et al (2018) Volume-based analysis of 6-month-old infant brain MRI for autism biomarker identification and early diagnosis. In: *International conference on medical image computing and computer-assisted intervention. Springer*, , pp 411–419
43. Li G et al (2014) Measuring the dynamic longitudinal cortex development in infants by reconstruction of temporally consistent cortical surfaces. *Neuroimage* 90:266–279

A Reinforcement Learning Based Adaptive Traffic Signal Control for Vehicular Networks



S. P. Krishnendhu, Mainampati Vigneshwari Reddy, Thulunga Basumatary, and Prabu Mohandas

Abstract The congestion of urban traffic is becoming one of the serious issues with the increase in vehicles and population in cities. The static time traffic controlling system fails to manage traffic and leads to heavy congestion and crashes on roads. To improve traffic safety and efficiency proactively, this study proposes a Adaptive Traffic Signal Control (ATSC) algorithm to optimize efficiency and safety simultaneously. The ATSC works by learning the Optimal Control policy via Double Dueling Deep Q Network (3DQN). The proposed algorithm was trained and evaluated on simulated isolated intersection Simulation of Urban Mobility (SUMO). The results showed that the algorithm improves both traffic efficiency and safety compared with static time traffic control technique by 42%.

Keywords Adaptive traffic signal control · Safety optimization · Reinforcement learning · Traffic simulation · Multi objective reinforcement learning · Single objective reinforcement learning

1 Introduction

The increase in population is leading to urban congestion on roads. Adaptive Traffic signal control that helps in reducing the waiting time of vehicles which leads to reducing fuel consumption of vehicles [1]. It is also more economical compared to the manual control by traffic police at intersections. The ATSC also maintains safety at junctions.

To accommodate the traffic in an efficient and safe way is the principal challenge to the traffic control system. Traffic control signals direct the vehicles to cross the junctions, but it fails to reduce the delay in the quick movement of traffic. When more vehicles approach from only one side, due to fixed green time for each lane, it also fails to reduce the queue length of vehicles.

S. P. Krishnendhu (✉) · M. Vigneshwari Reddy · T. Basumatary · P. Mohandas
Intelligent Computing Lab, Department of CSE, National Institute of Technology, Calicut, Kerala, India
e-mail: krishnendhusp@gmail.com

One of the most commonly suggested ways to solve this is by using an Reinforcement Learning (RL) based algorithm for ATSC. Few studies proposed some ATSC that are able to switch their objectives dynamically. But this type of ATSC is not suitable to optimize the safety and efficiency simultaneously. Because It can solve only one objective at a time. Another way to solve this is by using the Multi Objective Reinforcement Learning (MORL) based algorithm. In this kind, some of the work [2, 3] used a synthetic Q value, synthetic Q value is the Q value obtained from the Q values of the both objectives to account for Multi Objective simultaneously and some [4] used value function instead of synthetic reward where value function is the expected long-term reward.

RL is a goal based Machine Learning Algorithm. Similarly RL for ATSC has an Agent, that is traffic signal controller in our case, interacts with the traffic environment to fulfill the goal of reducing vehicle's waiting time and ensuring safety at the junctions. In this process Deep Q-Network (DQN) algorithm acts as a backend algorithm of the Adaptive Traffic Signal Controlling system. in which two Convolution Neural Networks (CNN) [5] act as probability finders to choose the correct signal phase according to the goal. Synthetic Q-value algorithm is used to normalize the highest probability of choosing a phase and lowest probability of occurring a crash and to find the average of both. Replay memory algorithm is used to store the executed input and the obtained rewards and calculated probabilities as tuples to train two CNNs.

2 Literature Survey

A single junction control can be of different types such as two-way junction, three-way junction and four-way junction or multiple way junction. Considering the two-way junction, there is not much traffic to control [6]. While on a three-way junction sometimes high traffic may arise. To solve this congestion, 'Traffic Circle' or automatic traffic control can help. On a Traffic Circle [7] in between the junction vehicles move round to the traffic circle to choose the right way for the destination. Vehicles turning on the traffic circle can be either left or right in different countries. This is because some countries drive on the left side of the road, others drive on the right. These traffic circles can optimize the congestion on roads but due to its huge area requirements it is not suitable to implement in a country like India.

A junction with multiple lanes i.e., four-way junction where vehicles are coming from all four directions. In this type of junction, a fixed timer controller leads to more waiting time for vehicles [6]. Some lanes may have more incoming vehicles in a pick hour. Which can lead to congestion and long queues of vehicles. To avoid the congestion and long waiting time for vehicles on junctions connected with multiple lanes, an automatic traffic signal control is required. An automatic traffic signal control guides the traffic to cross the junction by controlling the traffic signal and green time is also optimized [8] according to the traffic condition to increase the efficiency [9]. IR sensors with PIC microcontroller [10] can be used to achieve the

dynamic green signal time. In this IR sensors are placed on the road to detect the density of the vehicles and dynamic green signal time is calculated according to the traffic density on each lane. This work is tested only with one lane road.

The automatic traffic controller of fixed timers leads to more congestion on lanes. For example, peak hours where few lanes have high congestion and rest have low traffic, due to the fixed and long green signal time irrespective of congestion on the lanes leads to heavy congestion. When it comes to connected junctions, the waiting time of the vehicles doubles because vehicles should wait at each connected junction for the same fixed time [11, 14]. There is a need for Dynamic Priority based scheduling Adaptive Traffic signal Control which chooses Green time dynamically based on predicted traffic on connected junctions and also congestion on current junctions [8].

ATSC adapts to real time traffic scenarios and optimizes the traffic flow by dynamically changing the signal timings. ATSC reduces waiting time of vehicles, decreases the queue length, maximizes the vehicle throughput [12] Got positive results on solving the waiting time of the vehicles at a traffic junction. But optimizing traffic safety is not taken into consideration. Some of the work [2, 8] shows that the installation of ATSC reduces the number of crash occurrences. In contradiction to that a study on traffic conflicts [1] found that there is an increase of crashes after installation of ATSC. This raises the concern for ATSC's safety impact. To solve this issue this study supports and proposes a safety oriented ATSC.

The work [13] suggests to solve the problem with connecting the vehicles with the networks and sharing the traffic information with each other. This is a very good way to solve the traffic issue. But even after successful implementation of this research, a Junction traffic control system will be required to maintain the traffic because making every vehicle smart in our country like India will require long years.

Single Objective Reinforcement Learning (SORL) [1] is a goal-based Reinforcement Learning algorithm where only one objective is targeted to achieve as an end result. In SORL learning agents aim to learn an optimal policy via interactions with uncertain environments. Initially the agent doesn't know anything about the environment so the agent first explores the environment and learns a policy and then exploits. For example, the obtained normalized probability i.e. Q-value for the reducing waiting time objective is greater than that of the objective ensuring safety than the reducing waiting time objective is achieved. SORL is not suitable to achieve both safety and efficiency simultaneously.

Multi Objective Reinforcement Learning (MORL) [12] is a generalization of standard Reinforcement Learning algorithms where Multiple objectives are satisfied simultaneously [9]. To solve and optimize both safety and efficiency simultaneously we propose to MORL in this paper. In MORL Multi learning agents aim to learn an optimal policy via interactions with uncertain environments. To satisfy both objectives we first normalize the values and find the average of Q-value.

3 Problem Definition

The congestion of urban traffic is becoming one of the serious issues with the increase in the number of vehicles and population in cities. Let's take a real time scenario of a 4-way junction where green signal time is 5 min and let's assume that the time is 12AM and there is no traffic on the roads. A person waiting at the 4th lane should wait 15 min in-order to pass the junction. This in turn increases waiting time of the vehicles and may lead to pollution due to extra fuel consumption. There is no speed limit checker at junctions, there is no protected left turn to avoid crashes. Long intervals of green signal time leads to heavy congestion on roads and leads to a crash prone environment. Due to the current static time traffic control technique there are:

- increase in waiting time of vehicles at junctions.
- heavy congestion on roads.
- increase in crash occurrence on roads.

The proposed ATSC is a dynamic priority based algorithm unlike the current static timer controlling technique. The proposed ATSC using RL chooses the green signal time and appropriate signal phase dynamically leading to reducing waiting time of vehicles and ensuring safety.

4 Design

The environment of the traffic junction is implemented on a simulator SUMO based on a real-world four way intersection with free left condition. The junction is a typical four way intersection with three lanes each for incoming and outgoing vehicles on each of the ways.

4.1 Distance Level Assignment and Distance Estimation:

This distance level assignment is done to calculate the queue length of the traffic and to estimate the number of vehicles on each lane. We have shown the assumption of 3 distance levels i.e. 'A', 'B', 'C' on each edge of the road as shown in Fig. 1. Distance levels 'A', 'B', 'C' are of equal length. 'A' is the nearest to the junction.

Figure 2 shows the edge of a road with three lanes each for incoming and outgoing traffic. Every lane is of equal width. One lane each from incoming and outgoing are kept for free left for vehicles turning left. Other two lanes are used for going straight and turning right of vehicles. Each position 'A', 'B', 'C' can hold four bus size vehicles on it as shown in Fig. 3.

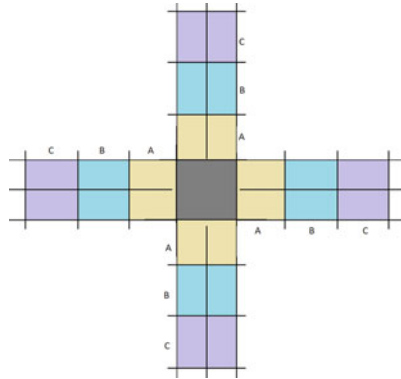


Fig. 1 Distance estimation

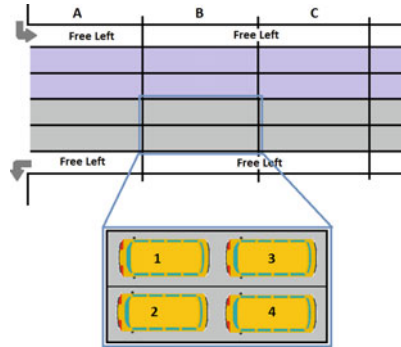


Fig. 2 Lane illustration

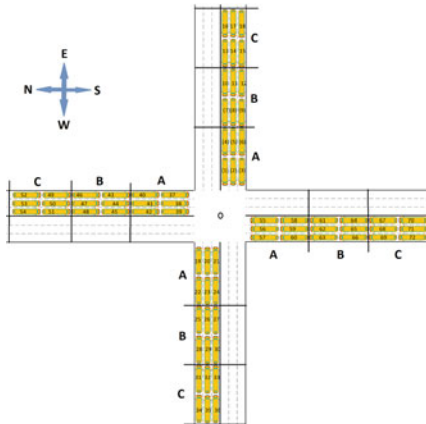


Fig. 3 Vehicle location illustration

4.2 Problem Statement with Notation and Design:

The waiting time of the vehicles and queue length of the vehicles are taken into consideration in this work as parameters to solve the problem. Because reducing the waiting time and reducing the queue length of vehicles at the junction are the key factors for increasing the efficiency of traffic control systems. Equations to find these parameters are given in Table 1.

- **Waiting time:** It is the total sum of waiting time of vehicles at the junction at that instance. Here waiting time is calculated to find reward, the main objective of the work is not to queue the vehicles at junctions, so in order to check it we calculate the waiting time as the total sum of vehicles at the junction.
- **Queue length:** It is the length of vehicles waiting on each lane to cross the junction at that instance.

Fix timer traffic signals control has a major issue with waiting time of vehicles. Let us assume a junction with many vehicles arriving at the junction. Let suppose 60 vehicles can cross the junction in 60 s of green time. If 121th vehicles arriving on east direction and wants to go to west direction then the waiting time of that vehicle is calculated in Table 2:

From the Table 2, it required 480 s to pass 120 vehicles from east incoming lanes. 121th vehicles will have to again wait for the next green signal. That would come after green time for the south, west and north direction. So again another waiting time for 121th vehicle is $60 + 60 + 60 = 180$ s. So the total waiting time for 121th vehicle is $480 + 180 = 660$ s.

Table 1 Parameters and equation

| Sl no | Parameter | Notation | Equation |
|-------|--------------|----------------|--|
| 1 | Waiting time | Wt | $Wt = \text{sum of waiting time of } (EA_v + EB_v + EC_v \dots + WA_v + WB_v + \dots + NA_v + NB_v + \dots + SA_v + SB_v + \dots)$ (Vehicle with position notation are shown at Fig. 4) |
| 2 | Queue length | Qe, Qw, Qn, Qs | $Qe = (EA_v + EB_v + EC_v + \dots)$ $Qw = (WA_v + WB_v + WC_v + \dots)$ $Qn = (NA_v + NB_v + NC_v + \dots)$ $Qs = (SA_v + SB_v + SC_v + \dots)$ |



Fig. 4 Symbol notation

Table 2 Waiting time for fix timer controller

| Green signal lane (60 s) | Count of vehicles crossing | | | |
|--|----------------------------|------------|------------|------------|
| | East | West | North | South |
| South | 0 | 0 | 0 | 60 |
| West | 0 | 60 | 0 | 0 |
| North | 0 | 0 | 60 | 0 |
| East | 60 | 0 | 0 | 0 |
| Total vehicles Crossed in 240 s | 60 | 60 | 60 | 60 |
| South | 0 | 0 | 0 | 60 |
| West | 0 | 60 | 0 | 0 |
| North | 0 | 0 | 60 | 0 |
| East | 60 | 0 | 0 | 0 |
| Total vehicles Crossed in 480 s | 120 | 120 | 120 | 120 |
| South | 0 | 0 | 0 | 60 |
| West | 0 | 60 | 0 | 0 |
| North | 0 | 0 | 60 | 0 |
| East | 60 | 0 | 0 | 0 |
| Total vehicles Crossed in 720 s | 180 | 180 | 180 | 180 |

4.3 Proposed Methodology

The proposed ATSC utilizes the DQN algorithm. DQN is a RL algorithm where CNN acts as functional approximators. The CNN agent interacts with the uncertain environment as shown in Fig. 5, to find an unknown underlying function that maps inputs to outputs. The proposed ATSC fulfills two objectives, reducing waiting time of vehicles and ensuring safety at junctions.

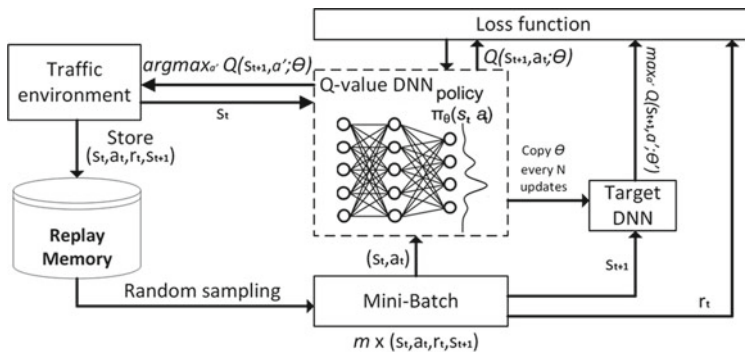


Fig. 5 Mechanism of DQN

The obtained Q values from the both DQN algorithms will be normalized and averaged using a weighted sum approach to satisfy both objectives simultaneously. ATSC using deep RL reduces waiting time of vehicles and ensures safety at junctions. The backend algorithm in ATSC Using RL is Deep Q-Network. In ATSC DQN is used to find an optimal function which maps inputs to outputs. The components of DQN are, Q network, Target Q network, Replay memory, Simulation environment, Mini batch, Loss function as shown in Fig. 5.

A Q network is used to find predicted Q values of the given state. and the target Q network is used to find predicted Q values of the next state. Loss function is used to find the error occurring between actual and predicted outputs. Mini batch and Replay memory is used to store the data to train the DQN networks.

The error obtained is the difference between the actual and predicted Q values.

The components of DQN are:

- Q network
- Target Q network
- Replay memory
- Simulation environment
- Mini Batch
- Loss function

All prior experience memory observations are saved as training data. Now we take a random batch of samples from training data and we store in a mini batch to avoid correlation we choose random samples of saved data. This Mini batch of training data is then sent to the Q-value DNN (Q network) and Target DNN (Target Q network). The Q network takes the current state and action from each data sample and predicts Q value for that particular action. The Target network takes the next state from each data sample and predicts the best Q value from that state. The predicted Q value, Target Q value, and the reward from the data sample is used to compute the loss to train Q network. Thus, the Q network is trained in continuous learning.

5 Junction Coordination

A method for getting dynamic traffic signal control to provide for the smooth movement of traffic with minimum waiting time of vehicles at the junction is proposed.

5.1 Setup

Environment is built on a simulator based on a real-world signalized intersection that we follow in India. The junction is a four way intersection with three lanes each for incoming and outgoing vehicles. Vehicles are permitted to move on green signal

time for that direction. For every edge one lane from incoming and outgoing lanes are kept for free left turn and vehicles are free left turn all the time. The green timing in all places in urban cities is chosen dynamically based on the density of vehicles on the lane. yellow timing is fixed in all places i.e. 5 s.

5.2 State, Reward and Action

The DQN algorithm (Agent) interacts with the Environment (Current Traffic scenario provided by SUMO). Based on the traffic scenario and the current traffic signal state agent takes an action (Providing different signal phases for different lanes). Agent (DQN algorithm) also receives rewards for taking actions, refer Fig. 6 for inputs and outputs. In other words, the reward value is positive when vehicles cross the junction more efficiently and safely.

- Reward for efficiency:

$$r = -(W_{t+1} + W_t) \tag{1}$$

- Reward for safety: $r = (R_t - R_{base})$. Where
 - R_t is the risk score obtained at time t and
 - R_{base} is the maximum risk score obtained during the pre-training process.

Risk score is the value determining the chance of occurrence of crash at junctions. Risk score is obtained from the history of accidents at junctions. The Risk score calculated on parameters like, green timing length, peak hours, restricted left turn etc. the parameters may vary from place to place, country to country. The calculation of Risk Score is important to maintain safety at junctions.

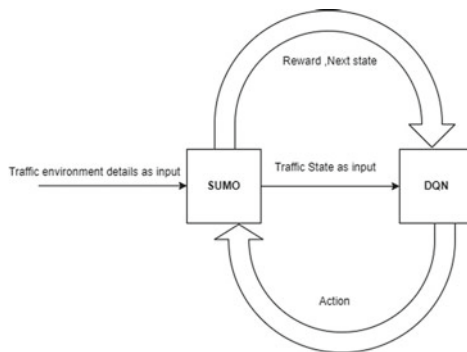


Fig. 6 Interaction of agent and environment

Risk score is calculated based on crash prone parameters like green interval length and traffic density etc.

Loss function:

$$L(\theta) = (((r + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta^{target})) - (Q(s, a; \theta^{pred}))))^2 \quad (2)$$

In the above equation $(r + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta^{target})) =$ Target Q value and $(Q(s, a; \theta^{pred}))$ is Predicted Q value.

State includes two components, that are traffic conditions at the junction and current activated signal phase. The action of the DQN algorithm (Agent) is to select the appropriate signal phase for the junction. Action of the agent follows some rule such as:

- Minimum green time.
- Maximum green time or maximum waiting time of vehicles.

These conditions are kept to ensure fair travel rights and to satisfy drivers expectations at the junction.

5.3 Density Calculation

Based on the incoming vehicles for every lane at the junction, density for that lane is calculated. Every assigned distance 'A', 'B', 'C' on each lane are of equal length and can accommodate 2 bus size vehicles. For every lane, the density of that lane is calculated according to the queue length of the vehicles on that lane. i.e., queue length of vehicles on one lane is up to 'B' position then the density of that lane is 'B' and if queue length of one lane is up to 'C' assigned position then its density is 'C'.

5.4 Pre Training

According to the design of the proposed algorithm there are two phases of pre-training. The objective of the first phase of pre-training is finding a baseline risk score. It is impossible to find the exact best risk score. Risk score is a measure of occurrence of crashes at junctions, which is obtained from existing records on crashes. Thus the hourly-average risk score of the benchmark scenario is utilized. This means the control agent is instructed to perform better than the benchmark signal controller. The second phase of pre-training is to find the range of Q values because we can't find an optimal function without knowing Q minimum and Q maximum values. During pre-training the agent observes the action taken by benchmark controller rather than taking actions based on it's own policy and updates the weights of Q networks by which we can find an optimal policy that works better than benchmark controller.

6 Work Plan

The backend algorithm of MORL is DQN. SUMO is used to generate traffic and state of the traffic will be sent to the DQN algorithm as you can see in Fig. 6. State of the traffic is a Matrix of 0's and 1's which is obtained by division of lanes into presence cells, if the vehicle is present in the cell then the value of the cell will be 1 or else 0. Action is a choice of traffic light phase from the possible phases at a junction. In DQN, Deep Neural Networks predicts the Q-values by Bellman's equation:

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a') \tag{3}$$

- The discount factor γ indicates the importance of future rewards.
- r means reward
- $Q(s, a)$ means the Q-value of state action pair s, a
- $\gamma \max_{a'} Q(s', a')$ is the highest Q-value of the next state action pair

Bellman's Equation tells that the expected Reward or probability of choosing an action of a particular state is the sum of the obtained reward r of the current (state, action) pair and the maximum expected reward of the next (state,action) pair u can refer Eq. 3. We need the maximum expected reward of the next state because choosing a state action pair will automatically take you to the possible states that should be chosen next, in those states we choose that state having a high expected reward refer Eq. 3. For example Let's assume there is a 4 way junction where we are in the current state lane 's1' from s1 we can go to s2 s3 and s4 states i.e. lanes. Bellman's equation says that, choose that state in s2,s3, s4 which yields maximum expected reward. In DQN we use rewards to train deep neural networks to find optimal policy. This is Multi Objective Reinforcement learning so there are two agents to satisfy two objectives so there are two types of rewards.

That means if the vehicles are in the queue, then the agent is penalized. For example, let's take a scenario of a 4-way junction. In the South Lane the waiting time of vehicles is 10 s and waiting time of vehicles in the east lane is 4 s and in the North Lane is 6 s and in the west lane is 2 s. Now 'Waiting time at t' i.e., cumulative waiting time of all vehicles at time t is $10 + 4 + 6 + 2 = 22$ s. Now the DNN agent comes and choses a lane which reduces waiting time of vehicles i.e., it chooses the lane south let's assume passes the green signal of 3 s time, the other lanes should wait for 3 s. Now go to the next step $t + 1$ calculate the cumulative waiting time of vehicles at time $t + 1$ i.e. $(4 + 3) + (6 + 3) + (2 + 3) = 21$. The reward for efficiency (refer Eq. 1)

$$\text{reward for efficiency} = - (\text{Waiting time of vehicles at time } t + 1) - (\text{Waiting time of vehicles at time } t).$$

$$r \Rightarrow (21 - 22) = 1$$

Therefore, DNN got positive reward that means the vehicles are not queued at lanes, if the agent gets negative reward DNN agent is penalized.

Reward calculation based on the safety:

$$r = (\text{risk score at } t - \text{risk score base}) \quad (4)$$

Risk score at t is the risk score at time t while base risk score is calculated during the per-training process refer Eq. 4 for risk score. If the safety performance is better than the baseline then the agent is rewarded. Two Q-values obtained for efficiency and safety are normalized to the same magnitude since the rewards and Q-values of the two objectives have a huge difference. By obtaining synthetic Q-value, both the objectives will be satisfied. MORL are computationally intractable for real time decision making models like Adaptive Traffic Signal control so we use weighted sum approach to obtain Synthetic Q-value function.

Replay memory is used to store the data to train Deep Neural Network. In DNN, Loss function(Eq. 2) is used to minimize the error by optimizing weights as u can refer to Eq. 2. Error is measured by the difference between the predicted result of Q-value and actual Q-value refer to Eq. 2. Theta is the weights of the policy pie of DNN, refer Fig. 5. We use two Deep Neural Networks, one to find target Q-value and one to predict Q-value and we use loss function to reduce error in prediction of Q-value which yields best action. The Q-Value DNN is used to match state action pairs i.e. to find the respective signal phase which reduces waiting time of vehicles and ensures safety at junctions by matching input state to the output action using function approximation.

7 Simulation Results

The simulation allowed us to dive into a large set of traffic management topics. It is carried out with the help of SUMO. Different traffic conditions are created for simulation where each vehicle has its own route, and moves individually through the network.

7.1 System Requirements

The minimum memory requirement for SUMO version 1.9.2 is 2 GB of RAM installed on a machine. Provided that machine has an Intel Dual-Core CPU at a minimum to run SUMO. One will need at least 1 GB of free disk space to install SUMO on a local machine. Above mentioned requirements match with small setup and configuration such as only one junction with limited number of vehicles simulation. Since SUMO can be used with an unlimited network size, an unlimited number

of simulated vehicles and an unlimited simulation time the system requirements can increase accordingly.

7.2 Evaluation Matrix

Evaluation matrix shows how the algorithm evaluates whether the objective is achieved or not. It's important to check the algorithm's correctness. We use the Deep Q-network algorithm to reduce waiting time of vehicles and to ensure safety at junctions. The W_{t+1} and W_t are the cumulative waiting time of current time $t + 1$ and previous time t of all vehicles, refer Eq. 1 for reward for efficiency. That means if the vehicles are in the queue then the agent is penalized. For example, let's take a scenario of a 4-way junction. In the South Lane the waiting time of vehicles is 10 s and waiting time of vehicles in the east lane is 4 s and in the North lane is 6 s and in the west lane is 2 s. Now 'Wt' i.e. cumulative waiting time of all vehicles is $10 + 4 + 6 + 2 = 22$ s according to Eq. 1.

Now the DNN agent comes and chooses a lane which reduces waiting time of vehicles i.e. it chooses the lane south let's assume passes the green signal of 3 s time, the other lanes should wait for 3 s. Now go to the next step $t + 1$ calculate the cumulative waiting time i.e. $(4 + 3) + (6 + 3) + (2 + 3) = 21$. $r = > -(21-22) = 1$ according to Eq. 1.

7.3 Static Traffic Density

Static traffic density is the calculation of the number of vehicles on the lane based on a static timer traffic controller controlling the traffic at junctions. As mentioned in the problem statement, the fixed and long green signal timing leads to heavy congestion on lanes i.e. there will be high traffic density on lanes due to static timer traffic controllers.

7.4 Dynamic Traffic Density

Dynamic traffic density is the number of vehicles on the lane based on the dynamic priority scheduling algorithm controlling the traffic at junctions. The density of lanes in Dynamic scheduling algorithm controlled junctions is lesser than that of static traffic density. Let's calculate the waiting time of vehicles to compare the static and dynamic densities. If the waiting time of vehicles at a given time is lesser, that means low density at the junction. Waiting time of vehicles using DQN of a 4-way junction: Let's take a scenario of a 4-way junction. In the South Lane the waiting time of vehicles is 10 secs and waiting time of vehicles in the east lane is 4 secs and in the

North lane is 6 s and in the west lane is 2 s. Now 'Wt' i.e. cumulative waiting time of all vehicles is $10 + 4 + 6 + 2 = 22$ s.

Let's assume the south lane is chosen to free traffic. The green signal is 3 s time, the other lanes should wait for 3 s. Now go to the next step $t + 1$ calculate the cumulative waiting time i.e. $(4 + 3) + (6 + 3) + (2 + 3) = 21$.

Now let's take the same scenario to calculate static traffic waiting time. Now let's assume the green signal time is 20 s, the other lanes should wait for 20 s. Therefore the waiting of vehicles is $(4 + 20) + (6 + 20) + (2 + 20) = 72$.

Therefore, the waiting time using dynamic scheduling algorithm is 21 and using static algorithm is 72. which indicates due to high waiting time in the static scheduling algorithm the vehicles will be queued at junctions increasing the density of vehicles at junctions, refer Eq. 1 for reward.

8 Conclusion

The static time traffic controlling system fails to reduce congestion and improve safety at junctions. To improve traffic efficiency and safety simultaneously an ATSC based on Deep reinforcement learning algorithm is proposed. The control agent takes the Real time traffic data as input (Real time traffic data is created from an open source simulator Simulation of Urban Mobility to test) and selects appropriate signal phases for every state to reduce waiting time of vehicles and crash occurrence at junctions. The MORL framework with DQN as backend algorithm is used to find optimal functions to map inputs to outputs. The weighted sum approach is used to work simultaneously on traffic efficiency and safety. The results showed that the algorithm improves both traffic efficiency and safety compared with static time traffic control technique by 42%.

This work still has space for improvement and can be extended by adding the ability to permit emergency vehicles such as ambulances to be given the highest priority. In addition the system can be made more efficient by connecting nearby junctions to share traffic data to each other to make better decisions.

References

1. Rasheed F, Alvin Yau K-L, Noor RMd (2020) Deep Reinforcement learning for traffic signal control: a review 8
2. Khamis MA, Gomaa W (2014) Engineering applications of artificial intelligence 29
3. van der Pol E, Oliehoek FA (2016) Coordinated deep reinforcement learners for traffic light control
4. Jin J, Ma X (2015) Transportation research procedia 10
5. Mohandas P, Anni JS, Thanasekaran R, Hasikin K, Azizan MM (2021) Object detection and movement tracking using tubelets and faster RCNN algorithm with anchor generation. Wirel Commun Mobile Comput

6. Essa M, Sayed T (2020) Self-learning adaptive traffic signal control for real-time safety optimization 146
7. Singh L, Tripathi S, Arora H (2009) Time optimization for traffic signal control using genetic algorithm
8. Jin J, Ma X (2015) Adaptive group-based signal control by reinforcement learning 10
9. Houli D, Zhiheng L, Yi Z (2010) Multiobjective reinforcement learning for traffic signal control using vehicular Ad Hoc network. Article number: 724035
10. Ghazal B, Chahine KKK, Smart traffic light control system <https://doi.org/10.1109/EECEA.2016.7470780>
11. Jin Z, Zhang Y (2012) Traffic circle administration based on circuit principles and marginal benefit theory 17(Part A)
12. Khattak ZH, Mark JM, Fontaine MD (2018) Estimating safety effects of adaptive signal control technology 64
13. Chen, Zong JI, Smys S (2020) Optimized dynamic routing in multimedia vehicular networks. *J Inf Technol* 2(3):174–182
14. Manley Ed (2015) Estimating urban traffic patterns through probabilistic interconnectivity of road network junctions

The Structure of the Vulnerability Detection System on Web Servers



Aziz Ibrokhimov

Abstract In the article, vulnerabilities in web servers, the structure of the vulnerability detection system available on web servers, the types of attacks that can be carried out using vulnerabilities, OWASP top 10 has been studied and researched. Apart from these, vulnerabilities in web servers are listed in the dynamic view of the register.

Keywords Vulnerability · Web-server · Malware · Threat · Weakness · OWASP

1 Introduction

The existence of vulnerabilities in the information resources of information systems, along with the possibility of computer attacks and the spread of viruses, gives rise to various random errors and possible causes of loss of resources. Given the dynamics and complexity of the programs, the development of information technology, it is necessary to constantly improve the methods of solving this problem and dealing with new types of threats.

Weaknesses in web servers can never be completely eliminated. Because of the dynamic ports and users of data exchange on web servers, there are constant errors in information security. Therefore, it is recommended to use real-time security monitoring systems in all information systems.

2 Main Part

Web server vulnerabilities, on the other hand, are referred to as security vulnerabilities. A bug is a weakness in software development that has the potential to compromise information security. In this scenario, the security vulnerability employed is

A. Ibrokhimov (✉)
Amir Temur, Tashkent 108, Uzbekistan
e-mail: a.ibrokhimov@tace.uz

a vulnerability, resulting in information system or web server interruptions, threats, and attacks. That is why web server vulnerability detection technologies have always been in high demand. There are a number of security flaws in web servers all over the world.

In the realm of software security today, there are a variety of simple and complicated hierarchical groups. These are:

- threat and attack classification;
- malicious program classification;
- vulnerability classification and registers.

The necessity to periodically provide concise information about vulnerabilities for each kind and version of software products and environments is related to vulnerability classifications and registers on web servers. Major software developers (e.g., Adobe, Microsoft, RedHat) and various organizations (e.g., US-CERT, Open Information Security Foundation) create such registers, which compile statistics from all around the world on a regular basis. On the other hand, a number of registries have been established to collect vulnerabilities in software developed by multiple software providers into a common identification mechanism (for example, CVE-ID) [1, 2] (Table 1).

Weaknesses in web servers can be detected during or after detection of vulnerabilities [3]. Examples of such signs are the following signs.

- recurrence of certain events;
- incorrect commands at the current time;
- use of other vulnerabilities (existing vulnerabilities in the web server directly lead to other vulnerabilities);
- incorrect parameters of network traffic;
- unexpected attributes;

The structure of the vulnerability detection system on web servers is in the form of a three-tiered hierarchy [4] (Figs. 1 and 2).

The modern vulnerability identification framework for web servers is built by connecting vulnerabilities in a distributed manner to each other and to a central management server. By centralizing the vulnerability database, vulnerability detection systems with this topology can improve system speed. Weakness detection methods necessitate the simultaneous connection of three servers. These servers are:

- central analysis server;
- agent server;
- vulnerabilities database collection, processing, and management server;
- vulnerabilities database collection, processing, and management server;
- vulnerabilities database collection, processing, and management server;
- vulnerabilities database collection, processing.

Table 1 Register vulnerabilities on web servers

| No | Appearance | Examples | Features |
|----|--|--|--|
| 1. | Classification of malware | Miter MAEC (Enumeration and grouping of malicious software attributes)—implementation based on the list and characteristics of malware | A programming language that is common to group malicious programs, taking into account behavioral characteristics, type of attack, and so on |
| | | Kaspersky Classification is a classification of Kaspersky Lab | Malware grouping by methods of exposure |
| 2. | Register of vulnerabilities and groups of the software supply system | MITER CVE (Common Vulnerabilities and Exposures)—common vulnerabilities and vulnerabilities | Database of known vulnerabilities |
| | | NVD (National Vulnerability Database)—US National Vulnerability Database | Weaknesses database when using CVE identifiers |
| | | CERT Vulnerability Notes Database | Identified vulnerabilities and how to identify them |
| | | Guide bulletins: – Microsoft Newsletter IDs – Exploit-DB | Conclusions on identified vulnerabilities |
| 3. | Classification of security threats | OWASP Top Tente is one of the 10 most common web add-on threats | Threat classes associated with vulnerabilities in web applications |
| | | MITER CAPEC (Common Attack Pattern Enumeration and Classification) - list of common attack types | Extensive classification of types of attacks |
| | | WASC Threat Classification 2.0 is a classification of web application security consortium threats | Web Security Threats Focused on Applications |

As the number of vulnerabilities caused by web server vulnerabilities has grown year after year, a number of multinational efforts have been launched around the world to construct and maintain a database of web server and application vulnerabilities. Any web server and application should have vulnerability detection mechanisms. OWASP (Open Web Application Security Project) is a community that currently develops add-on vulnerabilities, add-ons and technologies to improve security systems. This organization is an ineffective community that implements the security of open web applications. According to the OWASP methodology, the vulnerabilities in web servers and applications are divided into 3 parts, through which it is

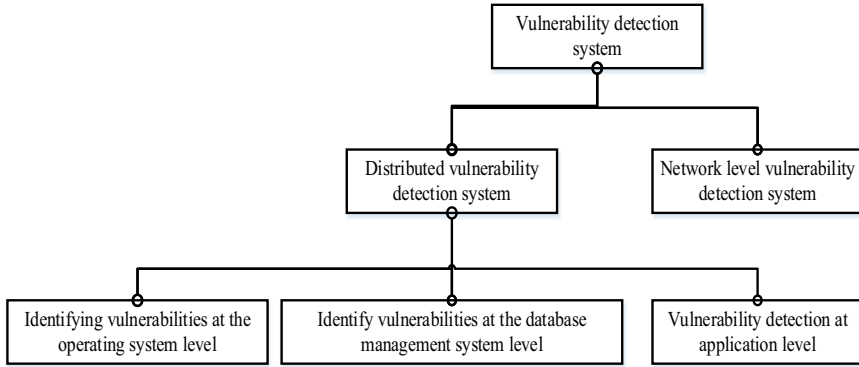


Fig.1 The structure of the system for identifying vulnerabilities in web servers

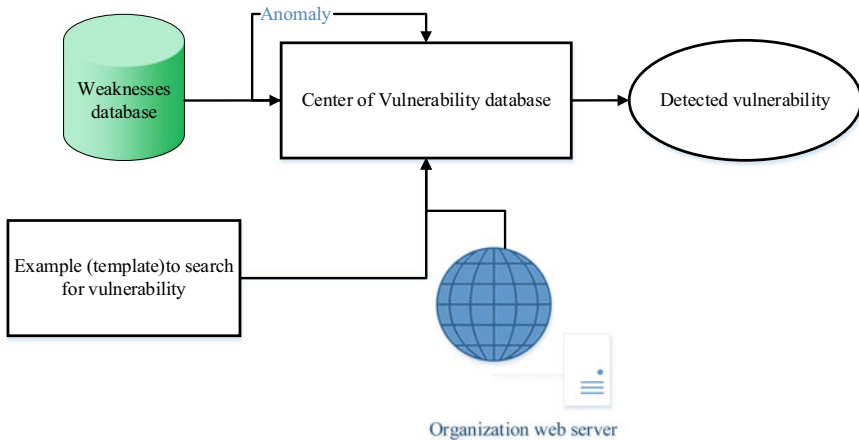


Fig. 2 The modern structure of the system for identifying existing vulnerabilities in web servers

possible to carry out a number of attacks. These types of vulnerabilities include [5, 6]:

- injection weaknesses;
- business logic vulnerabilities;
- weaknesses in session management;

The attacks, which are on the OWASP’s 2021 list of top 10, were as follows [6] (Figs. 3 and 4).

A01: 2021-Access control violations were fifth in 2017 and first in 2021. In other words, taking advantage of this vulnerability, the loss of information on web servers and the disruption of the information system are severely damaged. Identified vulnerabilities 3.81% One or more Common Weakness Enumerations (CWE) and

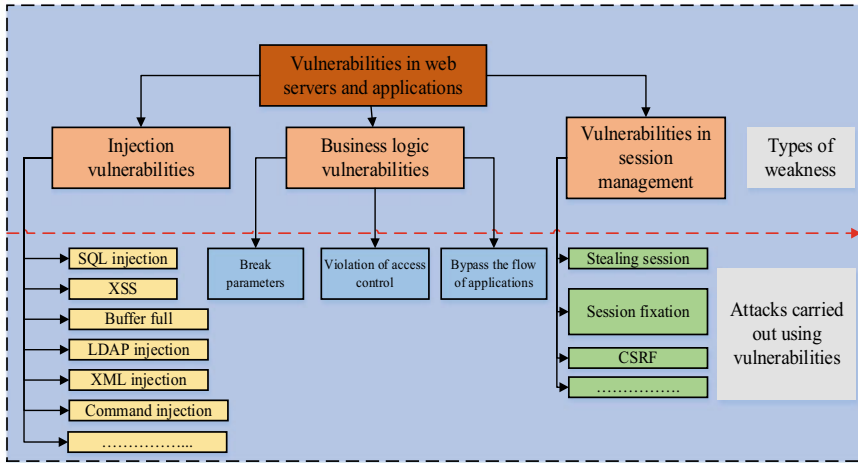


Fig. 3 Weaknesses in web servers and the types of attacks that can occur as a result

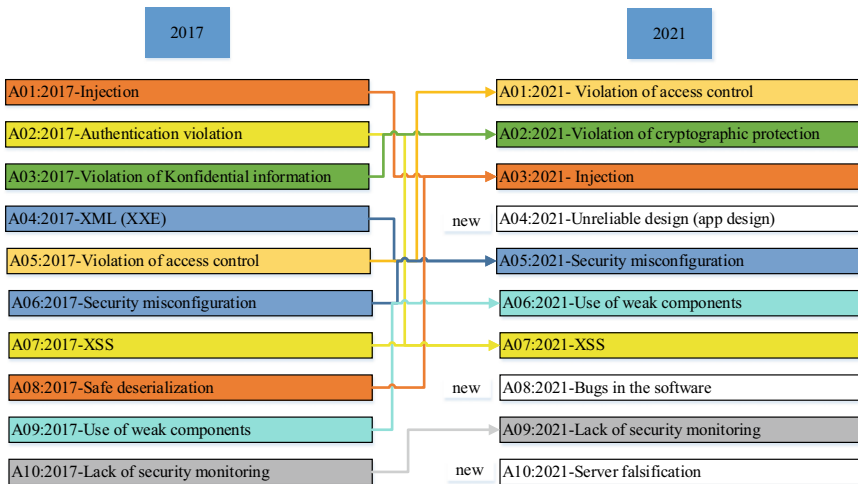


Fig. 4 OWASP’s top 10 list

more than 318 thousand in this risk category CWE cases have been identified. Access control violations are included in 34 web applications CWE is more common on web servers than other categories.

A02: 2021-Violation of cryptographic protection than in 2017.

Twice as many. This type of vulnerability has compromised the integrity and confidentiality of information on web servers.

A03: 2021-As a result of injection vulnerabilities, 33 CWEs included in web applications received 274,000 recurrences.

A04: 2021 and A05: 2021 - 33 CWE 208 thousand recurrences in web additions. There is no way to completely eliminate unreliable design weaknesses. This is because users prefer websites with a dynamic interface. Dynamic web pages cause unreliable design vulnerabilities.

A06: 2021-The use of vulnerable components was originally called the use of components with known vulnerabilities. Over the past two years, the number of professionals using such components has increased dramatically. This is to identify vulnerable components of the system without compromising the security mechanism and to allow unauthorized access to data on web servers.

A07: 2021-Multiple CWEs can be sent to a web server due to an authentication violation or an identification and authentication error. When using such methods, the attackers will be able to bypass both classical and modern methods of defense.

A08: 2021-Errors in the software. This is a new version of A07: 2021, and includes 10 new categories of CWEs that are not part of A07: 2021.

A09: 2021-Due to the inability to monitor the system in real time due to the lack of security monitoring, the information on the web servers is not able to record attacks, or if there is a monitoring system, it is possible to receive accurate information about the system due to system failures.

A10: 2021-Server forgery. Because it is a classic method, it is rarely used today. This can be done on web servers that are protected by classical methods. It is possible to identify servers using modern security methods. This is because two-way authentication allows this to be determined.

3 Conclusion

In conclusion, based on the results of comparative study, the black box method is currently the most effective method of discovering vulnerabilities in web servers. The vulnerability detection system for web servers and applications must be improved, and a black box approach based on the structure of a modern vulnerability detection system must be developed. This is because, according to OWASP statistics from the last ten years, the number of web vulnerabilities is increasing as hundreds of web apps are produced. Each new opportunity reduces the protection mechanism's effectiveness. If you want to add a new feature to a web application, you must first figure out what vulnerabilities the system has before you can implement it. Otherwise, the web server's entire operation will be jeopardized. That is, by adding a feature to a web application, you can disable existing features' security. As a result, the work must be organized in conjunction with the security officer.

References

1. Alonso G, Casati F, Kuno H, Machiraju V (2010) Web services: concepts, architectures and applications. first ed., Springer
2. Christey S, Martin RA (2007) Vulnerability type distributions in CVE. The MITRE Corp 1:1
3. Gray J (1993) The benchmark handbook: for database and transaction processing systems. Morgan Kaufmann Publishers Inc.
4. <https://www.checkmarx.com/2017/12/03/closer-look-owasp-top-10-application-security-risks/>
5. Bekmuratov TF, Botirov FB, Haydarov ED, Electronic spam filtering based on neural networks, chemical technology. Control Manag 3(93):59–65
6. Bekmuratov TF, Botirov FB, Development of structures of intellectual information protection system, chemical technology. Control Manag 6(90):63–71

Comparison of Image Blending Using Cycle GAN and Traditional Approach



Medha Wyawahare, Ninad Ekbote, Sameer Pimperkhede,
Atharva Deshpande, Pranav Bapat, and Ishan Aphale

Abstract GANs are a very interesting emerging domain. Implementation of the Neural style transfer technique can be done using particular types of these GANs. In this technique, a particular art style is superimposed over a certain image, resulting in the input image looking like a painting in the same style. A similar result can be achieved through manually mixing the input image and a painting using rudimentary image manipulation techniques. The goal of this paper is to analyse the difference between these outputs using a variety of techniques such as Discrete Cosine Transform and Fourier Transform along with others, to observe the amount of difference in the techniques. We conclude with a discussion on how the manually mixed image can be tweaked to most closely resemble the GAN output.

Keywords GAN · Style transfer · DCT · Similarity

1 Introduction

Image-to-Image translation has many applications, from the fields of photography to image synthesis the research of Image-to-image translation has been a vital part in the Machine Learning community. Image to image translation plays a major in image reconstruction, style transfer, Image segmentation etc.

Translating an image to another image is primarily done using Neural Style transfer [1]. It makes use of an input image along with a reference art image and outputs an image that makes it seem as though the input has been painted in the style of the art style photo. This is framed as a supervised problem.

The introduction of GANs [2] added another interesting alternative to this. CycleGAN [3] is one such popular approach to unpaired image to image translation. This architecture uses a discriminator and a generator for image synthesis. Both update themselves as the generator tries to trick the discriminator and the discriminator tries to better detect the generated images.

M. Wyawahare · N. Ekbote (✉) · S. Pimperkhede · A. Deshpande · P. Bapat · I. Aphale
Department of Electronics and Telecommunication, Vishwakarma Institute of Technology, Pune,
Maharashtra, India
e-mail: ninad.ekbote18@vit.edu

2 Background

If we see two different images, we sometimes naturally imagine a sort of mixture or a blended version of them. Image mixing is an important topic in the field of image processing and computer vision [4, 5]. Simple image mixing using the addition of images by multiplying with required weights may give gibberish results. A better approach of image mixing or essentially the style transfer is using the neural networks and generative adversarial networks (GANs) which perform intelligent mixing. This is a research topic under image-to-image translation and relatively new in field of computer vision. The need for comparing the manual and intelligent image mixing for particular application will be very useful as we can further tune the GAN to give better image mixing results for that specific application. The addition of two images in our Traditional Approach is based on the famous ‘Prisma’ Traditional Approach which uses simple image blending instead of neural style transfer. However, this is traditional image processing. Hence it is better to analyse the outputs of traditional image processing and intelligent mixing done by GAN to understand the specialities that the GAN can explore to mix images then traditional image processing.

3 Literature Review

Unpaired image to image translation [3] is an important topic under research in the field of computer vision and generative adversarial networks (GAN). Lots of research work has been done in the technique of data augmentation using DCGAN [6], whereas other different types of GANs like StyleGAN [7, 8], CycleGAN are under research yet have not been explored fully.

CycleGAN is a method for training unsupervised image interpretation models utilizing the GAN design with unpaired picture dataset from two unique areas. The two different domains may not be related to each other at all. CycleGAN has been carried out for scope of uses including season interpretation, object change, style move, and creating photographs from compositions. Some common implementations of CycleGAN which has been done are horse to zebra image translation, summer to winter season translation, style transfer of images and translating paintings to photographs.

Some of the traditional methods of image blending researched previously are Laplacian pyramid blending, Poisson’s blending and simple image addition. Some of the research is also done on modifying Poisson’s blending and smoothening the image boundaries [9]. Research has been done on designing Intelligent GAN network for image blending. The GP-GAN has been proposed for image blending methods and it leads to fine blending of high-resolution images with sharp boundaries [10] but exact analysis where the GAN makes image blending special remains unexplored.

This work focuses on use of CycleGAN for style transfer [11]. This is also achievable by an alternative image processing-based approach which includes mixing of two

images manually according to user defined ratio. Our research aims at comparison of outputs of both the approaches mentioned above using statistical analysis.

4 Problem Statement

Analysing and comparing image blending techniques of a GAN network and a Traditional Approach to find mathematical insights for further studies in field of Generative adversarial Networks. This research aims to find the similarity between the outputs of CycleGAN and the image blending Traditional Approach. Various Similarity indices are used for same purposes which are discussed further in this write-up.

5 Dataset

For comparison purposes we used an openly available dataset known as “I’m Something of a Painter Myself” [12]. The dataset contains two sets of data which are Landscape photos and Monet style photos. The Monet set contains 300 Monet images, their dimensions are $256 \times 256 \times 3$. The Landscape set contains 7028 images their dimensions are $256 \times 256 \times 3$. The images are unpaired. CycleGAN is used for both paired and unpaired Datasets.



Fig. 1 Shows the sample images taken from Landscape set and Monet set respectively. Sample from Landscape(left) image and Monet image(right)

6 Methodology

6.1 Generative Adversarial Network (GANS)

Generative Adversarial Network (GANS), are a way to deal with generative modelling utilizing deep learning techniques, for example, convolutional neural networks.

Generative modelling is an unsupervised machine learning technique. This method includes automatically discovering and learning the patterns and regularities in an input data. A GAN model consists of Two individual neural network, these two neural networks are called Generator and Discriminator. There are many types of GANS for the purpose of this research we will be focusing on Cycle GAN.

CycleGAN

The CycleGAN network is specially designed over an image-to-image translation of unpaired data. So, the training of image-to-image translation model is done automatically in CycleGAN training. The models are trained in an unsupervised way using an assortment of images from the source and target domain that are not connected in any way. For training, CycleGAN uses Cycle consistency loss for unpaired images [3].

A U-net based generator is used. The generator first down samples the input image and then up samples while establishing long skip connections. Skip connections are a way to help bypass the vanishing gradient problem by concatenating the output of a layer to multiple layers instead of only one. Instead of giving a single node, the discriminator gives a smaller 2D image with higher pixel values indicating a real classification and lower values indicating a fake classification. CycleGAN has four network models two Generator model and two discriminator model [3]. Following the function of Generators used in CycleGAN.

Generator 1: Transforms landscape images to Monet images. (collection1 to collection2).

Generator 2: Transforms Monet images to landscape images. (collection2 to collection1).

The discriminator and generator models for a CycleGAN are trained under typical adversarial loss like a standard CycleGAN model.

CycleGANs are additionally refreshed utilizing cycle consistency loss. The selection of Hyperparameters was taken from [3] and no changes were made during implementation.

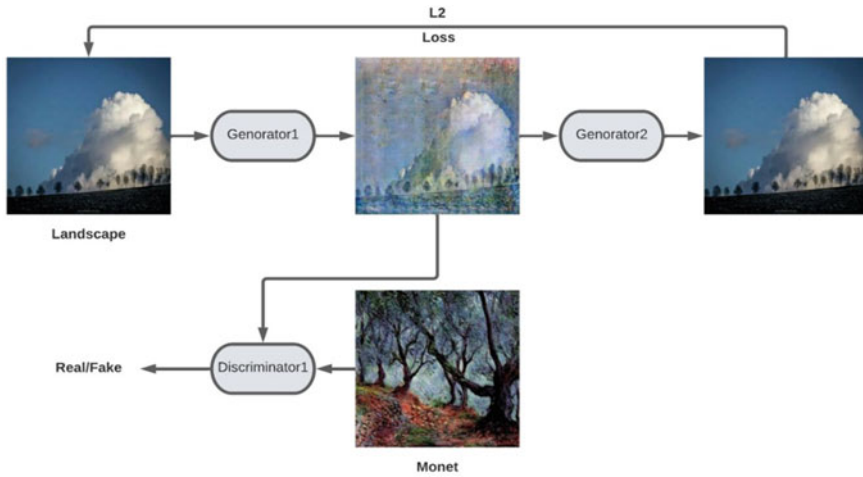


Fig. 2 Demonstrates the working of CycleGAN on actual dataset. Two Generator and two Discriminators are used for translation from Landscape to Monet and Monet to Landscape

6.2 Traditional Image Approach

We have developed an android application to achieve a traditional image generation approach. Our simple application designed for image mixing selects two photos from gallery, one as a foreground image and one as a background image. We adjust the weights and saturation of image by sliding the sliders. According to value of sliders images mix with each other. For mixing the image we just add the image pixel by pixel according to the weights given to them by slider. The Traditional Approach also has the facility to change the brightness or contrast of the image for better artistic looks and features. The slider is added to change the brightness value to each pixel of the image. As the user is allowed to adjust the values of sliders, this method requires the images to be of same size. Hence Traditional Approach has feature to resize the images. This is a method that we designed to grant people the maximum level of freedom in combining or blending through the traditional means. The output from Traditional approach will be referred as APP output. As an APP was used to implement traditional approach.

7 Result

We have developed an android application using specifications mentioned earlier and also trained a CycleGAN model for the same. In an application user can adjust the blending ratio of this image. CycleGAN model produces a blended image on its own.



Fig. 3 Demonstrates CycleGAN(left) output and APP out(Right)

We get the generated output from Traditional Approach and CycleGAN both and analyse them. We used FFT(DFT) and DCT as comparison techniques and structural similarity techniques for measurement of similarities between two images. According to outcome of analysis and comparison CycleGAN's output were better than APP's output.

8 Analysis

The images generated by the image blender Traditional Approach as well as the images generated with the help of CycleGAN need to be analysed and interpreted. This analysis will help us to determine the performance of both the systems and compare their outputs on a direct basis. We have used the following methods in order to investigate the similarity between the generated images.

Investigation of similarity between two images/signals is required in many application areas [13–18] and this can be done on the basis of following metrics.

8.1 Structural Similarity (SSIM)

The Structural Similarity (SSIM) index is a method for measuring perceptual similarity between two images. The SSIM index can be viewed as a quality measure of one of the images being compared, provided the other image is regarded as of perfect quality [19]. In the actual algorithm, SSIM is calculated on various windows of the images. Mathematically, the measure between two windows x and y of same size ($N \times N$) is given as:

$$\text{SSIM} = \frac{(2u_x u_y + c_1)(2\sigma_x \sigma_y + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (1)$$

In Eq. 2, x and y are the two images being compared and μ, σ are their mean and standard deviation respectively. $c1$ and $c2$ are constants. SSIM looks for similarities within pixels [19]; i.e., if the pixels in the two images line up and or have similar pixel density values. SSIM puts everything on the scale of -1 to 1 . A score of 1 means the images are very similar whereas -1 means they are very different [19]. The Structural Similarity Index metric extracts 3 features from an image: luminance, contrast and structure [20]. The comparison between the images is performed on the basis of these features.

In our case, let's consider the images shown in results section where the right image is Traditional Approach output and left image is CycleGAN blended image. SSIM index for these images is found out to be 0.01183 (rounded). Thus, we can see that the metric is quite close to 0 , indicating there is a lot of difference between the two images, of which image generated by CycleGAN is much better.

8.2 PSNR

PSNR shows a ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation [21]. PSNR performs direct pixel to pixel comparison and hence will give result as dissimilar images even if there is small change in contrast, luminance etc. unlike SSIM.

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \tag{2}$$

$$MSE = \frac{\sum M, N [I_{1(m,n)} - I_{2(m,n)}]^2}{M * N} \tag{3}$$

In Eq. 3, MSE means average of square of errors and R is the maximum fluctuation in image [21].

PSNR Analysis method: - Using traditional image mixing one foreground image is mixed with 10 background or Monet images. To carry out PSNR analysis the first 10 background images are mixed by directly adding them pixel by pixel with 1:1 ratio. Then this image mixture is mixed with one foreground image in 4 different ratios. The ratio background images are mixed with weights 50, 60, 70, 80, 90 etc. the PSNR value for each mixing ratio is observed and plotted.

PSNR Analysis Observation

Analysis Method 1

As the weight of foreground image increased, the PSNR value of CycleGAN to APP image also increased. This is due to the fact that 10 background images that are mixed together act as noise for the foreground image PSNR is metric of image similarity hence it shows the fact that, in CycleGAN image mixing the foreground images are

Table 1 PSNR values of APP output with respect to CycleGAN output

| | 50 | 60 | 70 | 80 | 90 |
|-----|-------|-------|-------|-------|-------|
| [1] | 40.79 | 41.44 | 41.72 | 41.51 | 41.33 |
| [2] | 41.05 | 39.56 | 37.85 | 37.08 | 35.98 |
| [3] | 41.83 | 41.52 | 40.64 | 40.03 | 38.99 |
| [4] | 41.87 | 39.99 | 38.24 | 37.27 | 35.93 |
| [5] | 38.92 | 38.58 | 37.97 | 37.76 | 37.11 |

more prominent and originality of image is retained. Table 1 shows the PSNR values of corresponding CycleGAN output and APP output of first 5 images 50, 60, 70, 80, 90 represents the proportion with which we created the APP output This showcases.

Analysis Method 2

The method 1 proves that increasing background image weight reduces similarity with CycleGAN but it doesn't necessarily prove CycleGAN image mixture is better. Hence 2nd Analysis method aims to prove that CycleGAN image mixture is better than APP image mixture. In this method PSNR score of CycleGAN mixture is calculated with original foreground image and same is done with APP image mixture with 80% foreground image weight and original image. These scores are compared with each other and plotted. The PSNR score of CycleGAN -Original is higher than APP-Original image in 18/20 images. In 4 images the scores are very close. This proves

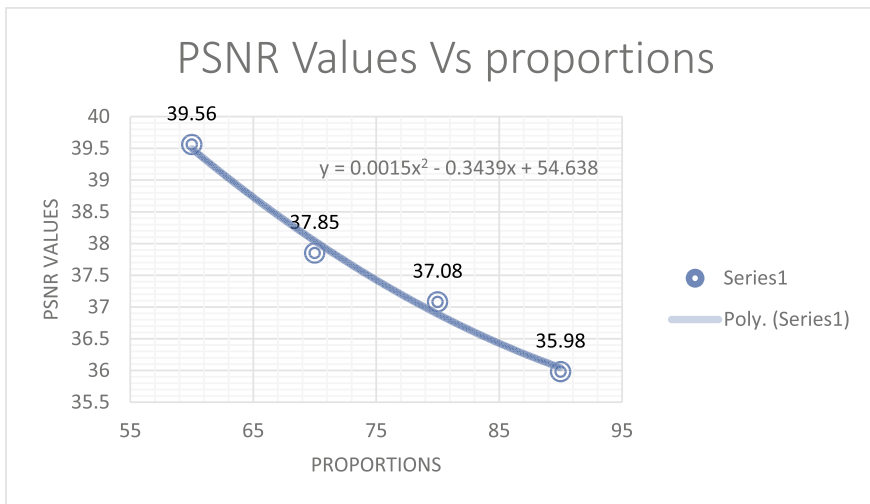


Fig. 4 Depicts variation of PSNR values with mixing proportions Graphs. X-axis in the graph shows mixing weight(proportions) of background images and Y-axis in the graph shows PSNR values of APP with respect to CycleGAN images

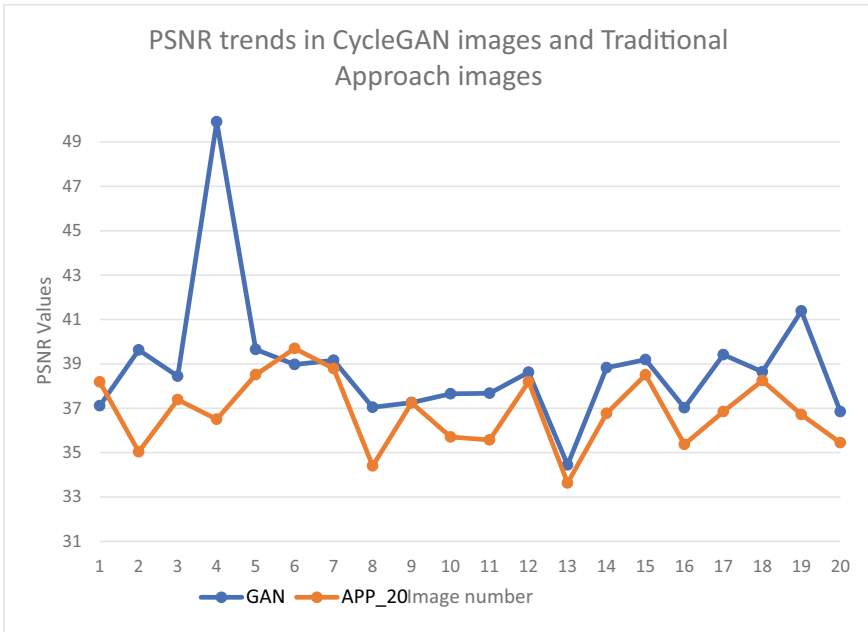


Fig. 5 X-axis proportions, Y-axis PSNR values)In the graph, x-axis shows the original image number and Y-axis shows the PSNR value with corresponding image number. The Red line shows the APP-original PSNR and blue shows the CycleGAN -original PSNR. The graph clearly shows CycleGAN -original graph lies mostly above to APP-original

that CycleGAN mixture is very similar to original image and hence proves to retain the property of original image unlike Traditional Approach image.

Results: - The SSIM score for all the images is close to 1 and lies in the range 0.8 to 1 where 0.98 is the maximum. Hence this method proves the images produced by CycleGAN and traditional image processing are similar perceptually and in terms of contrast and luminance. The Same analysis method 2 is also done for SSIM as of PSNR which gives similar results as that of PSNR.

However, the PSNR ratio is random and mostly indicate towards dissimilarity of images.

8.3 Fourier Transform

The Fourier transform is used in order to view the frequency components of the images as frequency domain view of images can give us more knowledge like the central frequency, magnitude, and phase plots [22]. Fourier analysis helps to decompose an image into its sine and cosine components which tells the highest frequency present and the average brightness (dc component) of the images. For the previous

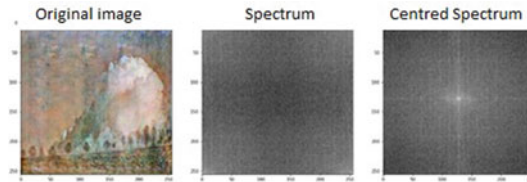


Fig. 6 (DFT For CycleGAN output image: - From left to right CycleGAN output, Spectrum, Central spectrum)

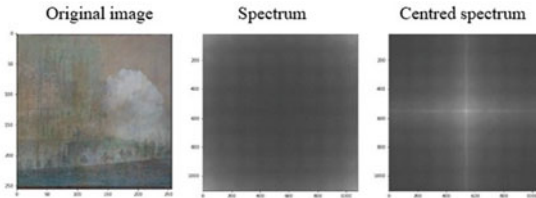


Fig. 7 (DFT For APP output image: - From left to right APP output, Spectrum, Central spectrum)

mentioned images, their grey-scale image, frequency spectrum and centred spectrum have been plotted.

Figures 6 and 7 shows the output from DFT. As we can see from these three images, the CycleGAN image spectrum consists of single centred frequency, whereas the APP output one has quite a few high frequencies in between [23]. These high frequencies can lead to poor image quality and thus degradation of image. Thus, the CycleGAN network also takes into account this fact of frequencies present in the image, while in the APP system we have no way to restrict this. This is a distinct advantage of the network as it learns through numerous input data images.

8.4 Discrete Cosine Transform

The discrete cosine transform (DCT) basically represents a finite length of data points in terms of cosine functions having various frequencies. In case of digital images, the high frequency components can be discarded and then image can be analysed [24]. It uses only cosine components for estimation and hence it is better for computational purposes. It is used in order to calculate values known as DCT coefficients which contain direct relation with the pixel values. The DCT performed on a $(N \times N)$ matrix, generally (8×8) matrix, of pixels results into a $(N \times N)$ square matrix of recurrence coefficients. Every component of the 8 pixel-by 8 pixel matrix contains the worth of the pixel at the relating (x, y) location. When these coefficients are plotted, we can directly view all the pixel frequencies in a single shot. By playing out the DCT on the input data, we have concentrated on the portrayal of an image in the upper left

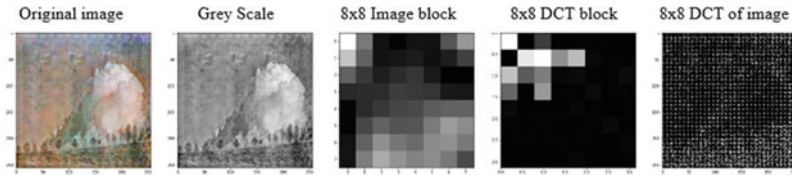


Fig. 8 (DCT of CycleGAN output)

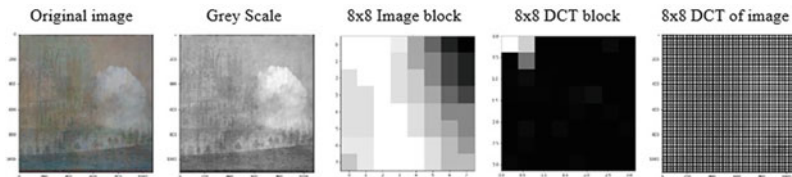


Fig. 9 (DCT of APP output)

coefficients of the result matrix, with the lower right coefficients of the DCT matrix containing fewer valuable data and thus redundant. For the below image (CycleGAN output), an 8×8 original pixel block and its corresponding DCT coefficient block has been plotted for visualization.

Figures 8 and 9 shows the output from DCT.

In Figs. 8 and 9 if we go from left to right, we will see the output (CycleGAN/APP) image, Grey scale, 8×8 Image block, DCT block and DCT'S of the image. Thus, we can see how from DCT block images, the high frequency components of the image are concentrated in the upper left corner coefficients in the coefficients plot.

From all the above metrics we can see that the image generated by CycleGAN is far better, smoother and sharper than that of the APP. The APP can only mix the images plainly to some extent but CycleGAN has the capability to improve the images a lot more as it has been trained on numerous images.

9 Discussion

DFT

Centered spectrum shows frequency components present in the image. Central part has low frequency components which is white in colour, while lines show dominant frequency components present, amount of white part at corners gives sharpness of edges present in image.

DCT

We centralized the spectrums of the Traditional Approach and CycleGAN images. One of the notable differences is the uniform illumination in the CycleGAN output. It also shows the presence of several strong horizontal components, indicating that edges are better detected. Looking at the CycleGAN output, we can see that the style has been transferred very well, while retaining the structure of the objects in the original image. On the opposite end, the rudimentary mixing of the Traditional Approach does not result in a good image. As a result, the spectrum shows no strong presence of any components. Converting both of these images back using Inverse FFT, we can see the structural integrities of both images. As expected, the CycleGAN image is well recovered, but the Traditional Approach image suffers from a degrade in quality after removing colour.

PSNR and SSIM

APP_20 stands for APP image where foreground image weight 80% and noise 20%.

Results of PSNR and SSIM testing can be analysed by taking into consideration three different cases. In first case, the PSNR score of CycleGAN image and original image is more than the PSNR score of APP_20 image and original image, this is the most frequent case encountered in the testing which suggest that the likelihood of a GAN generated image being more similar to original image than APP_20 generated image is high. Second case is an exceptional case, here the PSNR score of APP_20 image and original is more than that of CycleGAN image and original case. Third case like second case is also an exception, here the PSNR score of CycleGAN image and original image is much higher than the PSNR score of Traditional Approach_20 image and original image. SSIM testing follows the same pattern as that of PSNR testing resulting in the same conclusion as that of PSNR testing.

10 Conclusion

On a preliminary basis, the images generated by CycleGAN are quite better in quality, perception and in aesthetic point of view. Using PSNR, DFT analysis it is also mathematically proven that CycleGAN images have sharper edges and has more similarity with foreground image hence better mixing. Thus, the CycleGAN implemented here as high amount of capability and capacity to generate better images than just simply mixing foregrounds and backgrounds. Thus, we conclude given a foreground and background, CycleGAN can smartly blend the two images so that they are portrayed in quite good aesthetic fashion which has high chance of user selection.

11 Future Scope

The current implementation of the CycleGAN network does not focus on any specific application, however if required, the user can train the network according to his requirements, which are specialized and focused on a particular field or area, as in that case, the network will surely give more catered results and preferred images according to the styles used for training it.

References

1. Gatys LA, Ecker AS, Bethge M (2015) A neural algorithm of artistic style. arXiv preprint [arXiv:1508.06576](https://arxiv.org/abs/1508.06576)
2. Goodfellow I et al (2014) Generative adversarial nets. *Adv Neural Inf Process Syst* 27
3. Zhu J-Y et al (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE international conference on computer vision*
4. Coup S, Vetrova V, Frank E, Tappenden R (2019) Domain specific transfer learning using image mixing and stochastic image selection. Presented at the the sixth workshop on fine-grained visual categorization (FGVC6), computer vision and pattern recognition conference (EVPB 2019), Long Beach, CA
5. Gatys LA, Ecker AS, Bethge M (2016) Image style transfer using convolutional neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*
6. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)
7. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
8. Karras T et al (2020) Analyzing and improving the image quality of stylegan. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
9. Zhang L, Wen T, Shi J (2020) Deep image blending. In: *IEEE/CVF winter conference, applications of computer vision (WACV)*, pp 231–240
10. Zeng S, Hu H, Zhange J, Huang K (2019) GP-GAN towards realistic high resolution image blending. In: *27th ACM international conference on multimedia*
11. Tmenova O, Martin R (2019) CycleGAN for style transfer in X-ray angiography published in *springer*
12. Kaggle Team “I’m Something of a Painter Myself” version 1. <https://www.kaggle.com/c/gan-getting-started/data>
13. Pandian AP (2021) Review on image recoloring methods for efficient naturalness by coloring data modeling methods for low visual deficiency. *J Arti Intell* 3(3):169–183
14. Drori I, Cohen-Or D, Yeshurun H (2003) Example-based style synthesis. *2003 IEEE computer society conference on computer vision and pattern recognition, Proceedings, vol 2. IEEE*
15. Darney PE, Jeena Jacob I (2021) Rain streaks removal in digital images by dictionary based sparsity process with MCA Estimation. *J Innov Image Process* 3:174–189
16. Gireesan G, Mathew LS (2019) Stratified meta structure based similarity measure in heterogeneous information networks for medical diagnosis. In: *International conference on computational vision and bio inspired computing*. Springer, Cham, pp 66–70
17. Mahajan AD, Chaudhary S (2019) Image context based similarity retrieval system. In: *International conference on computational vision and bio inspired computing*. Springer, Cham, , pp 1265–1272
18. Irfan S, Ghosh S (2019) Efficient ranking framework for information retrieval using similarity measure. In: *International conference on computational vision and bio inspired computing*. Springer, Cham, pp 1344–1354

19. Rouse D, Hemami SS (2008) Understanding and simplifying the structural similarity metric. Proceedings international conference image process, pp 325328
20. Chen G, Yang C, Xie S (2006) Edge-based structural similarity for image quality assessment. Proceedings international conference acoustics. Speech Signal Process. pp. 14–19
21. Sara U, Akter M, Uddin MS (2019) Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. Proc J Comput Commun 7(3)
22. Narwaria M, Lin W, McLoughlin IV, Emmanuel S, Chia L (2012) Fourier transform-based scalable image quality measure. IEEE Trans Image Process 21(8):3364–3377
23. Agarwal S, Girdhar N, Raghav H (2021) A novel neural model based framework for detection of GAN generated fake images. 2021 11th international conference on cloud computing, data science & engineering (Confluence), pp 46–51
24. Tomosada H, Kudo T, Fujisawa T, Ikehara M (2021) GAN-based image Deblurring using DCT loss with customized datasets. IEEE Access 9:135224–135233

Botnet Attacks Detection Using Embedded Feature Selection Methods for Secure IOMT Environment



A. Karthick Kumar, K. Vadivukkarasi, R. Dayana, and P. Malarvezhi

Abstract Security remains a primary responsibility, despite the expanding number of IoT devices. The IoMT (Internet of Medical Things) ecosystem is subdivision of IoT which is made up of medical computing devices, software applications, and healthcare systems. The most critical criteria in an IoMT (Internet of Medical Things) environment are privacy and security. A security breach in a healthcare network might directly result in the death of patients, the Internet of Things (IoMT) necessitates enhanced protection. A botnet is a collection of internet-connected devices used to carry out security breaches, leak sensitive information, and give botnet attackers control over IoMT devices. Malware distributed via botnets often includes network communication features that allow attackers to connect with other threat actors via the botnet's vast network of infected devices. Using the UNSW-NB15 Dataset, we proposed a ESLRFBM (Efficient and Secured LASSO and RFBM) model for predicting botnet attacks in an IoMT scenario. In the UNSW-NB-15 dataset, we employed the embedded feature selection methods LASSO and RIDGE to detect botnet attacks. To detect security concerns, the IOMT has evolved and used machine learning technologies. We created and evaluated performance metrics such as Hamming loss and AUCROC scores using both classic and hybrid classifiers. By comparing and analysing the results using LASSO and RFBM properties, we can determine that our described model performs better.

Keywords IoMT (Internet of Medical Things) · Botnet Attacks · DoS · RIDGE · LASSO · Security · Health care systems

A. Karthick Kumar · K. Vadivukkarasi (✉) · R. Dayana · P. Malarvezhi
Department of ECE, SRM Institute of Science and Technology, KTR Campus, Chennai, India
e-mail: vadivukk@srmist.edu.in

A. Karthick Kumar
e-mail: ka8753@srmist.edu

R. Dayana
e-mail: dayanar@srmist.edu.in

P. Malarvezhi
e-mail: malarvip@srmist.edu.in

1 Introduction

Internet of Things (IoT) is a big revolution that is gaining traction across a variety of sectors, including education, business, personal surroundings, and healthcare [1, 2]. A secure IoT ecosystem is necessary to prevent sensitive information leaks and data theft. IoMT has made its way into healthcare over time, and it has a bright future ahead of it, where it will be widely used and widely distributed outside of hospital walls [3]. Botnet attacks are a major security concern in the IoT/IoMT industry, offering major risks to IoT smart devices and user data. Any exposed IoMT vulnerability can be used by cybercriminals to carry out a variety of unlawful operations, such as taking control of the medical device and accessing sensitive patient health, personal, and insurance data [3]. To establish a link between the attacker and the user, the botnet uses DNS (Domain Name System) (C&C). The best strategy to predict and avoid botnet attacks is to use feature engineering [4]. The model's accuracy and efficiency in terms of identification and classification are assessed using the CICDDoS 2019 dataset. Based on an analysis of the botnet's vulnerability form and vulnerability dissemination technique, a vulnerability mining and confrontation methodology for botnet has been developed [5, 6]. To predict cyber attacks, the UNSW-NB15 IoT botnet dataset is employed. With a false-negative rate of only 1.5%, the botnet detection system can detect up to 92% of the f1 score [7]. Our research attempts to close the knowledge gap so that a reliable botnet attack detection model can be developed. The selection procedures RIDGE and LASSO were employed in this work to extract the most important features [8]. AdaBoost, Decision Tree, Gradient Boosting, K- Nearest Neighbors, and Random Forest are among the supervised models used in this study, as are hybrid classifiers (Random Forests Bagging Method) RFBM and Decision Tree Bagging Method (DTBM). The findings are compared to those of previous studies.

The following is the paper's structure: The research's objectives are outlined in Sect. 2. A overview of employing classifiers and hybrid approaches to forecast botnet attacks is included in Sect. 3. In Sect. 4, the suggested Research Strategies as well as many performance metrics are detailed. Section 5 discusses the system's implementation performance comparisons. The results and discussions as explored in Sect. 6, respectively. Section 7 concludes with several suggestions for future research.

2 Research Goals

The goal of this research is to develop a reliable botnet attack prediction strategy in an IoMT setting.

- (i) In this research, data from the UNSW-NB15 Dataset was used.
- (ii) RIDGE and LASSO Regularization procedures are two ways for obtaining the most significant properties based on rank values. This also aids in the resolution of overfitting and underfitting concerns in machine learning.
- (iii) Bagging and boosting are two methods for reducing testing time and increasing execution speed.
- (iv) The findings of the Ridge and LASSO approaches were used to assess the performance of various models.

3 Literature Review

The UNSW NB-15, In this Research is the IoT Botnet attacks dataset, which was used to detect cyber attacks. Machine learning algorithms are utilized to effectively detect botnet attacks. Machine learning methods, particularly shallow and deep learning, are being utilized to combat tweet-based botnets [9]. In tough covert DDoS attack scenarios; we put the proposed detection and mitigation technique to the test. According to empirical data, oversampling based on GANs (generative adversarial networks) is useful for learning adversarial evasion attempts on botnet detectors in advance [10].

The Outsourced Support Vector Machine Scheme with High Privacy Protection is a support vector machine solution for Internet of Things implementation. With an accuracy of 84%, it's on level with the GSVM (General SVM) [11]. For Internet of Medical Things applications, a privacy-preserving data gathering and analysis system has a 95.56% average packet delivery ratio. In the proposed SDP (Software Defined Perimeter) describes network access controls and connection between SDNs (software defined Network infrastructures). PS: This is a DDoS attack test bed scenario (Port Scanning attacks). The SDN and SDP offer 75% network throughput while protecting against flooding attacks [12]. Nagarathna Ravi et al. offer LEarning-driven DETection Mitigation (LEDEM), a semi-supervised machine learning-based solution for detecting and mitigating DDoS attacks. DDoS attacks are now being detected at a rate of 96% [13].

4 Research Strategy

The discussion is centred on developing a machine learning model for predicting botnet attacks in the IoMT environment.

4.1 The Proposed Model

The healthcare industry is confronted with a number of issues. At all times, medical and patient information must be kept private. Medical IoT devices provide patients with unique procedures, medical device monitoring, and better care at a lower cost. Many smart medical devices, on the other hand, have undisclosed security features. In the context of IoMT, healthcare medical devices offer enormous benefits, but they also pose substantial security and privacy concerns. Cybercriminals can penetrate the hospital's network and get unauthorized access to sensitive personal and medical data by exploiting holes in IoMT devices.

The processes and strategies mentioned below show how to identify botnet attacks in an IoMT context.

Step: 1-Choose the UNSW NB15 Dataset for the proposed model since it contains 49 unique features.

Step: 2-The trained Random Forest Bagging Method (RFBM) hybrid model uses features from the uploaded data to generate inputs. The trained model only processes a subset of attributes.

Step: 3-'0' and '1' are generated by Botnet Attack detection. '0' indicates that the condition is normal, '1' indicates that a botnet attack has been identified.

Step: 4 -Data is uploaded to a database and used to construct a trained model to improve the accuracy of hybrid classifiers and trained models.

In Fig. 1 the Data gathering and preprocessing to detect botnet attacks in IoMT environment. Ridge and LASSO are two alternate feature selection algorithms that are intended to address overfitting concerns and shorten execution times. To analyze the dataset and choose the best model, many training models have been provided. The RFBM classifier was deemed the most successful, with a 99.21% ROC value and 4.9% of hamming loss values.

4.2 Performance Metrics

In Fig. 2 Performance metrics can be used to evaluate the machine learning model's effectiveness in hamming loss and ROCAUC score values.

True Positive-TP: Botnet attacks were correctly identified.

True Negative-TN: No botnet attacks were identified.

False Positive-FP: Botnet attacks that were incorrectly identified occurred.

False Negative-FN: Botnet attacks were incorrectly identified as normal attacks.

The Hamming loss L_{Hamming} between two samples is defined as: y_{ij} is the predicted value for the j th label of a given sample, y_j is the matching real value, and n_{labels} is the number of classes or labels.

$$\text{Hamming Loss} = L_{\text{H}}(xx') = f(x) = \frac{1}{n_{\text{labels}}} \sum_{j=0}^{n_{\text{labels}}-1} 1(\hat{y}_j \neq y_j) \quad (1)$$

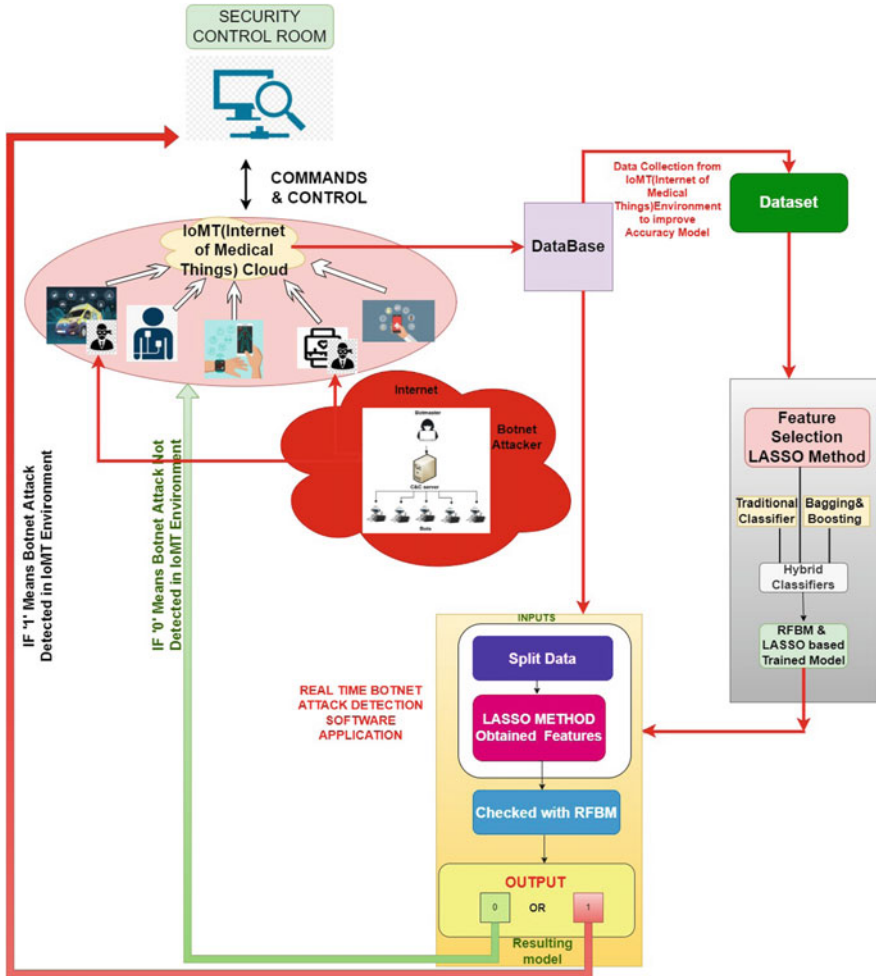


Fig. 1 Proposed model working diagram

| | | NEGATIVE VALUES | |
|------------------|-------------------|-----------------|----------|
| | | POSITIVE | NEGATIVE |
| PREDICTED VALUES | NEGATIVE POSITIVE | TP | FP |
| | NEGATIVE | FN | TN |

Fig. 2 Confusion matrix

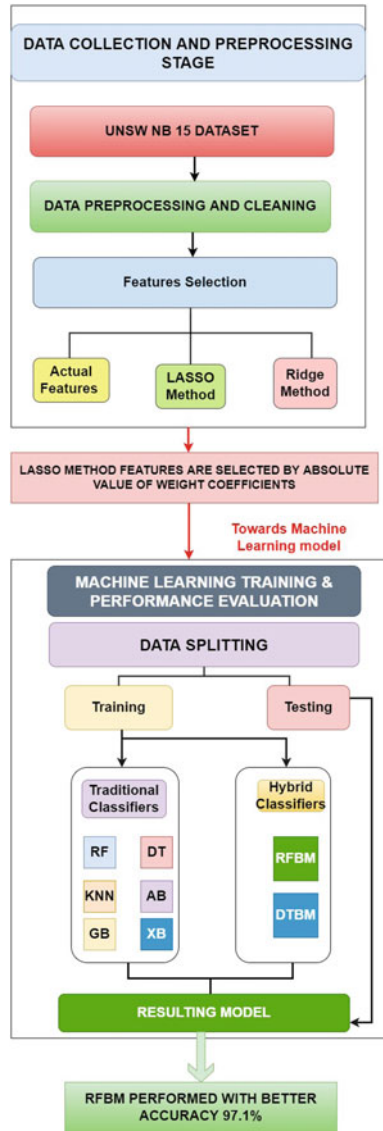


Fig. 3 General framework model

For a prediction model utilising multiple probability thresholds, ROC Curves highlight the trade-off between true positive rate and false positive rate.

$$TPR = TP / (TP + FN) \tag{2}$$

$$FPR = FP/(FP + TN) \tag{3}$$

4.3 *Justification of Proposed Model*

The suggested model demonstrates how feature extraction techniques may be applied to botnet attack prediction. Figure 3 shows how botnet attacks are detected in the HDO environment. Medical equipment connects to cloud servers, which store the collected data, in the Internet of Medical Things (IoMT). In the RFBM (Random Forest Bagging Method) model, feature extraction techniques such as LASSO are used. Following the selection of attributes, the RFBM model was used to forecast attacks. If the output is '0,' there are no botnet attacks in IoMT devices; if the output is '1,' the botnet attacker has set up a botnet and taken control of IoMT/IoT devices. Traditional classifiers like Random Forest, Decision Tree, KNN, XGBoost, ADABOOST, GRADBoost, and hybrid classifiers like RFBM and DTBM are utilised.

5 Implementation

To implement the model, which was constructed in Google Colab and Python, simple modules such as Panda, Pyplot, and Scikit-learn were employed. UNSW-NB15 is a network intrusion dataset having training set 175,341 records, whereas the testing set having 82,332 records, which includes botnet attacks and normal records. It is a computer network security dataset having 2,540,044 examples of real time and anomalous (also known as attack) network behaviour [14, 15]. From raw network packets, Argus and Bro-IDS software gathered 49 features, including packet and flow- based data fuzzers, analysis, backdoors, DoS, exploits, generic, reconnaissance, shellcode, and other sorts of attacks are among the nine categories of attacks [16]. Table 1 describes The UNSW NB-15 data set statistics include the simulation duration, total source and destination bytes, number of normal and abnormal records. Fuzzers, analysis, backdoors, DoS, exploits, generic, reconnaissance, shellcode, and other types of assaults are among the nine types of attacks included.

Table 1 Dataset records distributions

| Name of features descriptions | |
|-------------------------------|---|
| Normal | Actual data transaction between source to destination |
| Fuzzers | Accessing the device at random with malicious, software codes |
| Analysis | Attacker damage the network by malicious events |
| Backdoors | Bypasses standard authentication mechanisms |
| Exploits | It is a piece of software that creating a flaw software, hardware |
| Generic | A approach that works against all blockciphers |
| Reconnaissance | To discover information about the network's susceptible areas |
| Shellcode | To propagate to additional systems, the attacker copies itself |
| Worms | To steal sensitive information by taking advantage of security software |

6 Feature Selection Techniques

Certain variables in a dataset are nearly never beneficial for developing a machine learning model in the actual world [17]. The process of minimising the number of input variables in a prediction model is known as machine learning feature selection. The benefits of feature selection strategies are as follows: It enables the machine learning system to learn at a faster rate. We favour embedded methods, LASSO, and Ridge over the other feature extraction techniques because they are more accurate. The wrapper methods are highly costly and have a high chance of overfitting. Filters do not put any algorithm to the test; instead, they rank the initial characteristics according to their relationship to the problem (labels) and only select the top of the list.

Figure 4 All Machine Learning approaches that utilize feature selection during their training stage are included in embedded methods. Embedded approaches combine filtering and wrapper techniques

6.1 Ridge Feature Selection Method

Ridge and Lasso are two popular regularisation approaches for creating compact models with a large number of features. Ridge regression employs a technique known as 'L2 regularisation. It adds a penalty to the squared magnitude of the coefficients. In the optimization objective, it adds a sum of squares of coefficients factor. 'L2 regularisation' is a technique used in ridge regression. It penalises the coefficients' squared magnitude.

$$\text{Objective} = \text{RSS} + \alpha^* (\text{sum of square of coefficients}) \quad (4)$$

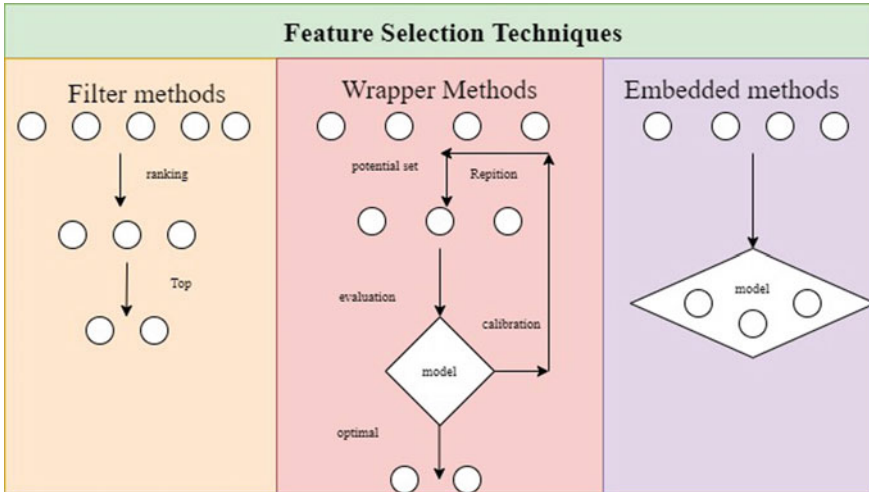


Fig. 4 Feature selection techniques

The importance of RSS minimization versus sum of squares of coefficients minimization is determined by the parameter Alpha.

6.2 Lasso Feature Selection Method

In Lasso regression, L1 regularisation is utilised, which imposes a penalty equal to the amount of the coefficients absolute value. The practice of applying a penalty to the various parameters of a machine learning model for reduction of the model’s freedom and avoid overfitting is known as regularisation. Lasso regression is used to accomplish Least Absolute Shrinkage and Selection Operator (LASSO(L1) regularisation [18, 19]. It adds a factor to the total of absolute coefficient values in the optimization target.

$$\text{Objective} = \text{RSS} + \alpha^*(\text{sum of absolute of coefficients}) \tag{5}$$

The fundamental difference between ridge and lasso regression is that for the same alpha values, the coefficients of lasso regression are significantly smaller than those of ridge regression. Lasso does feature selection in addition to decreasing coefficients.

7 Ensemble Methods of Machine Learning

IoMT demands improved protection Because, unlike other businesses, a security breach in a healthcare network can directly result in the loss of lives,. Ensemble approaches can be used to cope with some of the variables, such as uncertainty and bias. The bagging and boosting ensemble processes were used to get more accurate results. When the goal is to reduce the variance of Decision Tree classifiers, bagging is employed. Their Decision Tree is trained using randomly selected subsets of data. The average of all tree predictions is applied. It aids in the reduction of overfitting as well as the proper handling of data with larger dimensionality. Three ensemble hybrid models based on DT, RF, and KNN are created with the help of the Bagging approach. The bagging algorithm is used in Random Forest, which is an Ensemble Learning approach. Before integrating the findings, it grows as many trees as possible on a portion of the input. As a result, overfitting and variation in decision trees are reduced, resulting in improved accuracy.

8 Results and Discussions

Table 2 shows the LASSO score for the 32 most relevant features for botnet attack prediction. Table 3 The sum of squares of coefficients factor is included in ridge feature selection approaches with 5 features. Five machine learning classifiers and two hybrid techniques were applied to all attributes of the UNSW NB-15 dataset. A 2×2 confusion matrix was built to produce the different performance metrics and compare all of the strategies. A method for evaluating binary classification challenges is the Receiver Operator Characteristic (ROC) curve. The AUC is a summary of the ROC curve that assesses the ability of a classifier to distinguish between classes.

Table 2 Features selected by LASSO Regularization techniques

| Feature name | Feature description | Feature score |
|------------------|---|---------------|
| dur | Record total duration | 0.01218 |
| proto | Protocol for transactions | 0.04265 |
| service | Contains the network services | 0.3968 |
| state | Contains the state and the procedure that is depending on it | 1.9459 |
| spkts | Count of packets sent from the source to the destination | 0.001823 |
| dpkts | Count of packets from source to destination | 0.03429 |
| sttl | Time to live value from source to destination | 0.010662 |
| dttl | Time to live value from destination to source | 0.1324 |
| sloss | Retransmitted or discarded source packets | 0.007141 |
| dloss | Retransmitted or discarded destination packets | 0.04422 |
| sinpkt | Arrival time of source interpacket (mSec) | 0.0002885 |
| djit | Jitter at the destination (mSec) | 0.0000872 |
| swin | TCP window advertisement value (source) | 0.09747 |
| dwin | TCP window advertisement value (destination) | 0.02605 |
| tcprtt | Round-trip time for establishing a TCP connection | 1.0237 |
| synack | In a TCP connection, the time between SYN and SYN ACK packets | 1.0912 |
| smean | Mean of packet size transmitted from source | 0.000668 |
| dmean | Mean of packet size transmitted from source | 0.0009179 |
| Trans_depth | Represents the http request/response transaction's pipelined depth | 0.031816 |
| Ct_srv_src | Connections with the same service,destination address | 0.02833 |
| Ct_state_ttl | Connections with each state based on a particular range of source/destination time to live values | 1.8746 |
| Ct_dst_ltm | Time based, connections with the same destination address | 0.001814 |
| Ct_src_dport ltm | The connections that have the same source, destination port | 0.02561 |
| Ct_dst_sport ltm | The connections with the same destination, source port | 0.3483 |
| Ct_dst_src_ltm | Time based, same source address | 0.069 |
| Is_ftp_login | If the user and password are used to access the FTP session, then 1; else, 0 | 1.00298 |
| Ct_ftp_cmd | Flows in an FTP session that have a command | 0.0008131 |
| Ct_flw_http mthd | The flows in the http service that have methods like Get, Post | 0.1862 |
| Ct_src_ltm | Time based, connections with the same source address | 0.01592 |
| Ct_srv_dst | Time based, connections with the same service and destination address | 0.039 |
| Is_sm_ips_ports | This variable has a value of 1 if the source, destination IP addresses are the same and the port numbers are the same, or 0 | 1.237 |

Table 3 Features selected by RIDGE regularization techniques

| Feature name | Feature description | Feature score |
|--------------|--|---------------|
| sbytes | Bytes transferred from the source to the destination | 0.00000225 |
| dbytes | Packet count from destination to source | 0.00000309 |
| rate | Record total duration | 0.0000155 |
| dload | bits per second of the destination | -0.0000036 |
| sinpkt | Arrival time of source interpacket (mSec) | -0.0000014 |

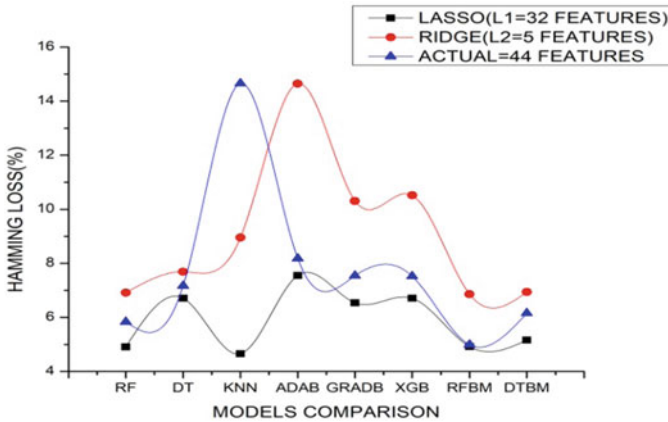


Fig. 5 Hamming loss values comparison

8.1 Hamming Loss Values Comparison on Different Methods

The fraction of labels that are wrongly predicted is known as the Hamming loss, and it is used to calculate the average Hamming loss or Hamming distance between two samples (Fig. 5).

A better classifier has a lower hamming loss value. The lowest hamming loss rate is generated by RFBM with 4.9% in LASSO regularization with 32 features and highest hamming loss value generated by Adaboost with 7.5%. In RIDGE with 5 features lowest error rate generated by RFBM with 6.8%.

8.2 ROCAUC Values Comparison on Different Methods

The Receiver Operator Characteristic (ROC) curve is a probability curve that displays TPR vs. FPR, and effectively separates the "signal" from the "noise". The highest ROCAUC value is generated by RFBM with 99.21% in LASSO regularization with 32 features and lowest value generated by Decision Tree with 93.49%. Figure 6

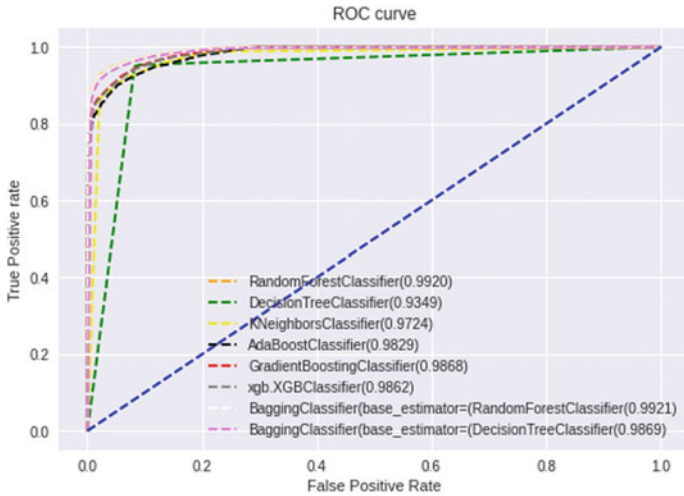


Fig. 6 ROCAUC values comparison lasso values comparison

RIDGE with 5 features highest ROCAUC vaule generated by RFBM with 98.58% and lowest value generated by Decision Tree with 92.35%.

Figure 7 RIDGE with 5 features highest ROCAUC vaule generated by RFBM with 98.58% and lowest value generated by Decision Tree with 92.35%

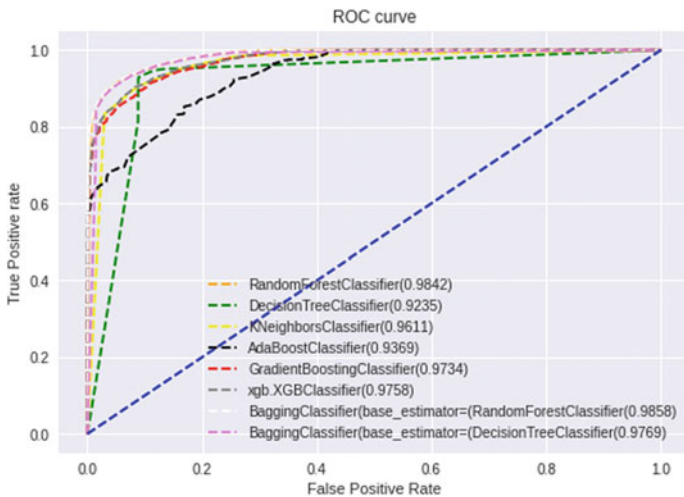


Fig. 7 ROCAUC values comparison ridge values comparison

Table 4 Comparison between existing and proposed system

| Model | LASSO | RIDGE | Actual | Other works |
|-------|-------|-------|--------|-------------|
| RF | 95.1 | 93 | 91.1 | 96 [9] |
| DT | 94.84 | 92.3 | 94 | 94 [12] |
| KNN | 92.7 | 91.0 | 90.3 | 90 [14] |
| ADAB | 93.9 | 86.5 | 87.1 | 77 [13] |
| GRADB | 92.1 | 86.6 | 87 | 91.1 [13] |
| XGB | 93.4 | 89.4 | 88.1 | 90 [16] |
| RFBM | 97.1 | 93.1 | 93 | 89 [19] |
| DTBM | 96 | 92 | 91 | 89.1 [19] |

9 Comparison with Existing Techniques

GridSearchCV has been employed in our proposed system, in order to achieve a better precision and accuracy (Table 4).

10 Conclusion and Feature Works

The capacity to detect botnet attack risk properly could have a big impact on the IoT/IoMT world. This study shows that the LASSO feature selection technique can be used to make accurate predictions. Several machine learning algorithms can be utilized with it. With 32 features, RFBM achieved a 99.21% ROCAUC score and lowest hamming loss rate is generated with 4.9%. In future, To handle datasets with a considerable amount of missing data, we aim to generalize the model using several feature extraction algorithms. Deep Learning algorithms are another potential option to consider.

References

1. <https://www.softwaretestinghelp.com/best-iot-examples/>. Accessed 27 September 2021
2. <http://www.thalesgroup.com/en/markets/digital-identity-and-security/iot/iotsecurity#:text=IoT%20security%20is%20the%20practice,confidentiality%20of%20your%20IoT%20solution>. Accessed 21 September 2021
3. Li W, Jianjin, Lee J-H (2019) Analysis of Botnet domain names for IOT cyber security. IEEE Access 7:94658–94665. <https://doi.org/10.1109/ACCESS.2019.2927355>
4. Panda M, Mousa AAA, Hassanien AE (2021) Developing an efficient feature engineering and machine learning model for detecting IoTBotnet cyber attacks. IEEE Access 9:91038–91052. <https://doi.org/10.1109/ACCESS.2021.3092054>
5. Alharbi A, Alsubhi K, Botnet detection approach using graph-based machine learning. IEEE Access 9:99166–99180. <https://doi.org/10.1109/ACCESS.2021.3094183>

6. Khan AY, Rabialatif, Seemablatif, Shahzaibtahir, Tanzilasaba (2019) Malicious insider attack detection in IoTs using data analytics. *IEEE Access* 7:11743–11753. <https://doi.org/10.1109/ACCESS.2019.2959047>
7. Arunansivanathan, Hassanhabibgharakheili, Vijaysivaraman (2020) Managing IoT cyber-security using programmable telemetry and machine learning. *IEEE Trans Netw Serv Manag* 17(1): 60–74. <https://doi.org/10.1109/TNSM.2020.297121>
8. Al-Fuqaha M, Guizani M, Mohammadi MA, Ayyash M (2015) Internet of things: a survey on enabling technologies, protocols, and applications. *IEEE Commun Surv Tutor* 17(4):2347–2376. <https://doi.org/10.1109/COMST.2015.2444095>
9. Hussain F, Rasheedhussain, Syedalihassan, Ekramhossain (2019) Machine learning in IoT security: current solutions and future challenges. *Arxiv* 1904.05735 V1 [Cs.Cr]
10. Moustafa N, Turnbull B, Choo K-K (2019) An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet Things J* 6(3):4815–4830. <https://doi.org/10.1109/JIOT.2018.2871719>
11. Kaifanhuang, Yang L-X, Yang X, Yongxiang, Tang YY (2020) A low-cost distributed denial-of-service attack architecture. *IEEE Access* 8:42111–42119. <https://doi.org/10.1109/AC-CESS.2020.2977112>
12. Wazid M, Das AK, Rodrigues JJPC, Shetty S, Park Y (2019) IoMT malware detection approaches: analysis and research challenges. *IEEE Access* 7:182459–182476. <https://doi.org/10.1109/ACCESS.2019.2960412>
13. Chaabouni N, Mosbah M, Zemhari A, Sauvignac C (2019) Network intrusion detection for IoT security based on learning techniques. *IEEE Commun Surv Tutor* 21(3). <https://doi.org/10.1109/COMST.2019.2896380>
14. Chu Z, Han Y, Zhao K (2019) Botnet vulnerability intelligence clustering classification mining and countermeasure algorithm based on machine learning. *IEEE Access* 7:182309–182319. <https://doi.org/10.1109/ACCESS.2019.2960398>
15. Shakya S (2021) A self monitoring and analyzing system for solar power station using IoT and data mining algorithms. *J Soft Comput Parad* 3(2):96–109
16. Harkare A, Potdar V, Mishra A, Kekre A, Harkare H (2021) Methodology for implementation of building management system using IoT. In: *Evolutionary computing and mobile sustainable networks*. Springer, Singapore, pp 939–948
17. Rahal BM, Santos A, Nogueir M (2020) A distributed architecture for DDoS prediction and bot detection. *IEEE Access* 8:159756–159772. <https://doi.org/10.1109/AC-CESS.2020.3020507>
18. Randhawa RH, Aslam N, Alauthman M, Rafiq H, Comeau F (2021) Security hardening of botnet detectors using generative adversarial networks. *IEEE Access* 9:78276–78292. <https://doi.org/10.1109/ACCESS.2021.3083421>
19. Sidi L, Nadler A, Shabtai A (2020) MaskDGA: an evasion attack against DGA classifiers and adversarial defenses. *IEEE Access* 8:161580–161592. <https://doi.org/10.1109/AC-CESS.2020.3020964>

An Accuracy Based Comparative Study on Different Techniques and Challenges for Sentiment Analysis



Radha Krishna Jana and Saikat Maity

Abstract The use of social media has gained more popularity in the modern society. People provide their own opinions on different subjects such as, social issues, news, events and products, on the web. Sentiments are generated by large number of people in social media. The user's sentiments are useful in uplifting a society in educational, commercial and business aspects. Sentiment analysis is the procedure to analyze the sentiment of textual matter in social networks. Sentiment Analysis constitutes a tool to take out the requisite content from the web. This analysis is also performed to collect the opinion for the assessment. Various techniques exist to execute this job. Different challenges are observed when the sentiment analysis is computed. This paper surveys on the respective methods for Sentiment Analysis and the different challenges faced by each technique.

Keywords Sentiment analysis · Sentiment classification · Machine learning · Lexicon-based · Hybrid method · Accuracy

1 Introduction

Social relationships occur in a social media between users, over which multiple interactions arise. These interactions can be represented in a graph which signify information cascade [1]. Mobile social networking is a form of social networking where people with common interests meet and converse using a mobile phone. In Mobile social networking, communication occurs between different age and gender based people [2]. Online social networks take part in a major responsibility in the spread of information at a very large dimension [3]. A group of users interact and debate with each other across many subjects in online social network [4]. Rumors are mainly spread through media, internet or through relationships between individuals within social relationship networks [5]. Nowadays, the spreading and manipulating information over social networks have become a critical issue from various points

R. K. Jana · S. Maity (✉)
JIS University, Kolkata, West Bengal, India
e-mail: saikat.maity@jisuniversity.ac.in

of view [6]. The social networks show a major role on appropriate platform for information diffusion and knowledge sharing. Negative news diffusion spread faster than positive news diffusion in online social networks [7].

Sentiment Analysis (SA) is a versatile field of research in social network. It consists of social science, psychological science, machine learning and natural language processing. This analysis is a process to survey people's sentiments, feeling, and cognition. The number of social media users increases every day. The global users may reach up to 4.41 billion by 2025 [8]. Different types of contents such as textual matter, picture, audio and video data are loaded on the social media [9]. These types of data need to be prepared into a functional statistics with help of latest techniques [10]. Sentiment of text data is processed by Natural Language Processing (NLP). After that, it is sorted out into negative sentiment, positive sentiment or natural sentiment [11]. Document, sentence and feature levels are the various levels of analysis in sentiment analysis [12]. Machine learning, Lexicon-based and hybrid approaches are the key techniques for sentiment analysis.

The aim of this paper is to study different sentiment analysis techniques and compare their accuracy. This paper also presents different challenges to each techniques. Different topics related to the research in this paper is discussed sequentially. Section 2 describes the previous work. SA tasks are discussed in Sect. 3. Section 4 expresses the level of analysis in SA. Different techniques for SA are discussed in Sect. 5. Section 6 investigates the different challenges. Different techniques of SA are experimented and compared in Sect. 7. Section 8 concludes the paper with the conclusion and explains the future research in the work.

2 Related Work

In this segment, previous research work related to sentiment analysis have been reviewed. It also concentrates on different new methods and algorithms for sentiment analysis. Sentimental Orientation (SO) of opinion is a text classification method used for the management of complex structural phrases [13]. Emotions may depend on the amount of the expression and the type of the text [14]. The theoretical analysis of sentiment analysis techniques is based on noun, adverbs, adjectives [15]. Sentiment analysis was executed by applying an existing constructed annotated corpus, using quotation sentences by Raza et al. [16]. Different views of the sentiment analysis were examined with a focus on Czech by Arote Rutuja et al. [17].

Most of the research study discussed in this paper are based on sentiment analysis using machine learning. Machine Learning algorithms were applied for learning, analysis, and classification based on customer experience of product data by Yi and Liu [18]. For sentiment analysis, comparative tests of classification algorithms based on machine learning algorithm was performed by Yogi and Paudel [19]. Supervised machine learning performed sentiment analysis for fake positive or negative review identification [20]. Authors of Gujar and Pardeshi [21] proposed that the machine learning approach sets the path to social media for opinion mining by the Twitter API.

Paper [22] described that Machine learning research decides the Bitcoin's predictable price path in USD.

The research study [23] used the Support Vector Machines (SVM), Random Forest and Naive Bayes, to perform sentence level sentiment analysis of news stories and blogs [23]. The authors of Bansal et al. [24] analyzed the YouTube feedback for strong predictions using decision tree, K Nearest Neighbors and SVM. Artificial Neural Networks (ANN) and the unigram were performed as an extractor to achieve the high accuracy in paper [25]. To increase new score on customer product feedback, NLP, text data mining and clustering techniques were used by the authors of Kauffmann et al. [26]. Context analysis was used to categorize certain terms shifting from one dataset to another dataset [27]. The review [28] focused on transfer learning applications in sentiment analysis. The technique described in Sadhasivam and Kalivaradhan [29] used a combination of Naive algorithms, SVM and ensemble to determine the precision and speed of the output of the algorithm [29]. Natural language processing assisted to measure the tweet emotions, where they are positive, neutral, and negative, and then were graded according to the emotions by Suryawanshi et al. [30]. The logistic regression of Naïve Bayes was used to assess feelings, and were categorized according to the better precision of the scientific classification in paper [31]. Naïve Bayes algorithms can predict the e-sports sentiment as per Ardianto et al. [32].

In recent years, Fuzzy methods are introduced for sentiment analysis. A fuzzy based knowledge engineering model has been formed for sentiment classification of sentences by the reserchers of Yazdavar et al. [33]. For more refined outputs, the fuzzy membership degrees in sentiment analysis were used by Jefferson et al. [34]. For Detecting sentiment in social media, the Fuzzy rule-based systems were analyzed by Basha et al. [35]. The fuzzy logic technique was applied to classify the data with different types of commitment by Ghani et al. [36].

A few research on hybrid methods for SA have been studied. A hybrid model developed by Lai and Raheem [37] experimented on hotel review of customers feedback. The model introduced by Revathy [38] was used for online product reviews to prevent fake assessment. A dictionary base method was proposed by Rani and Gill [39] to improve the accuracy of classification.

Nowadays, sentiment analysis techniques are used for sentiment analysis of covid tweets. A huge number of people were affected by Covid 19 occurrence in the world. During that period, people shared their experiences in social media. A model was proposed by Chakraborty et al. [40] to analyse the feedback of people [40]. A deep learning technique was used in Sitaula et al. [41] for analysing Nepali covid tweets. In paper [42], a deep learning model was used for the analysis of public sentiments related to covid tweets which produced a better accuracy than SVM, LSTM & LR. A recurrent neural network based model used to classify the covid tweets was proposed by Nemes and Kiss [43]. A machine learning technique was used for the sentiment analysis of covid tweets in paper [44]. Paper [45] explains that the experiments for SA are appropriate only to the English language. The literature survey conducted shows that many research works have enhanced the accuracy of sentiment analysis.

3 Sentiment Analysis Tasks

There are various tasks in sentiment analysis which is shown in Fig. 1.

3.1 Subjective Classification

This is the main task in sentiment analysis. This task finds out the existing originality in the textual matter [46]. The main aim of this classification is to control undesired data [47]. This classification notices subjective clues and words. It transfers feeling or opinion like ‘economical’, ‘difficult, and ‘outstanding’ [39]. These indications are habituated to categorize text attributes as objective or subjective. A rule-based classifier that classifies a sentence into a subjective classification with the help of subjective clues has been introduced [48].

3.2 Sentiment Classification

Sentiment classification is the most essential research task in the sentiment analysis. The main sub-tasks of sentiment classification are polarity determination, cross-language classification and cross-domain classification. The aim of the polarity determination is to find out the opinion polarity of textual matter. This subtask is categorized as negative or positive [49]. A novel method has been introduced for polarity determination in sentiment classification which has various process involved, such as, collection of data, preprocessing, stop words removing, tokenization, stemming, part of speech tagging, review, scoring features, sentence scoring, review scoring, star ratings and classification [50]. The sentiment of a reference domain is anticipate by cross-domain analysis. A model is trained on original language dataset in cross-language analysis which is being tested on other languages [51].

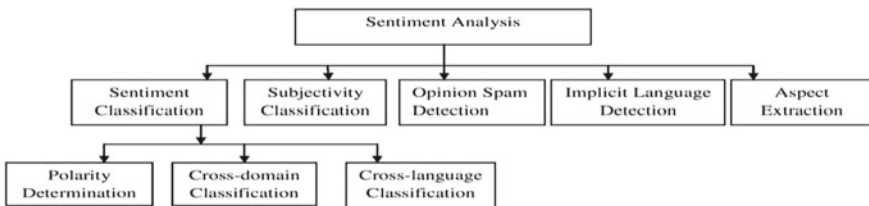


Fig. 1 Sentiment analysis tasks

3.3 *Aspect Extraction*

This method is used to recover the views of the reference text. People's opinions are needed to recognize for detailed sentiment analysis [52]. This method is essential for sentiment analysis in social network. There are no predefined content. Multiple methods exist for aspect extraction. Frequency-based analysis is the most traditional method for aspect extraction [51]. Syntax-based methods detect aspects utilizing the syntactic relations.

3.4 *Opinion Spam Detection*

For the increasing advancement of e-business, opinion spam detection has been generated. It is a superior topic in the area of sentiment analysis. The objective of this method is to point out the false opinions. It also points out the characteristics like assessment of the content, assessment of metadata, and practical knowledge about the product [52]. There are various process for detecting false opinion such as, collection of data, processing data and feature selection. After that, the machine learning algorithm is applied for detecting false opinions. Machine learning techniques are applied to review the content to disclose untrue views [51].

3.5 *Implicit Language Detection*

It detects comedy and criticism. The aim of the implicit language detection is knowing the information of an psychological feature. There are few conventional techniques for implicit language detection considered to investigate appearance, facial expression and punctuation mark [53].

4 *Level of Analysis*

Document, sentence, and aspect levels are the level of analysis in sentiment analysis. Figure 2 shows the different level of analysis in sentiment analysis that are discussed in the following paragraphs.

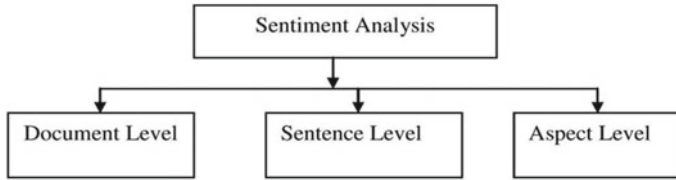


Fig. 2 Level of analysis

4.1 Document Level

This is the first level of sentiment analysis. This level is working on a document. The full document is taken to resolve the polarity. This level classifies the available opinion as a positive or negative sentiment. The important advantage of this level is maximum polarity of a specific feature can be received. This level did not receive the users interest and disinterest.

4.2 Sentence Level

This is another level working on a particular sentence. The level gives the mutual opposition of each sentence. Subjective sentence consists of users views and opinion. When a sentence does not suggest any opinion then it is called neutral. When a sentence is neutral it is called objective sentence. Subjective and objective classification are the important advantage of this level.

4.3 Aspect Level

This level can discover public interest and disinterest in the document level and in the sentence level. The outcome of this level are positive or negative or target value. A new algorithm (LDA) has been invented for aspect level analysis [51].

5 Traditional Approaches

Different approaches are used for sentiment analysis which have been picturised in Fig. 3.

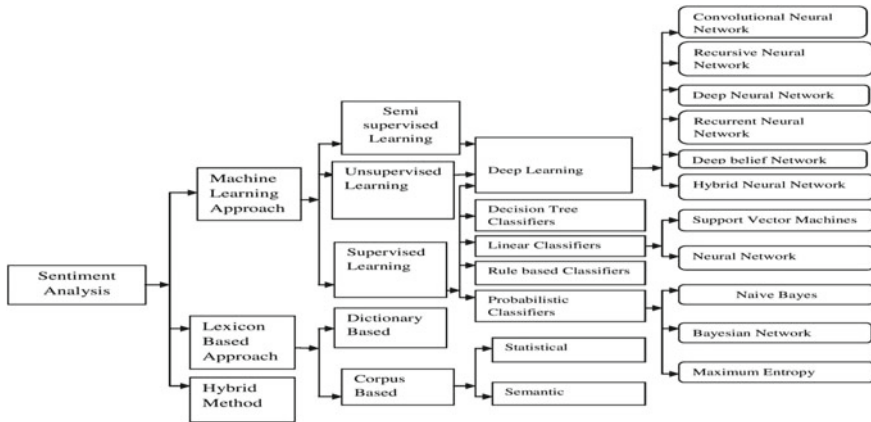


Fig. 3 Sentiment analysis approaches

5.1 Machine Learning Approach

Computer systems are adjusted and modified with experience, with the help of this learning technique. It also meets with the information theory, cognitive science and probability theory. This method is performed on trained data and test data. Lot of categorized data are require for executing this technique. This learning method is categorized as unsupervised, semi-supervised and supervised.

The unsupervised learning techniques are unlabeled similar information grouped into bunch [54]. Semi-supervised learning techniques are used for unlabelled and labelled data. Supervised learning is the most important learning methods. This technique is based on labeled dataset which is provided to the model for processing. For getting significant output, the labeled dataset are trained. Supervised learning is better than semi-supervised and unsupervised techniques.

The most important Machine learning algorithms for sentiment analysis are Support Vector Machines (SVM), Maximum Entropy, Naive Bayes, Decision Tree, N-gram, K-NN, Artificial Neural Network (ANN), Weight K-NN, K-Nearest, Logistic Regression (LR), Random Forest and Bayesian Network.

Deep learning algorithms have an important role in sentiment analysis. The main deep learning algorithms are Recursive Neural Network (RNN), RNTN, Simple Recurrent Networks (SRN), Convolutional Neural Network (CNN), Gated Recurrent Units, Latent Rating Neural Network, and Long-Short Term Memory.

There are various machine learning classifiers related to sentiment analysis. These are SVM, Naive Bayes and K-Nearest Neighbour.

SVM: It is a non-probabilistic binary linear classifier. The mathematical expression of SVM is shown in Eq. 1.

$$Y = f(x) = W^T X + b = \sum_{i=1}^n W_i X_i + b \quad (1)$$

Naïve Bayes: It is a probabilistic classifier that calculates probability using the Bayesian approach. Equation 2 represents the mathematical expression of Naive Bayes.

$$P(A/B) = [P(B/A) \times P(A)]/P(B) \quad (2)$$

where, A refers to class and B refers to data.

K-Nearest Neighbor: It classifies the data which are closest to each other. Equation 3 represents this algorithm.

$$D(q, p) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (3)$$

5.2 Lexicon Based Approach

This method uses opinion dictionary with its opinions words and match the words with the data to determine opinion polarity. They allocate sentiment of words as positive and negative. This method primarily depends on an opinion dictionary, and examines all the documents. Positive words are outstanding, nice etc. and negatives words are ‘bad’, ‘ugly’, ‘scary’ etc. Dictionary and corpus based techniques are the sub classification of this approach.

5.2.1 Dictionary-Based

The user gathers a set of opinion words in this approach, and then the seed list is processed by them. The user starts searching for dictionary to find synonyms and antonyms to define text. This technique cannot deal with domain oriented words.

5.2.2 Corpus-Based

This approach supplies lexicon affiliated words to a specific domain. This lexicon creates a set of source opinions. The authors established a polarity recognition method for distinguishing the polarities of few parties [55]. This lexicon based approach categorize sentiment at sentence level with the help of machine learning approach.

It gives better results than traditional machine learning and Lexicon techniques [56]. This technique establishes better result than support vector machine approach. The amount of dictionary data leads to an important role in this lexicon-based approach. The authors used more amount of data for higher accuracy value [57].

5.3 Hybrid Approaches

Machine Learning and Lexicon-based approaches are used in Hybrid-based approaches. There are various new lexicon-based techniques initiated to improve the classification of opinion. This hybrid approach assess the automatic recognition of the sentiments of english text [58]. The authors has given an SLU pipeline for managing entities and intents. The authors also established the effectiveness of new idea across the real world tasks [59]. This research demonstrates that hybrid approach provide higher accuracy than other classifiers [60]. There are various performance parameters related to sentiment analysis such as, precision, recall, F measure and accuracy.

Precision: It is the ratio of rightly grouped positive files to the total number of positive grouped files. Equation 4 gives the mathematical expression of precision.

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (4)$$

Recall: It is the ratio of total number of rightly categorized positive files to the total number of positive files. Equation 5 gives the mathematical presentation of recall.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (5)$$

The F measure: It is the harmonic mean of precision and recall. Equation 6 gives the mathematical expression of F measure.

$$\text{F measure} = 2 * \text{P recision} * \text{Recall}/(\text{Precision} + \text{Recall}) \quad (6)$$

Accuracy: It is the ratio of rightly categorized documents to the total number of documents. Equation 7 gives the mathematical demonstration of accuracy.

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (7)$$

Here, TP refers true positive, FP refers false positive, TN refers true negative and FN refers false negative.

6 Challenges

Sentiment analysis is taking out the sentiment words from the unstructured text inputs and categorize them into negative, positive or neutral. The sentiment topics must be defined from the input data by using the hypothesis testing methods. Then domain independent words such as =people or =said are to be removed. Normally phrases, adjectives and adverbs can show the sentiment of the text. The primary task of sentiment classification is to recognize and extract the sentiment topics from the unstructured data. The following challenges are considered in the past few years.

- a. Recognizing subjective portions of text
- b. Correlating sentiment with specific keywords
- c. Thwarted expressions
- d. Domain dependence
- e. Sarcasm detection
- f. Entity recognition
- g. Indirect negation of sentiment
- h. Structure
- i. World knowledge
- j. Opinion identification

7 Comparative Study

In this study, 39 previous researches are used for the comparison with various sentiment analysis techniques like Machine learning, Lexicon-based and Hybrid approaches. The first comparison investigates which technique gives the better accuracy. The second comparison examines sentiment analysis challenges and also related accuracy. Table 1 shows the study on different parameters related to different sentiment analysis technique.

Here, 15 latest research papers have been studied.

Table 1 Different sentiment analysis techniques

| Ref. No. | Methods | Algorithms | Datasets | Accuracy | Limitations |
|----------|------------------|---------------------------------------|------------------------------|----------|--|
| [61] | Machine learning | Random forest, SVM, Naive bayes | Twitter | 99 | Model performs good with the help of high-dimensional feature spaces and less data |
| [62] | Machine learning | Deep learning | Wordnet | 93.33 | During learning process, hidden layers become slow |
| [63] | Machine learning | SVM, NN, Naive Bayesian, KNN | Tweet review on Airline data | 88 | Not tested on large dataset |
| [31] | Machine learning | Logistic regression | IMDB | 88.4 | Algorithm gives better result on Count Vectorizer method |
| [64] | Machine learning | Supervised, unsupervised, Naive Bayes | Twitter data | 70 | The model was not tested on the frequent data |
| [57] | Lexicon based | Dictionary | Twitter and instagram | 87.78 | Dataset needs to be large for better performance |
| [65] | Lexicon based | Dictionary | Stadford | 72 | This model was not tested on GPOMS & opinion finder lexicon |
| [66] | Lexicon based | VADER | Snippets | 77 | This model works better in social media than text blob |
| [67] | Lexicon based | Improved lexicon-based method | Twitter | 88.2 | Domain independent polarity determination was not tested on multiple domains |
| [68] | Lexicon based | Semantic | Twitter | 95 | This model experimented only sports related tweets |
| [69] | Hybrid | Machine learning and lexicon-based | Memes | 86.53 | Model is not tested on gif, cartoon and animation memes |
| [70] | Hybrid | CNN, SVM | Amazon | 94.3 | Do not perform for domain dependent |

(continued)

Table 1 (continued)

| Ref. No. | Methods | Algorithms | Datasets | Accuracy | Limitations |
|----------|---------|---|---------------------------|----------|---|
| [71] | Hybrid | Random forest, Logistic regression, Bernouli NB, Multinomial NB | Surveys and websites | 90.93 | This model use only binary classification |
| [72] | Hybrid | DNN, LR, SVM | Hotel and Persian Product | 86.29 | Not tested on unclassified sentences |
| [60] | Hybrid | NB, SVM, GA | Restaurant | 93 | This model use only heterogeneous ensemble method |

The average accuracy of 87.74% for machine learning approach, 83.99% for lexicon approach and 90.21% for hybrid approach have been obtained. Figure 4 shows that the hybrid approach achieves better accuracy than lexicon based and machine learning approaches.

Table 2 shows the mapping between the accuracy and different challenges in sentiment analysis.

Figure 5 shows the average accuracy of different sentiment analysis challenges. The lowest average of accuracy is 78.80 for Huge lexicon.

Figure 6 presents the maximum accuracy with different challenges. Figure shows that the domain dependent gives the maximum accuracy and Huge lexicon gives the minimum accuracy.

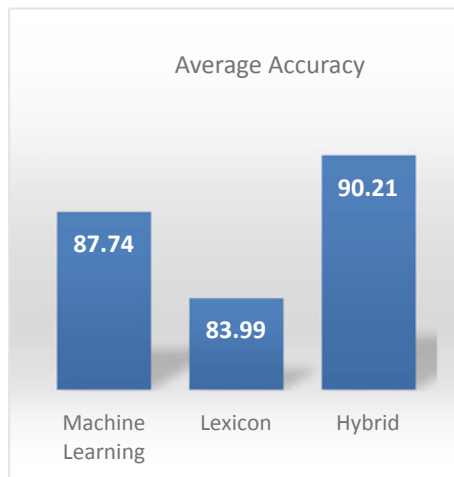


Fig. 4 Average accuracy of sentiment analysis approaches

Table 2 SA challenges

| Ref. No. | Challenges | Methodology | Data set | Accuracy |
|----------|--|---|-----------------------------------|------------|
| [72] | Word knowledge | Machine learning | Hotel | 86.29% |
| [73] | Word knowledge | Machine learning | ISEAR and other 9 datasets | 74.6% |
| [74] | Domain dependent | Machine learning | Kaggle toxicity detection dataset | 98.46 |
| [75] | Domain dependent | Machine learning | Twitter sentiment dataset | 89.46% |
| [76] | World knowledge + Lexicon + feature extraction | Machine learning | Real dataset | 83.5% |
| [77] | Domain dependent | SemEval-2013 | Twitter and SMS | 69% |
| [78] | NLP overheads | | Weibo microblog | 80% |
| [79] | Spam reviews | Lexicon-based | Store#364, | 85.7% |
| [80] | NLP overheads + Domain dependent | Lexicon-based | Social media website | 93.9% |
| [81] | Spam reviews | n-gram | 800 opinions | Nearly 90% |
| [82] | Domain dependence | Machine learning | Movie review | 90 |
| [83] | Domain dependence | Machine learning | | 94% |
| [84] | Huge lexicon | Machine learning | Movie | 82.30 |
| [85] | NLP overheads | Machine learning | Tweets | 82.7 |
| [86] | NLP overheads | NLP and Machine learning | Blog, review | 83% |
| [87] | Domain-dependence | n-gram | 560 Chinese review | 65 |
| [88] | World knowledge | | Game | 80% |
| [89] | Domain dependence | n-gram | Hotel | 83 |
| [90] | Data Sparsity | CNN | | 77.73 |
| [91] | Domain dependence | Dependency-Sentiment-LDA-Markov chain | Amazon product review | 70.7 |
| [92] | Huge lexicon | Lexicon based technique | Twitter | 73.5% |
| [93] | Data sparsity | Ensemble and Hybrid based SA analysis algorithm | Twitter | 85 |
| [94] | Feature and keywords extraction | NLP and Data mining | GooglePlay | 91% |

(continued)

Table 2 (continued)

| Ref. No. | Challenges | Methodology | Data set | Accuracy |
|----------|--------------|--------------|-----------|----------|
| [95] | Huge lexicon | New approach | Sentences | 75.9 |

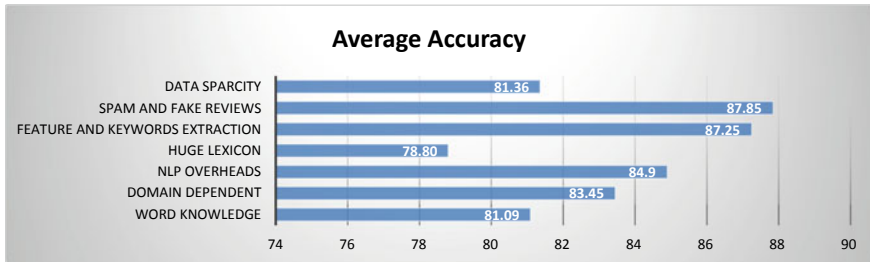


Fig. 5 The average accuracy of each SA challenges

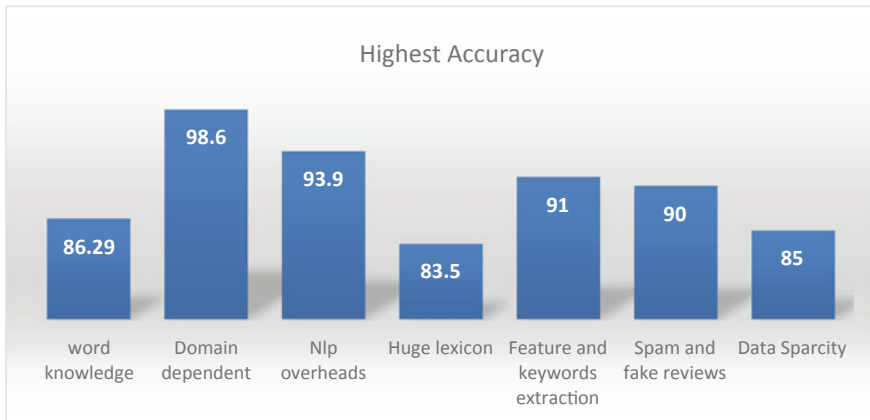


Fig. 6 Maximum accuracy of SA challenge

8 Conclusion

This study establishes that the hybrid approach achieves better accuracy than machine learning and lexicon based approaches. Here, a proper classification technique is selected based on the nature of data for designing a classification model. Moreover, a proper subset of feature is selected to increase the performance of the model without degrading the accuracy. There are other areas like classification of streaming and temporal data, neural networks and fuzzy based classifier that can be further investigated on the latest area of research.

The most desirable factor in sentiment analysis is accuracy of result. This study presents different factors such as, word knowledge, domain dependent, NLP overheads, huge lexicon, data sparcity, Feature extraction and Spam reviews, that impact the accuracy of sentiment analysis. These factors are addressed in 24 research papers. Huge lexicon has the lowest average accuracy and more studies are required on huge lexicon to increase the accuracy. This study may be helpful in defining the relevant factors when designing a classification model with accuracy and others performance parameters.

In future, hybrid approaches will be concentrated on, to enhance the performance of sentiment analysis. Furthermore, more investigation on different challenges and specially huge lexicon challenge will be performed. Additionally, huge number of recent research studies would be examined to conclude a model that constitutes for a higher performance than other models.

References

1. Taxidou I (2013) Realtime analysis of information diffusion in social media. In: Proceedings of the VLDB Endowment, vol 6
2. Dong Y, Yang Y, Tang J, Yang Y, Chawla NV (2014) Inferring user demographics and social strategies in mobile social networks. In: Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining. New York, pp 15–24
3. Dotey A, Rom H, Vaca (2011) CInformation diffussion in social media. Final Project CS224W Stanford University
4. Flamino J, Szymanski BK (2019) A reaction-based approach to information cascade analysis. In: Proceedings of 2019 IEEE international conference on computer communication and networks. Valencia, Spain, July 29–August 1, pp 1–9
5. Ibrahim RA, Hefny HA, Hassanien AE (2016) Controlling rumor cascade over social networks. In: Proceedings of the international conference on advanced intelligent systems and informatics. Advances in Intelligent Systems and Computing 533
6. Ibrahim RA, Hassanien AE, HefnyHA (2018) Controlling social information cascade: a survey, Book Chapter, Bog data analytics, pp 17
7. Arazkhani N, Meybodi MR, Rezvanian A (2019) influence blocking maximization in social network using centrality measures. In: IEEE 5th international conference on knowledge-based engineering and innovation. Tehran, Iran
8. Statista (2021) Number of social media users worldwide 2017–2025. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
9. Giri KJ, Lone TA (2014) Big data-overview and challenges. Int J Adv Res Comput Sci Softw Eng 4(6)
10. Sivarajah U, Kamal MM, Irani Z, Weerakkody V (2017) Critical analysis of big data challenges and analytical methods. J Bus Res 70:263–286
11. Agarwal B, Mittal N, Bansal P, Garg S (2015) Sentiment analysis using common-sense and context information. J Comput Intell Neurosci
12. Mishra N, Jha CK (2012) Classification of opinion mining techniques. Int J Comput Appl 56(13)
13. Jabbar J, Urooj I, Wu J, Azeem N (2019) Real-time sentiment analysis on e-commerce application. Int Conf Netw Sens Control
14. Tuhin RA, Paul BK, Nawrine F, Akter M, Das AK (2019) An automated system of sentiment analysis from bangla text using supervised learning techniques. Int Conf Comput Commun Syst

15. Sultana N, Kumar P, Patra M, Chandra S, Alam SKS (2019) Sentiment analysis for product review. *Int J Soft Comput* 9:1913–1919
16. Raza H, Faizan M, Hamza A, Mushtaq A, Akhtar N (2019) Scientific text sentiment analysis using machine learning techniques. *Int J Adv Comput Sci Appl* 10
17. Arote Rutuja S, Gaikwad Ruchika P, Late Samidha S, Gadekar GB (2020) Online shopping with sentimental analysis for furniture shop. *Int Res J Modern Eng Technol Sci* 02(05):1–8
18. Yi S, Liu X (2020) Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review. *Complex Intell Syst*
19. Yogi TN, Paudel N (2020) Comparative analysis of machine learning based classification algorithms for sentiment analysis. *Int J Innov Sci Eng Technol* 7
20. Yadav N, Arora MS (2020) The performance of various supervised machine learning classification algorithms in sentiment analysis of online customer feedback in restaurant sector of hospitality industry. *Int J Technol Res Eng* 7
21. Gujar MA, Pardeshi NG (2020) Review on a sentiment analysis and predicting winner for indian premier league using machine learning technique. *Int Res J Modern Eng Technol Sci* 2:963–967
22. Raju SM, Tarif AM (2020) Real-time prediction of BITCOIN price using machine learning techniques and public sentiment analysis
23. Shirsat V, Jagdale R, Shende K, Deshmukh SN, Kawale S (2019) Sentence level sentiment analysis from news articles and blogs using machine learning techniques. *Int J Comput Sci Eng* 7
24. Bansal A, Gupta CL, Muralidhar A (2019) A sentimental analysis for youtube data using supervised learning approach. *Int J Eng Adv Technol* 8
25. Gamal D, Alfonse M, El-Horbarty ESM, Salem AB (2019) An evaluation of sentiment analysis on smart entertainment and devices reviews. *Int J Inf Theor Appl* 26:147–164
26. Kauffmann E, Peral J, Gil D, Ferrández A, Sellers R, Mora H (2019) Managing marketing decision-making with sentiment analysis: an evaluation of the main product features using text data mining. *Sustainability*
27. Aziz AA, Starkey A, Madi EN (2020) Predicting supervise machine learning performances for sentiment analysis using contextual-based approaches. *IEEE Acces* 8:17722–17733
28. Liu R, Shi Y, Ji C, Jia M (2019) A survey of sentiment analysis based on transfer learning. *IEEE Access* 7:85401–85412
29. Sadhasivam J, Kalivaradhan RB (2019) sentiment analysis of amazon products using ensemble machine learning algorithm. *Int J Math Eng Manag Sci* 4:508–520
30. Suryawanshi R, Rajput A, Kokale P, Karve SS (2020) Sentiment analyzer using machine learning. *Int Res J Modern Eng Technol Sci* 2
31. Sentamilselvan K, Aneri D, Athithiya AC, Kumar PK (2020) Twitter sentiment analysis using machine learning techniques. *Int J Eng Adv Technol* 9
32. Ardianto R, Rivanie T, Alkhalifi Y, Nugraha FS, Gata W (2020) Sentiment analysis on e-sports for education curriculum using naive bayes and support vector machine. *J Comput Sci Inf* 13:109–122
33. Yazdavar AH, Ebrahimi M, Salima N (2016) Fuzzy based implicit sentiment analysis on quantitative sentences. *J Soft Comput Dec Supp Syst* 3
34. Jefferson C, Liu H, Cocea M (2017) Fuzzy approach for sentiment analysis. In: *IEEE international conference on fuzzy systems*. Italy
35. Basha SM, Zhenning Y, Rajput DS, Iyengar NChSN, Caytiles RD (2017) Weighted fuzzy rule based sentiment prediction analysis on tweets. *Int J Grid Distrib Comput* 10:41–54
36. Ghani U, Bajwa IS, Ashfaq A (2018) A fuzzy logic based intelligent system for measuring customer loyalty and decision making. *Symmetry* 10
37. Lai ST, Raheem M (2020) Sentiment analysis of online customer reviews for hotel industry: an appraisal of hybrid approach. *Int Res J Eng Technol (IRJET)* 7
38. Revathy R (2020) A hybrid approach for product reviews using sentiment analysis. *Adalya J* 9
39. Rani S, Gill NS (2020) Hybrid model for twitter data sentiment analysis based on ensemble of dictionary based classifier and stacked machine learning classifiers-SVM, KNN and C5.0. *J Theor Appl Inf Technol* 98

40. Chakraborty AK, Das S, Kolya A (2021) Sentiment analysis of Covid-19 tweets using evolutionary classification-based LSTM model. *Adv Intell Syst Comput* 1355
41. Sitaula C, Basnet A, Mainali A, Shahi TB (2021) Deep learning-based methods for sentiment analysis on Nepali COVID-19-related tweets. *Comput Intell Neurosci*
42. Chintalapudi N, Battineni G, Amenta F (2021) Sentimental analysis of COVID-19 tweets using deep learning models. *Infect Disease Rep* 13:329–339
43. Nemes L, Kiss A (2020) Social media sentiment analysis based on COVID-19. *J Inf Telecommun* 5:1–15
44. Machuca CR, Gallardo C, Toasa RM (2021) Twitter sentiment analysis on coronavirus: machine learning approach. *J Phys Conf Series*
45. Azeez NA, Victor OE, Junior UE (2021) Sentiment analysis of COVID-19 Tweets. *Fudma J Sci* 5:566–576
46. Kasmuri E, Basiron H (2017) Subjectivity analysis in opinion mining—a systematic literature review. *Int J Adv Soft Comput Appl* 132–159
47. Kamal A (2013) Subjectivity classification using machine learning techniques for mining feature-opinion pairs from web opinion sources. *Int J Comput Sci Issues* 10:191–200
48. Yue L, Chen W, Li X, Zuo W, Yin M (2019) A survey of sentiment analysis in social media. *Knowl Inf Syst* 60:617–663
49. Wang G, Sun J, Ma J, Xu K, Gu J (2014) Sentiment classification: the contribution of ensemble learning. *Decision Supp Syst* 57:77–93
50. Kausar S, Huahu X, Ahmad W, Shabir MY, Ahmad W (2020) A sentiment polarity categorization technique for online product reviews. *IEEE Access* 8
51. Lighthart A, Catal C, Tekinerdogan B (2021) Systematic reviews in sentiment analysis: a tertiary study. *Artif Intell Rev*
52. Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowl Based Syst* 89:14–46
53. Filatova E (2012) Irony and sarcasm: corpus generation and analysis using crowdsourcing. In: *LREC*
54. Li G, Liu F (2013) Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. *Appl Intell*
55. Sharma S, Bansal M (2020) multilingual lexicon based approach for real-time sentiment analysis. *Int J Recent Technol Eng* 9
56. Eng T, Nawab MRI, Shahiduzzaman KM (2021) Improving accuracy of the sentence-level lexicon-based sentiment analysis using machine learning. *Int J Scient Res Comput Sci Eng Inf Technol* 7:57–69
57. Kaswidjanti W, Himawan H, Silitonga PDP (2020) The accuracy comparison of social media sentiment analysis using lexicon based and support vector machine on souvenir recommendations. *Test Eng Manag* 83:3953–3961
58. Dadhich A, Thankachan B (2021) Sentiment analysis of amazon product reviews using hybrid rule-based approach. *Int J Eng Manuf* pp 40–52
59. Price R, Mehrabani M, Gupta N, Kim Y-J, Jalalvand S, Chen M, Zhao Y (2021) A hybrid approach to scalable and robust spoken language understanding in enterprise virtual agents. In: *Proceedings of NAACL HLT 2021: Industry track papers*, pp 63–71
60. Chauhan M, Paneri D (2021) Comparison of different hybrid approaches used for sentiment analysis: survey. *Int J Scient Res Eng Trends* 7
61. Andoh J, Asiedu L, Lotsi A, Chapman-Wardy C (2021) Statistical analysis of public sentiment on the ghanaiian government: a machine learning approach. *Adv Human Comput Interaction*
62. Qaiser S, Yusoff N, Ali R, Remli MA, Adli HK (2021) A comparison of machine learning techniques for sentiment analysis. *Turkish J Comput Math Edu* 12:1738–1744
63. Mourabit YE, Habouz YE, Zougagh MLH (2020) A new sentiment analysis system of tweets based on machine learning approach. *Int J Scient Technol Res* 9
64. Iyer KBP, Kumaresh S (2020) Twitter sentiment analysis on coronavirus outbreak using machine learning algorithms. *Eur J Mol Clin Med* 7

65. Al-Shabi MA (2020) Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions mining. *Int J Comput Sci Netw Secur* 20
66. Bonta V, Kumaresh N, Janardhan N (2019) A comprehensive study on lexicon based approaches for sentiment analysis. *Asian J Comput Sci Technol* 8:1–6
67. Yurtalan G, Koyuncu M, Turhan C (2019) A polarity calculation approach for lexicon-based Turkish sentiment analysis. *Turkish J Electr Eng Comput Sci* 1325–1339
68. Wunderlich F, Memmert D (2020) Innovative approaches in sports science—lexicon-based sentiment analysis as a tool to analyze sports-related twitter communication. *Appl Sci*
69. Prakash TN, Aloysius A (2021) Hybrid approaches based emotion detection in memes sentiment analysis. *Int J Eng Res Technol* 14:151–155
70. Raviya K, Mary Vennila S (2020) A hybrid deep learning approach for sentiment analysis using CNN and improved SVM with multi objective swarm optimization for domain independent datasets. *Int J Adv Trends Comput Sci Eng* 9
71. Joshi S, Patel M (2020) Sentiment detection on news data using hybrid approach. *J Eng Sci* 11:1148–1153
72. Dashtipour K, Gogate M, Li J, Jiang F, Kong B, Hussain A (2020) A hybrid persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks. *Neurocomputing* 380:1–10
73. Stojanovski D, Strezoski G, Madjarov G, Dimitrovski I (2015) Twitter sentiment analysis using deep convolutional neural network. *Int Conf Hybrid Artif Intell Syst*
74. Srivastava S, Khurana P, Tewari V (2018) Identifying aggression and toxicity in comments using capsule network. In: *Proceedings of the first workshop on trolling, aggression and cyber bullying*, pp 98–105
75. Dufourq E, Bassett BA (2017) Evolutionary deep networks for efficient machine learning. In: *2017 pattern recognition association of South Africa and robotics and mechatronics*. IEEE
76. El-Din DM, Mikhtar HMO, Ismael O (2015) Online paper review analysis. *Int J Adv Comput Sci Appl* 6
77. Kiritchenko S, Zhu X, Mohammad SM (2014) Sentiment analysis of short informal texts. *J Artif Intell Res* 50
78. Tang D, Qin B, Liu T, Shi Q (2014) Emotion analysis platform on Chinese microblog
79. Peng Q, Zhong M (2014) Detecting spam review through sentiment analysis. *J Softw* 9:2065–2072
80. Al-Kabi MN, Alsmadi IM, Gigieh AH, Wahsheh HA, Haider MM (2014) Opinion mining and analysis for Arabic language. *Int J Adv Comput Sci Appl* 5
81. Ott M, Choi Y, Cardie C, Hancock JT (2011) Finding deceptive opinion spam by any stretch of the imagination. In: *Proceedings of the 49th annual meeting of the association for computational linguistics*. Portland, Oregon, pp 309–319
82. He Y, Lin C, Alani H (2011) Automatically extracting polarity-bearing topics for cross-domain sentiment classification. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*. Portland, Oregon
83. Kanayama H, Nasukawa T (2006) Fully automatic lexicon expansion for domain-oriented sentiment analysis. In: *Proceedings of the 2006 conference on empirical methods in natural language processing*. Sydney, pp 355–363
84. Mudinas A, Zhang D, Levene M (2012) Combining lexicon and learning based approaches for concept-level sentiment analysis. In: *WISDOM'12, Beijing, China*
85. Mohammad SM, Turney PD (2010) Emotions evoked by common words and phrases: using mechanical Turk to create an emotion lexicon. In: *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*. Association for Computational Linguistics. Los Angeles, California, pp 26–34
86. Boiy E, Mones M-F (2009) A machine learning approach to sentiment analysis in multilingual web texts. *Inf Retrieval*
87. Yang J, Hou M (2010) Using topic sentiment sentences to recognize sentiment polarity in Chinese reviews. In: *CIPS-SIGHAN joint conference on Chinese language processing*

88. Balahur A, Steinberger R (2009) Rethinking sentiment analysis in the news: from theory to practice and back. In: Troyano, Cruz, Díaz (eds.), Womsa, pp 1–12
89. Kasper W, Vela M (2011) Sentiment analysis for hotel reviews. In: Proceedings of the computational linguistics-applications conference, Jachranka, pp 45–52
90. Yuan Z, Wu S, Wu F, Liu J, Huang Y (2018) Domain attention model for multi-domain sentiment classification. *Knowl Based Syst* 155:1–10
91. Li F, Huang M, Zhu X (2010) Sentiment analysis with global topics and local dependency. In: 24th AAAI conference on artificial intelligence, vol 24
92. Kaushik C, Mishra A (2014) A scalable lexicon based technique for sentiment analysis. *Int J Found Comput Sci Technol* 4
93. Alsaeedi A, Khan MZ (2019) A study on sentiment analysis techniques of twitter data. *Int J Adv Comput Sci Appl*
94. Guzman E, Maalej W (2014) How do users like this feature? A fine grained sentiment analysis of app reviews. In: 22nd IEEE international requirements engineering conference
95. Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on Human Language technology and empirical methods in natural language processing, pp 347–354

Pseudo-CT Generation from MRI Images for Bone Lesion Detection Using Deep Learning Approach



S. Sreeja and D. Muhammad Noorul Mubarak

Abstract One of the most reliable MR-based attenuation correction techniques utilised in clinical practice for PET quantification is Magnetic Resonance (MR) image segmentation. The brain cannot be divided into distinct tissue groups since the intensity levels of bone and air signals are so comparable. The purpose of this research is to overcome the issue of attenuation correction, which will allow pseudo-CT images to be generated from MRI data for the diagnosis of bone lesions. The image registration stage has been shifted to the pre-processing stage to do this. The Gray-Level Co-occurrence Matrix (GLCM) and Gray-Level Run-Length Matrix (GLRLM) were discovered to be features. Finally, the UNet++ , a deep learning model, which produces the most similar bone areas as depicted in CT scans, has innovatively been used for pseudo-CT creation.

Keywords Magnetic resonance image (MRI) · Computed tomography (CT) · Pseudo-CT · Positron emission tomography (PET)

1 Introduction

PET (Positron Emission Tomography) is a well-known imaging method for examining real-time molecular data. The PET acquisition procedure, on the other hand, results in homogeneous bias, lowering image resolution. Attenuation maps are often created from CT images because there is a clear transition between CT intensity and attenuation coefficients [1]. Magnetic Resonance Imaging (MRI) is often regarded as the greatest imaging technology for structural brain research due to its superior soft-tissue contrast, high spatial resolution, anatomical and functional information, and lack of ionising radiation. MR images are also used to diagnose, plan therapy, and monitor the progression of a variety of neurological illnesses, including brain tumours. Alzheimer's disease, Parkinson's disease, multiple sclerosis, stroke and other neurodegenerative illnesses are examples of neurodegenerative disorders.

S. Sreeja (✉) · D. M. N. Mubarak
Department of Computer Science, University of Kerala, Karyavattom campus, Trivandrum
695581, Kerala, India
e-mail: sreebhav@gmail.com

In comparison to CT, soft-tissue contrast in MRI improves target delineation precision and repeatability in a range of anatomic locations. A treatment planning technique based only on MRI might save money, limit radiation exposure to patients, and speed up clinical workflow by eliminating systemic MRI-CT co-registration issues. Despite these benefits, MRI lacks the unique and quantitative electron density information required for precise dosage estimations and the development of reference images for patient setup. As a result, in this MRI-only treatment planning technique, electron density information from CT must be produced from the MRI by generating a pseudo-CT (PCT) [2].

The development of precise methodologies for estimating PCT from MRI is crucial. Atlas-based techniques, classification-based methods, sequence-based methods, and learning-based approaches are the four types of estimating methods available. Techniques based on atlases use single or several atlases with deformable registrations. Electron density can also be assessed using pattern recognition algorithms [3]. The fundamental issue is that their precision is dependent on the accuracy of inter-subject registration [4]. Many techniques are unable to identify bone from air consistently due to the uncertain intensity link between bone and air in MRI [5]. Classification-based systems manually or automatically divide MR images into bone, air, fat, and soft-tissue classes, then assign a consistent electron density to each class. However, because the intensity connection between bone and air in MRI is unknown, these methods may not be able to discriminate bone from air accurately.

In sequence-based techniques, PCTs are constructed using intensity data from standard MR, specialized MR sequences such as the Ultra-short Echo Time (UTE) [6–8], or a mix of the two. The current imaging quality of UTE sequences, on the other hand, makes distinguishing between blood arteries and bone problematic. Learning-based techniques use a training data set to create a CT-MRI map, which is then used to predict PCT for the target MRI [9–11]. A new technique for learning from UTE MR Sequences using Gaussian Mixture Regression (GMR) has just been published. The quality of the GMR approach was unsatisfactory due to the lack of spatial information.

The brain is difficult to separate into different tissue classes because the strength levels of bone and air signals are so near. As a consequence, the study's main contribution is to overcome the problem of attenuation correction in bone lesion identification by creating pseudo-CT images from MRI images. This was accomplished using the pre-processing component of the image registration procedure. Finally, the UNet++ model has been newly adapted for pseudo-CT generation, which produces the most similar bone regions as shown in CT images.

The following is the structure of the paper: Sect. 2 analyses current relevant studies, Sect. 3 details the suggested technique, Sect. 4 deliberates on the results and Sect. 5 closes the study with the conclusion.

2 Literature Survey

Some current works of literature on pseudo-CT generation, attenuation adjustments based on machine learning, and deep learning approaches are covered in this part.

Leynes et al. [12] proposed combining patient-specific multi-parametric MRI, such as Dixon MRI and proton-density-weighted zero echo time MRI (ZTE MRI),

with a deep learning system to create pseudo-CT images in real-time (ZeDD CT). A 3-T time-of-flight PET/MRI system was used to scan 26 people. The RMSE for evaluating bone lesions with PET was lowered in half (10.24% for Dixon PET and 2.68% for ZeDD PET). Liu et al. [13] developed and assessed deep learning approaches for MR imaging-based attenuation correction in brain PET/MR imaging. Before being tested on 10 people, the model was trained using 30 retrospective three-dimensional T1-weighted head scans, and the produced faux CT scan was compared to an obtained CT scan. Deep MRAC was used to identify PET reconstruction difficulties (-0.71.1). Yang Lei et al. [14] developed a machine learning technique for creating synthetic CT (sCT) images for MRI-only radiation therapy planning. The dense-cycle-GAN architecture was constructed using the dense-block approach on a cycle-generative adversarial network (cycle-GAN) architecture in this study. Data from 14 patients who had both was used to test the suggested approach. In addition to constructing an initial pseudo-CT, Leynes et al. [15] used MR data to create pseudo-CT uncertainty estimations using a Bayesian deep convolutional neural network. For bone lesions, UpCT-MLAA demonstrated an acceptable but slightly higher RMSE (3.4% versus 2.7%) than ZeDD-CT. Wiesinger et al. [16] investigated the morphological presentation and segmentation of cranial bone structure using Proton Density (PD)-weighted zero echo time (ZTE) imaging. Torrado-Carvajal et al. [17] built a network to link the four Dixon MR images (water, fat, in-phase, and out-of-phase) to their 2D CT equivalents. Except for bone tissue, the absolute mean relative change findings of the DIVIDE method were less than 2% on average in every region of concern. Balasubramaniam et al. [18] focused on the fundamental characteristics of ECG telemetry and offers an internet of things (IoT)-based application for monitoring patient health in both an indoor and outdoor setting. The IoT module enables data management and data transfer in the telemetry, and gathered data is saved in the cloud to improve service quality. Vijayakumar et al. [19] employed image augmentation and restoration to discover skin diseases, particularly skin cancers, which are extremely rare and difficult to diagnose. Despite the availability of modern technologies for image capture, images are still exposed to noise and imperfections, making it difficult to assess the real information that is to be shown. It's especially difficult to discern between the healthy and the damaged areas in uncommon skin malignancies like Merkel cell skin cancer, which grows swiftly if not treated early. Sungeetha et al. [20] developed a deep learning based approach for automatic segmentation for Lung CT Image. Koresh et al. [21] proposed a capsule network based hybrid algorithm for OCT corneal image classification and segmentation.

Through analysis, it is observed that, Leynes et al. [12] ZeDD CT does not perform as well with soft-tissue lesions as with bone lesions. RIRE [23] It's probable that the co-registered CT scan doesn't represent a real quantitative reconstruction standard. Yang et al. [14] High time complexity and [15] performed only for three patients, which is not applicable for real-time. Wiesinger et al. [16] High variability on CT acquisition in terms of vendor, resolution, and filter. Torrado-Carvajal et al. [17] The generated model may not be applicable to a pediatric population. Thus the author was motivated to develop a technique for accurately extracting the bone lesion from magnetic resonance images.

3 Methodology

The proposed methodology includes the processes of pre-processing where the image registration is being concentrated, then the feature extraction has been performed for extracting up to third-order features and finally, the pseudo-CT generation has been performed with the deep learning approach for bone lesion detection. The block diagram of pseudo-CT generating framework using MRI images is shown in Fig. 1.

The Retrospective Image Registration Evaluation (RIRE) dataset provided the MRI and CT images used in this study. The RIRE dataset includes CT, MRI, T1-weighted, T2-weighted, and PET images from the modalities described in Table 1. While certain files, such as OASIS and the Alzheimer’s Neuroimaging Initiative [22] have both MRI and CT data for the same individual, public databases having both modalities for the same person are uncommon. MRI and CT images from the same subject are only found in the RIRE project [23] and the Cancer Imaging Archive [24]. Since the RIRE project uses a uniform data format, we employed it for our investigation. In some cases, PET scans can be utilised in combination with CT scans. Although some MRIs had been fixed, we did not use them. When they provided the most participants, T1-weighted MRI and CT were chosen as input and

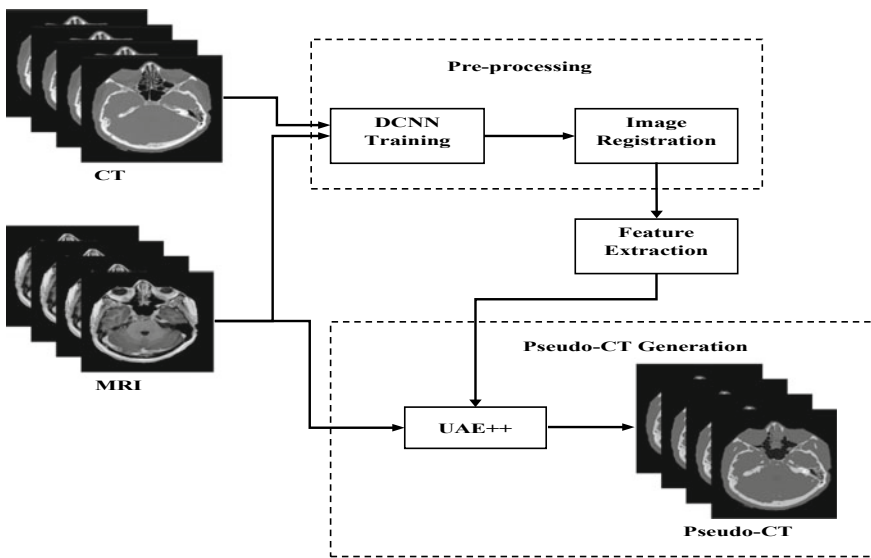


Fig. 1 Block diagram of the proposed methodology

Table 1 RIRE subject counts in terms of image modalities

| CT | MRI PD | T1w | T2w | MRI MP RAGE | PET |
|----|--------|-----|-----|-------------|-----|
| 17 | 14 | 19 | 18 | 9 | 8 |

goal data. Providing several MRIs as multi-channel input, on the other hand, would be an intriguing experiment.

3.1 Pre-Processing of MRI and CT Images

T1-w and T2-w images from each MR volume are converted into two-dimensional slices. In 25 volumes, there are 149 slices, but only 112 slices in four volumes. The N4 bias correction technique [25] is used in MR images to reduce bias field distortions. The intensity normalising approach is then used to decrease variation in MR images from various patients. The brightness of the MR images is increased to a value between [0, 100]. CT images are processed in the same way as other imaging techniques. There were 2982 2D images in the collection. Only 2328 slices were used after removing sections of the start and end slices of each volume because they contained tiny pixels that correspond to brain tissue and the remainder of the pixels correlate to the backdrop.

3.1.1 Image Registration

In multi-modal image analysis, image registration is critical. Since we processed the MRI and CT images for pseudo-CT generation, we've focused on image registration. It's a method for merging information from MRI and CT images into a single coordinate system. It is now feasible to compare and integrate data from a variety of sources directly. Several studies have developed ground truth data using established algorithms to exclude human input [26, 27].

3.2 Feature Extraction

The image characteristics were collected as part of the procedure to construct feature vectors for each split area. This research resulted in a precise and complete mapping of the transition from MR to CT areas. As a result, a neighbourhood was generated for each pixel in the MR image, and the imaging tissue's properties were obtained in that region. One of the most important aspects used in image pattern identification is texture, which is the event of a linked group of pixels. Texture characteristics provide information about a region's intensity variation. Tables 2, 3 and 4 list the grey-level co-occurrence matrix-based features, Table 3 lists the grey-level run-length matrix-based features, and Table 4 lists the histogram statistical features. The intensity of the grey values in the pixels in the reference image is used to produce first-order statistical indices that are unaffected by their spatial connection. The normalised histogram may be analysed as a probability density function and quantified using a variety of statistical methods. Although a histogram-driven statistical indices study

Table 2 Grey-level co-occurrence matrix-based features

| Features | Equation |
|-------------|--|
| Contrast | $\sum_{a=1}^N \sum_{b=1}^N (a - b)^2 P(a, b)$ |
| Correlation | $\frac{\sum_{a=1}^N \sum_{b=1}^N (a - m_R)(b - m_C) P(a, b)}{\sigma_R \sigma_C}$ |
| Energy | $\sum_{a=1}^N \sum_{b=1}^N P(a, b)^2$ |
| Smoothness | $\sum_{a=1}^N \sum_{b=1}^N \frac{P(a, b)}{1 + a - b }$ |
| Entropy | $\sum_{a=1}^N \sum_{j=1}^N P(a, b) P(a, b)$ |

Table 3 Grey-level co-occurrence matrix-based features

| Features | Equation |
|------------------------------|---|
| Short-run emphasis | $\frac{1}{n_R} \sum_{a=1}^M \sum_{b=1}^N \frac{P(a, b)}{b^2} = \frac{1}{n_R} \sum_{b=1}^N \frac{P_R(b)}{b^2}$ |
| Long run emphasis | $\frac{1}{n_R} \sum_{a=1}^M \sum_{b=1}^N P(a, b) \cdot j^2 = \frac{1}{n_R} \sum_{b=1}^N P_R(b) \cdot b^2$ |
| Grey-level non-uniformity | $\frac{1}{n_R} \sum_{a=1}^M \left(\sum_{b=1}^N P(a, b) \right)^2 = \frac{1}{n_R} \sum_{a=1}^M P_G(a)^2$ |
| Low gray-level run emphasis | $\frac{1}{n_R} \sum_{a=1}^M \sum_{b=1}^N \frac{P(a, b)}{a^2} = \frac{1}{n_R} \sum_{a=1}^M \frac{P_G(b)}{a^2}$ |
| High grey-level run emphasis | $\frac{1}{n_R} \sum_{a=1}^M \sum_{b=1}^N P(a, b) \cdot a^2 = \frac{1}{n_R} \sum_{a=1}^M P_G(a) \cdot a$ |

Table 4 Histogram statistical features

| Features | Equation |
|--------------------|--|
| Smoothness | $R = 1 - \frac{1}{1 + \sigma^2}$ |
| Standard deviation | $\mu_2 = \sum_{a=0}^{L-1} (z_i - m)^2 \rho(z_i)$ |
| Third moment | $\mu_3 = \sum_{a=0}^{L-1} (z_i - m)^3 \rho(z_i)$ |
| Uniformity | $U = \sum_{a=0}^{L-1} \rho^2(z_i)$ |
| Entropy | $e = -\sum_{a=0}^{L-1} \rho(z_i) \rho(z_i)$ |

of image texture gives important information on geographical distribution, it lacks data on the relative spatial connection. The grayscale image was used to establish four distinct orientations (0°, 45°, 90°, and 135°). The matrix’s dimension is equal to the image’s existing grey level’s maximum value. For some statistical indicators, the matrix may be seen as a probability density function.

Since it is a probability density function, the GLRLM would be utilised to generate third-order statistical moments for texture feature extraction. The number of runs in a run-length matrix that contains pixels of grey level I, where N is the maximum run length and M is the number of grey levels that contain pixels of run-length b. As

run-number vectors, the run-length run-number vector $P_R(b)$ and the grey-level run-number vector $P_G(b)$ are also provided. The cumulative distribution of the number of runs with grey level I and run-length b is shown in these two vectors.

$$P_R(b) = \sum_{a=1}^M P(a, b) \quad (1)$$

$$P_G(b) = \sum_{a=1}^N P(a, b) \quad (2)$$

Since the bone and air regions in MRI are comparable, polar coordinates are an independent characteristic that may be addressed in the training phase. The MRIs have the same intensity levels, but in the polar coordinate system, the border between the skull and the air space shifts and is more visible. After feature extraction, the Principal Component Analysis (PCA) method is employed to reduce data dimensionality linearly. Instead of examining all attributes, the PCA may find the essential components and analyse a subset of more valuable information combinations that have the largest influence on the final result. The following sequence will help you identify each major component:

$$PCA_a = a_{a1}X_1 + a_{a2}X_2 + \dots + a_{an}X_n \quad (3)$$

The desired component, related eigenvector, and initial independent variables are represented by PCA_a , a_{ij} and X_n respectively.

3.3 Pseudo-CT Generation

In this research, the UNet++ architecture [28] was employed to create pseudo-CT. The UNet++ model intends to achieve multiple bio-medical activities by introducing Dense block and convolution layers between the encoder and decoder of the original UNet architecture.

Between each U-net++ layer, however, there are several skip connection nodes. Each skip connection unit gets all feature maps from preceding units on the same level, and an upsampled feature map from the one below it. Throughout the network levels, each convolutional layer is made up of multiple three-dimensional filters that extract high-level brain features. As a result, each level is made up of a large number of thick blocks. The design of the UNet++ architecture for pseudo-CT generation is depicted in Fig. 2.

The UNet++ model is differing from conventional UNet in three ways. UNet++ having the convolution layers and dense skip connections on skip pathways, also having deep supervision, which enables model pruning. These advancements bridge the semantic gap between encoder and decoder feature maps, and then improves

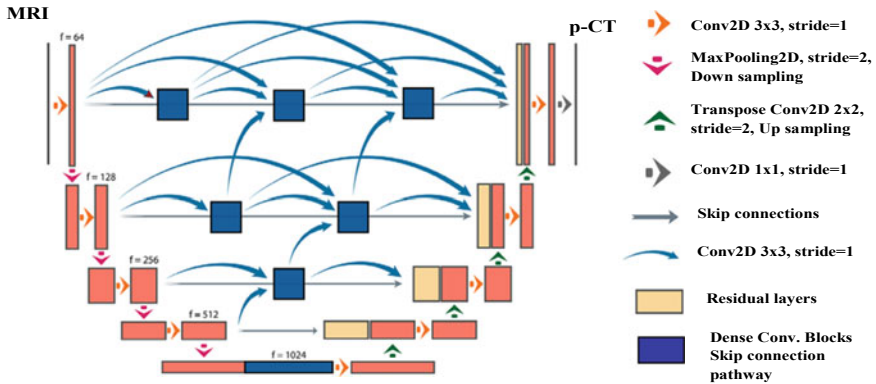


Fig. 2 UNet++ architecture

the gradient flow also improves the comparable performance by using only one loss layer.

The skip connection unit operates as follows, where X denotes the feature map and i and j denote the index down the contracting path and across the skip connections, respectively.

$$X^{a,b} = \begin{cases} H(X^{a-1,b}), & b = 0 \\ H\left(\left[X^{a,k}\right]_{k=0}^{b-1}, U(X^{a+1,b-1})\right), & b > 0 \end{cases} \quad (4)$$

Convolution and activation are represented by $H(\cdot)$, up-sampling is represented by $U(\cdot)$, and concatenation is represented by $[\cdot]$. The number of intermediary skip connection units is related to the number of layers, and it decreases linearly as the path gets shorter. The pseudo-CT image was created at the last layer of the expansion route with the accurate identification of bone lesions.

3.4 Evaluation Methods

The voxel-wise Mean Absolute Error (MAE) and the Mean Error (ME) are the most fundamental and extensively used measuring scales for evaluating the quality of a pseudo-CT image:

$$MAE_{\text{vox}} = \frac{1}{N_C} \sum_{n=1}^{N_C} |CT(n) - pCT(n)| \quad (5)$$

$$ME_{\text{vox}} = \frac{1}{N_C} \sum_{n=1}^{N_C} CT(n) - pCT(n) \quad (6)$$

4 Results and Discussion

The model was created using Windows 10, an AMD Ryzen 7 5800H-9 5900HS CPU, Nvidia GeForce RTX 3060-3080 graphics card, up to 32 GB of RAM, and MATLAB for modelling. For a single pCT from MRI, one epoch of neural network training takes 50 iterations to run through all of the training data once, resulting in a computation time of 37 s. The Mean Absolute Error (MAE) and Mean Error (ME) values for the total body were calculated to accomplish this. The RIRE dataset was utilised to create pseudo-CT investigations, as previously stated. MRI image of 12 subjects were taken into account for pseudo-CT generation.

The output image derived from the UNet++ model is depicted in Fig. 3 as a pseudo-CT generation, in which the bone lesions are successfully predicted from the input MRI image. The best accurate identification of bone lesion was achieved by performing absolute image registration with DCNN and feature extraction up to third-order statistical features with GLCM and GLRLM features.

The mean square error obtained by the proposed UNet++ based pseudo-CT generation for bone lesion detection has been depicted in Fig. 4. The mean absolute error is obtained as 10.105 at the training phase and 9.72 at the testing phase.

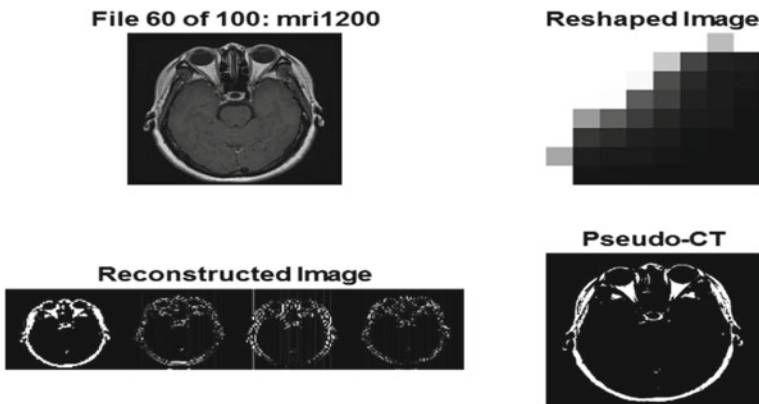


Fig. 3 UNet++ output as generated pseudo-CT image

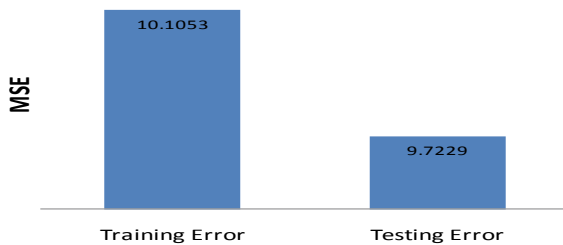


Fig. 4 Mean square error

5 Comparison Analysis

The suggested model’s performance was compared to that of existing deep learning models such as VGG16, ResNet, AlexNet, DenseNet, CNN, and UNet. Figure 5 show the computing time of several models for generating pseudo-CT.

Figure 6 represents the comparison of SSIM similarity measures between the original CT and the resulting pseudo-CT images. From the figure, it is understood that the UNet ++ can obtain 91% of similarity to the original CT image.

The above Fig. 7 represents the comparison of mean absolute error (HU) and mean square error (HU) of the proposed UNet++ based deep learning framework with the existing models for a pseudo-CT generation.

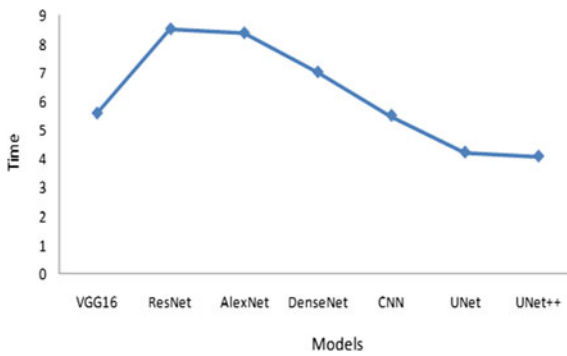


Fig. 5 Comparison of Computation

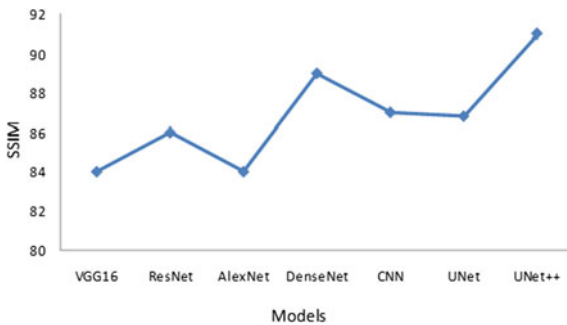


Fig. 6 Comparison of SSIM time

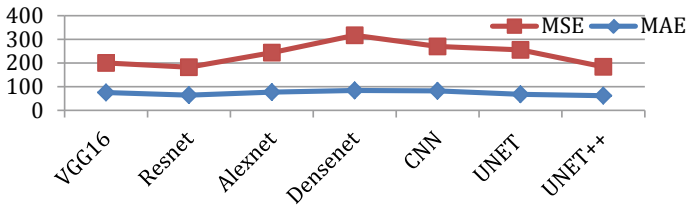


Fig. 7 Comparison of MAE (HU) and MSE(HU)

6 Conclusion

This research suggests a novel way for producing pseudo-CTs in order to use magnetic resonance imaging to detect bone abnormalities. In contrast, the images were registered using a deep convolutional neural network. To improve prediction accuracy, third-order statistical characteristics called the Grey-Level Co-occurrence Matrix (GLCM) and grey-level run-length matrix (GLRLM) are used. Finally, the UNet++ model was utilised to produce pseudo-CT scans from MRI images for the first time. The results revealed that the training and testing mean square error for the obtained pseudo-CT image are 10.105 and 9.72 respectively.

References

1. Fei B, Yang X, Nye JA, Aarsvold JN, Raghunath N, Cervo M, Stark R, Meltzer CC, Votaw JR (2012) MR/PET quantification tools: registration, segmentation, classification, and MR-based attenuation correction. *Med Phys* 39(10):6443–6454
2. Andreasen D et al (2015) Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. *Med Phys* 42(4):1596–1605
3. Uha J et al (2014) MRI-based treatment planning with pseudo CT generated through atlas registration. *Med Phys* 41(5):051711
4. Hofmann M et al (2008) MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration. *J Nucl Med* 49(11):1875–1883
5. Khateri P et al (2015) Generation of a four-class attenuation map for MRI based attenuation correction of PET Data in the head area using a novel combination of STE/Dixon-MRI and FCM clustering. *Mol Imaging Biol* 17(6):884–892
6. Aitken AP et al (2014) Improved UTE-based attenuation correction for cranial PET-MR using dynamic magnetic field monitoring. *Med Phys* 41(1):012302
7. Yousefi Moteghaed N, Mostaar A, Azadeh P (2021) Generating pseudo-computerized tomography (P-CT) scan images from magnetic resonance imaging (MRI) images using machine learning algorithms based on fuzzy theory for radiotherapy treatment planning. *Med Phys* 48(11):7016–7027
8. Kronthaler S, Boehm C, Feuerriegel G, Börner P, Katscher U, Weiss K, Makowski MR, Schwaiger BJ, Gersing AS, Karampinos DC (2021) Assessment of vertebral fractures and edema of the thoracolumbar spine based on water-fat and susceptibility-weighted images derived from a single ultra-short echo time scan. *Magn Reson Med* 87(4):1771–1783

9. Gholamiankhan F, Mostafapour S, Arabi H (2021) Deep learning-based synthetic CT generation from MR images: comparison of generative adversarial and residual neural networks. [arXiv:2103.01609](https://arxiv.org/abs/2103.01609)
10. Matsuo H, Nishio M, Nogami M, Zeng F, Kurimoto T, Kaushik S, Wiesinger F, Kono AK, Murakami T (2021) Unsupervised-learning-based method for chest MRI-CT transformation using structure constrained unsupervised generative attention networks. [arXiv:2106.08557](https://arxiv.org/abs/2106.08557)
11. Fei Y, Zhan B, Hong M, Wu X, Zhou J, Wang Y (2021) Deep learning-based multi-modal computing with feature disentanglement for MRI image synthesis. *Med Phys* 48(7):3778–3789
12. Leynes AP, Yang J, Wiesinger F, Kaushik SS, Shanbhag DD, Seo Y, Hope TA, Larson PEZ (2018) Zero-echo-time and Dixon deep pseudo-CT (ZeDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. *J Nuclear Med* 59(5):852–858
13. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB (2018) Deep learning MR imaging-based attenuation correction for PET/MR imaging. *Radiology* 286(2):676–684
14. Yang L, Wang T, Liu Y, Higgins K, Tian S, Liu T, Mao H, Shim H, Curran WJ, Shu H-K, Yang X (2019) MRI-based synthetic CT generation using deep convolutional neural network. In: *Proceedings of SPIE 10949, Medical Imaging 2019: Image Processing*, 109492T
15. Leynes AP, Ahn SP, Wangerin KA, Kaushik SS, Wiesinger F, Hope TA, Larson PE (2020) Bayesian deep learning uncertainty estimation and pseudo-CT prior for robust maximum likelihood estimation of activity and attenuation (UpCT-MLAA) in the presence of metal implants for simultaneous PET. MRI in the pelvis. [arXiv: 2001.03414](https://arxiv.org/abs/2001.03414)
16. Wiesinger F, Bylund M, Yang J, Kaushik S, Shanbhag D, Ahn S, Jonsson JH et al (2018) Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning. *Mag Res Med* 80(4):1440–1451
17. Torrado-Carvajal A, Vera-Olmos J, Izquierdo-Garcia D, Catalano OA, Morales MA, Margolin J, Soricelli A, Salvatore M, Malpica N, Catana C (2019) Dixon-VIBE deep learning (DIVIDE) pseudo-CT synthesis for pelvis PET/MR attenuation correction. *J Nuclear Med* 60(3):429–435
18. Balasubramaniam V (2020) IoT based biotelemetry for smart health care monitoring system. *J Infor Technol Digital World* 2(3):183–190
19. Vijayakumar T (2019) selective image enhancement and restoration for skin cancer identification. *J Innov Image Proc (JIIP)* 1(01):1–10
20. Sungheetha A, Rajesh SR (2020) Comparative study: statistical approach and deep learning method for automatic segmentation methods for lung CT image segmentation. *J Innov Image Process* 2:187–193
21. Koresh HJD, Chacko S, Periyanyagi M (2021) A modified capsule network algorithm for oct corneal image segmentation. *Pattern Recog Lett* 143:104–112
22. Alzheimer's Disease Neuroimaging Initiative (2004) ADNI Database. <http://adni.loni.usc.edu>
23. RIRE-Retrospective Image Registration Evaluation. <https://www.insight-journal.org/rire/>
24. Cancer Imaging Archive (2014). <http://www.cancerimagingarchive.net>
25. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010) N4ITK: improved N3 bias correction. *IEEE Trans Med Imag* 29(6):1310–1320
26. Nyul LG, Udupa JK, Xuan Z (2000) New variants of a method of MRI scale standardization. *IEEE Trans Med Imag* 19(2):143–150
27. Zhang X, Dong H, Gao D, Zhao X (2020) A comparative study for non-rigid image registration and rigid image registration. [arXiv:2001.03831](https://arxiv.org/abs/2001.03831)
28. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2019) Unet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans Med Imag* 39:1856–67

Simulation of Topology Based VANET Routing Protocols Using NS3



Mahabub Subhani Pathan and K. Annapurna

Abstract Nowadays more number of road accidents are happening due to increasing traffic. Intelligent Transportation system (ITS) proposes a system called Vehicular Ad hoc Network (VANET), a special class of mobile ad hoc network (MANET), which avoids traffic congestion problem and ensures safe road transportation. In VANETs vehicles are smart and they communicate with each other with the help of IEEE 802.11P protocol. VANETs have some unique characteristics like high dynamic topology, variable network density, frequently disconnected network, etc. Designing an effective routing protocol for such type of network is always a challenging task. This paper gives a brief introduction of the traditional routing protocols DSDV, OLSR, DSR and AODV and their performance is evaluated based on the parameters like MAC/PHY layer overhead and number of packets received. The simulations are performed using NS3 and SUMO.

Keywords VANET · MANET · ITS · DSDV · OLSR · DSR · AODV

1 Introduction

Because of the rapid growth in the usage of vehicles, many road accidents are occurring nowadays. This may be because of over speed, lack of coordination and control and inefficient traffic signaling system. It is possible to prevent these road accidents if one can able to alert the vehicle driver about a danger situation ahead of time. To make this possible, the vehicles should be made intelligent. VANETs, as a key part of Intelligent Transportation System (ITS), allow communication between on-road vehicles to share some useful information that can be used for reducing congestion, providing road safety, and betterment in traffic flow [1, 2].

VANETs, a type of infrastructure less networks, allow communication between vehicles without the need of a central coordinator. Because of their unique characteristics establishing a route between such fast moving vehicles is a challenging task to

M. S. Pathan (✉) · K. Annapurna
Vignan's Foundation for Science, Technology and Research (Deemed to Be University),
Vadlamudi, Guntur, Andhra Pradesh, India
e-mail: subhani15@gmail.com

address in VANETs [3]. Moreover, different routing approaches can be used based on the application scenario and network situation [4]. Hence designing of novel and efficient routing protocols is an important research area in VANETs [5]. In addition, it will be good if the routing protocols consider the priority of data in to account [6]. For real time data transmissions [7, 8], it will be more helpful if one is able to predict the traffic based on history [9].

The different applications of VANET include real-time traffic query, vehicle safety, parking spot searching, on-road advertising and to provide safety and convenience to drivers.

In this paper the traditional routing protocols, Destination Sequenced Distance Vector (DSDV) routing, Optimized Link State Routing (OLSR), Dynamic Source Routing (DSR) and Ad-hoc On- demand Distance Vector (AODV) routing, are compared and analyzed for MAC/PHY layer overhead and number of packets received. The analysis of these standard routing protocols can be useful to look at the pros and cons of the corresponding protocol which is further useful in the design of novel and more powerful algorithms for efficient routing in VANETs.

1.1 Key Research Areas and Challenges in VANETs

- (a) Mobility
- (b) Data administration and storage
- (c) Security and privacy
- (d) Quality of Service delivery
- (e) Heterogeneity and standardization
- (f) Routing protocols

The remainder of the paper is laid out as follows: Sect. 2 deals with the classification of VANET routing protocols, Sect. 3 gives brief description of DSDV, OLSR, DSR and AODV routing protocols, performance metrics are discussed in Sect. 4, simulation parameters are listed in Sect. 5, methodology is given in Sect. 6, simulation findings are reviewed in Sects. 7 and 8 concludes the paper.

2 Classification of VANET Routing Protocols

According to route formation strategy there are two broad classifications of VANET routing protocols. They are position based routing and topology based routing [10]. Figure 1 shows the classification.

The former is also called location based routing, geographic routing, or geo-routing. In this type, the position or location information of nodes is used to select the next best neighbor. Hence no need to create and retain route prior to data transmission.

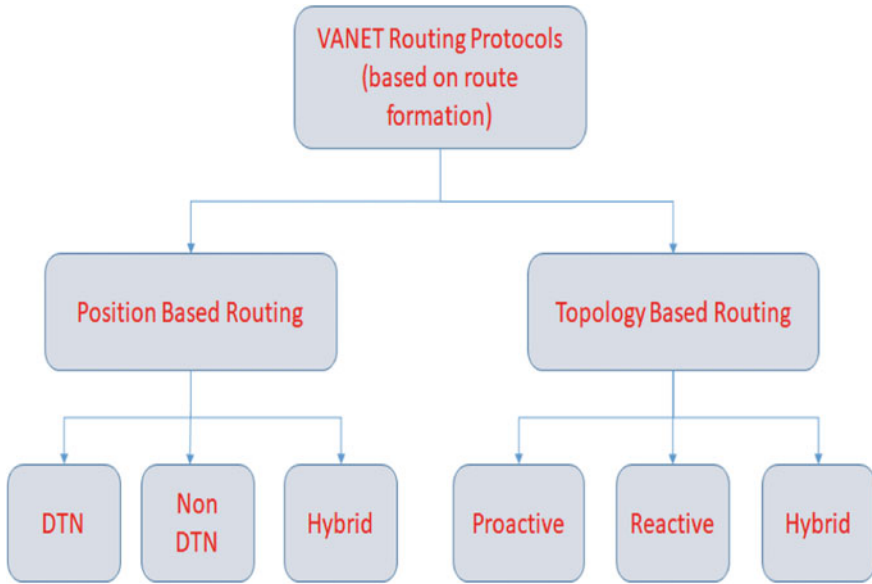


Fig. 1 Classification of VANET routing protocols

Delay Tolerant Network (DTN), non-DTN and hybrid types are the three subclasses of position based routing.

On the other hand, topology-based routing methods use the knowledge of network's existing links to forward packets. Routes are formed prior to data transfer with the help of control packets. As a result, every node of the network has the complete knowledge of network topology prior to data transfer. Proactive, Reactive and Hybrid are the subtypes of this routing technique.

In Proactive routing every node holds a table which contains the total network's topological information and all nodes update these table periodically. Whereas, in reactive or demand-driven routing, routes are created on demand i.e. when necessary. The protocols DSDV and OLSR are proactive whereas the DSR and AODV are of reactive type.

3 Description of Routing Protocols

3.1 Destination Sequenced Distance Vector (DSDV) Routing

In DSDV, each mobile host (MH) has a routing table that contains list of all possible destinations, next hop to each destination, shortest distance to each destination and a sequence number generated by destination MH for each entry [11]. Sequence numbers are used to distinguish between old and new routes, thereby avoiding the

routing loop problem. The routes with latest sequence number are always preferred. In DSDV, routing tables are updated periodically. And there exists another routing table that contains the settling time which is a reference for the node to advertise update information.

Because of the mobility of nodes in the network, links between them may be broken in some situations. In that case, all the routes through that node are allocated infinity metric and a new sequence number is broadcasted throughout the network. This mechanism gives assurance of loop-free routes; however, it introduces route fluctuation when multiple nodes transmit their update packets at irregular time intervals. This problem can be solved if every node waits for a predefined time (average settling time) available from a route-settling-time-table before broadcasting any route update message.

3.2 Optimized Link State Routing (OLSR)

OLSR is the modification of traditional Link State Routing (LSR) algorithm. In LSR each MH advertises its link state data to all the neighbors connected to it, which creates a lot of redundancy [12]. In contrast, the OLSR avoids these unnecessary transmissions by selecting particular nodes to retransmit the link state information. Each node of the network decides which of its neighbors can flood the link state packet. These selected neighbor nodes are known as Multi Point Relay (MPR) nodes [13]. Hence, rather than blind flooding, the packets are now retransmitted only by MPR nodes. Periodic updation of topological information is done in OLSR.

In OLSR, hello messages are used to get the link state data and also they contain the addresses of all connected neighbors of a node. Similarly, Topology Control (TC) messages contain broadcasting data. TC messages provide the information regarding the topology of the network that is stored by each node in its topological table. TC messages also include the information of MPR nodes. A routing table is generated based on this information.

3.3 Dynamic Source Routing (DSR)

A DSR network can organize itself independently. The DSR algorithm's routing process is divided into two parts: route discovery and route maintenance [14].

Route Discovery:

Whenever a MH (Source) does not find any route to a desired MH (Destination) from its routing table then it enters a route discovery phase. A route request (RREQ) packet, containing source and destination addresses and a unique ID to represent the request, is broadcasted by the source MH. The RREQ packet received with a new ID will only be forwarded by the neighbors in order to reduce redundancy. In case, the

destination node or an intermediate node with a route to the destination receives this RREQ packet then they respond with a route reply (RREP) packet. The RREP packet includes the complete path from source MH to destination MH obtained from the RREQ packet. Alternate paths are chosen in case the selected path is disconnected.

Route Maintenance:

Route maintenance phase involves the transmission of route error (RERR) packets and acknowledgements. When there is a transmission error in the data link layer, RERR packets are generated by the nodes that detects the link failure. After receiving a RERR packet, all the nodes remove the hop which is in error from their route caches. Acknowledgements ensure the proper functioning of route links.

3.4 Ad-Hoc On-Demand Distance Vector (AODV) Routing

AODV is said to be pure on-demand routing protocol in which routes are established when necessary and MHs which are not a part of the selected route do not hold the routing information or exchange routing tables [15]. Its routing method is hop-to-hop.

AODV routing follows the same procedure of route discovery and route maintenance similar to DSR. Additionally, it borrows the concept of sequence numbers that is used in DSDV routing. Sequence numbers are very much useful in identifying stale routes from new routes, thereby avoiding the routing loop problem. That is why AODV routing is also called a combination of DSDV and DSR routing techniques. In case of link failure either due to the mobility of source node itself or mobility of a MH on the active path, the source node may reinitiate the route discovery process if it still wants to communicate with that destination. In AODV, routing overhead is reduced as the number of broadcasts are minimized. It supports unicast routing only and no multipath routing.

4 Performance Metrics

Now, let us evaluate the performance of DSDV, OLSR, DSR and AODV routing methods with respect to MAC/PHY overhead and number of packets received.

4.1 MAC/PHY Layer Overhead

Generally, the useful information of a VANET is shared in the form of Basic Safety Messages (BSM). And the routing packets generated by routing protocols carry the updated routing information. As the routing packets do not carry any useful

information about the application so they create an overhead on network bandwidth which may in turn reduces the system performance. Hence, we call this overhead on the network bandwidth as MAC/PHY overhead [16]. The MAC/PHY layer overhead can be calculated as

$$\text{MAC/PHY layer overhead} = \frac{(\text{Total PHY bytes} - \text{Total App bytes})}{\text{Total PHY bytes}}$$

4.2 Number of Packets Received

This parameter gives the total number of packets received by the destination successfully.

5 Simulation Parameters

The simulation parameters are shown in Table 1.

6 Methodology

Step 1. Open the OSM web wizard interface and select the desired area for simulation.

Step 2. Select the number of cars, buses, trucks and simulation time.

Step 3. Generate the traffic scenario of selected area with the help of SUMO GUI. This will generate ‘osm.sumocfg’ file.

Step 4. Convert the osm.sumocfg file into ‘trace.xml’ file.

Table 1 Simulation parameters

| Parameter | Value(s) |
|-------------------|--------------------------|
| Simulators | NS-3.27 and SUMO |
| MAC protocol | IEEE 802.11P |
| Vehicular nodes | 39 and 104 |
| Node speed | 30 m/s |
| Mobility model | Random Waypoint |
| Packet size | 64 bytes |
| Simulation time | 16 and 20 s |
| Transmit power | 20 dBm |
| Routing protocols | DSDV, OLSR, DSR and AODV |

Step 5. Then Convert this XML file into TCL (Tool Command Language) file (mobility.tcl).

Step 6. Attach this ‘mobility.tcl’ file to the “vanet routing compare.cc” file of NS3.

Step 7. Run the code for different protocols (DSDV, OLSR, DSR, and AODV) which generates different files (.log file,.mobfile,.csv file, etc.).

Step 8. use the ‘.csv’ file of the corresponding protocol to generate graphs for MAC/PHY overhead and packets received with respect to simulation time with the help of GNU PLOT.

7 Simulation Results and Analysis

The simulations are performed using NS-3.27 and results are generated for two different traffic densities consisting of 39 and 104 nodes respectively.

7.1 MAC/PHY Layer Overhead

The Figs. 2 and 3 represent the MAC/PHY layer overhead of the discussed four routing protocols with respect to simulation time for 39 and 104 nodes respectively. It is clear that the routing protocol DSR is having least MAC/PHY layer overhead when compared to DSDV, OLSR and AODV in both the cases. Whereas DSDV is having lesser overhead than AODV and OLSR in low traffic density region and highest overhead approximately same as AODV in high traffic density region.

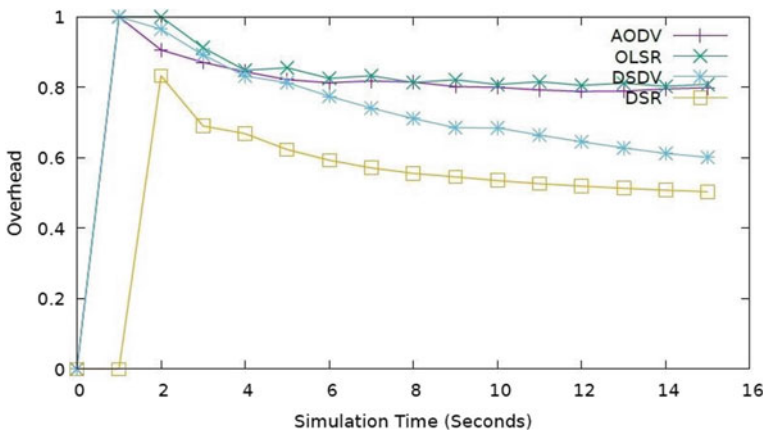


Fig. 2 Simulation time versus MAC/PHY layer overhead (39 nodes)

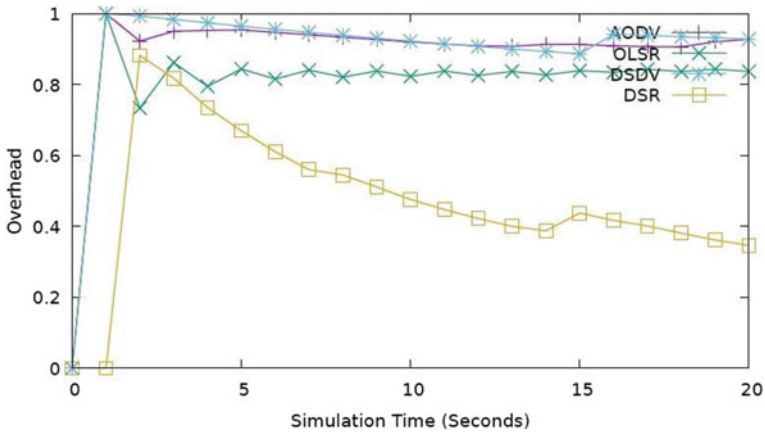


Fig. 3 Simulation time versus MAC/PHY layer overhead (104 nodes)

7.2 Number of Packets Received

From Figs. 4 and 5 it is clear that the routing protocols DSR and OLSR are receiving constant (over time) and highest number of packets when compared to the remaining two, DSDV is receiving constant (over time) and least number of packets and AODV is fluctuating between low and high values in case1. Whereas in contrast to case-1, DSR is receiving least number of packets than others in case-2, while the remaining three protocols are approximately receiving the same high number of packets.

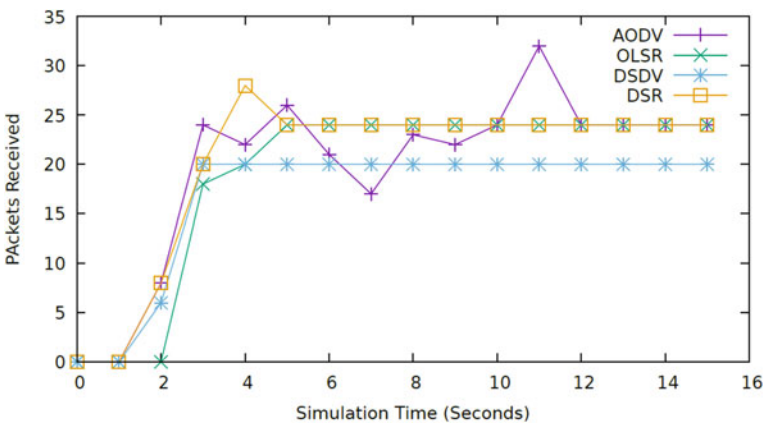


Fig. 4 Simulation time versus packets received (39 nodes)

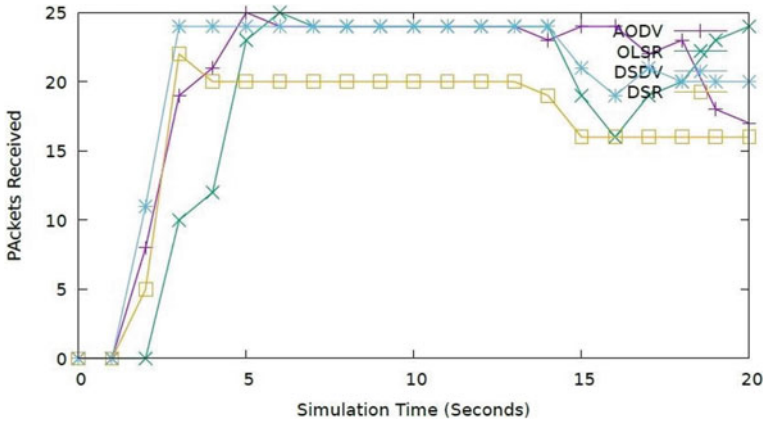


Fig. 5 Simulation time versus packets received (104 nodes)

8 Conclusion

The above results indicate that the DSR routing protocol is having least MAC/PHY overhead for both low and high density network of nodes and highest packets receive rate for low density of vehicles. However, it exhibits poor receive rate for a high density network of nodes. Hence one can conclude that DSR routing protocol is best suited for low density of nodes.

Similarly, OLSR performs better in high density network of nodes since it has highest packets receive rate with moderate MAC/PHY layer overhead.

AODV and DSDV are also suitable for high density networks as they are having highest packets receive rate, but one should compromise for MAC/PHY layer overhead.

Even though a wide variety of routing protocols have been developed for VANETs still there is a requirement to design novel and efficient routing protocols that will give better performance in dynamic environment conditions like VANET.

References

1. Qureshi KN, Abdullah AH (2013) Topology based routing protocols for vanet and their comparison with manet. *J Theor Appl Inf Technol* 58(3). ISSN: 1992–8645, E-ISSN: 1817–3195
2. Salunke K, Ragavendran U (2018) An overview of privacy preserving schemes for VANET. In: *International conference on intelligent data communication technologies and internet of things*. Springer, Cham, pp 1103–1112
3. Abbas F, Fan P, Khan Z (2018) A novel reliable low-latency multipath routing scheme for vehicular ad hoc networks. *EURASIP J Wirel Comm Netw*. <https://doi.org/10.1186/s13638-018-1292-1>

4. Pandey PV, Swaroop A, Kansal V (2019) A concise survey on recent routing protocols for vehicular ad hoc networks (VANETs). In: International conference on computing, communication, and intelligent systems (ICCCIS), ISBN: 978-1-7281-4826-7/19
5. Chen JIZ, Smys S (2020) Optimized dynamic routing in multimedia vehicular networks. *J Inf Technol* 2(03):174–182
6. Kumar KS, Annapurna K, Ramanjaneyulu BS (2015) Supporting real-time traffic in cognitive radio networks. In: International conference on signal processing and communication engineering systems, pp 482–485. <https://doi.org/10.1109/SPACES.2015.7058201>
7. Annapurna K, Ramanjaneyulu BS (2021) Timeout-aware inter-queuing for QoS provisioning of real-time secondary users in cognitive radio networks. In: Micro-electronics, electromagnetics and telecommunications lecture notes in electrical engineering, vol 655. Springer, Singapore, pp 693–700
8. Ramanjaneyulu BS, Annapurna K (2021) Supporting real-time data transmissions in cognitive radio networks using queue shifting mechanism. *Int J Embedded Real-Time Commun Syst* (2021)
9. Annapurna K, Ramanjaneyulu BS, Chaitanya CL, Hymavathi T (2017) Spectrum prediction in cognitive radio networks using neural networks. *Int J Control Theory Appl* 10(28):143–148
10. Nair CR (2016) Analysis and comparative study of topology and position based routing protocols in VANET. *Int J Eng Res General Sci* 4(1). ISSN: 2091–2730
11. Ghawy MZ, AL-Sanabani MA (2017) Application and performance analysis of DSDV routing protocol in Ad-Hoc wireless sensor network with help of NS2 knowledge. *Global J Comput Sci Technol E Netw Web Secur* 17(1). ISSN: 0975–4172
12. Jacquet P, Muhlethaler P, Clausen T et al (2003) Link state routing in wireless ad-hoc networks. In: MILCOM'03 proceedings of the IEEE conference on military communications. Washington, DC
13. Clausen T, Jacquet P (2013) Optimized link state routing protocol (OLSR). In: RFC-3626, IETF Networking Group
14. Johnson DB, Maltz DA (2001) Dynamic source routing in ad hoc wireless networks. In: Mobile computing. Springer, U.S
15. Maurya PK, Sharma G, Sahu V, Roberts A, Srivastava M (2012) An overview of AODV routing protocol. *Int J Modern Eng Res (IJMER)* 2(3):728–732. ISSN: 2249–6645
16. Afzal K, Tariq R, Aadil F, Iqbal Z, Ali N, Sajid M (2021) An optimized and efficient routing protocol application for IoV. *Hindawi Math Prob Eng* 2021(9977252):32. <https://doi.org/10.1155/2021/9977252>

A Fast Algorithm for Hunting State-Backed Twitter Trolls



Shaaban Sahmoud , Abdelrahman Abdellatif , and Youssef Ragheb

Abstract In recent years, state-backed troll accounts have been adopted extensively by many political parties, organizations, and governments to negatively influence political systems, persecute perceived opponents, and exacerbate divisiveness within societies. Thus, the need for an automatic state-backed troll classification system has increased. Various algorithms have been proposed in the literature to handle this problem, but a majority of them consider all types of trolls as one type which decreases the performance of classification algorithms. Our goal in this paper is to design a thorough method for detecting state-backed trolls on Twitter with the ability to work efficiently in any case regardless of the language, the location, and the purpose of the troll account. For accurate classification, a set of novel effective and powerful features from various categories are proposed. To train our algorithm, we gathered a large and relevant dataset from Twitter. The results show that the proposed algorithm achieves high classification accuracy (approximately 99%) and has the ability to classify state-backed troll accounts regardless of the language or the location of the account.

Keywords Troll detection · State-backed trolls · Online antisocial behaviors · Social media · Twitter troll accounts

1 Introduction

Twitter is one of the most widespread and popular social media platforms in the world. It employs short text messages to allow people to share their opinions and feelings publicly with their followers and followings [1]. Twitter's popularity has shown up dramatically over the past few years after its extensive usage in disseminating political opinions in political upheaval such as the so-called Arab Spring (2011–2013) [2], American elections (2016 and 2020) [3], Iranian protests (2019–2020), Iraq Protests (2019–2020) [4], and the Arab Gulf crisis against Qatar (2017–2020) [5]. Indeed, there are various reasons why people prefer Twitter over other social media networks

S. Sahmoud (✉) · A. Abdellatif · Y. Ragheb
Computer Engineering Department, Fatih Sultan Mehmet Vakif University, Istanbul, Turkey
e-mail: ssahmoud@fsm.edu.tr

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
G. Ranganathan et al. (eds.), *Pervasive Computing and Social Networking*, Lecture Notes in Networks and Systems 475, https://doi.org/10.1007/978-981-19-2840-6_49

643

during civil unrest or political problems (1) Twitter is characterized by many specific capabilities (Hashtags, retweets, quote tweets, ...) that makes it an extremely exciting tool for politicians and decision-makers [6] (2) most of the Twitter users keep their posts public, which means they allow full access to their tweets to all (3) there are no restrictions for the friends or followers and you can follow anyone's tweet (Facebook allowed a maximum of 5,000 friends.). Consequently, many politicians are often found more active on Twitter than other social media tools such as Facebook and Instagram. Unfortunately, the tremendous influence of Twitter on the world's political community has revealed a new threat against freedom of expression and democracy if it is misused as a tool of social control and influencing public opinion [7]. Usually, this abnormal behavior is regulated and controlled on Twitter using large numbers of accounts called Trolls [8, 10]. Unfortunately, after the emergence of the tremendous impact that can troll accounts make, many countries allegedly began to use and adopt these accounts and strategies to serve their specific agenda instead of preventing it, which contributed to the emergence of a new type called state-backed trolls or state-sponsored trolls [12, 15]. Another possible term that has been commonly used when discussing the private networks of state-backed trolls is Troll Farms. It refers to a considerable number of anonymous accounts or trolls linked to a certain government, political party, or organization and work to implement a specific political agenda [16–18]. Usually, these accounts work in groups and with pre-coordinated media plans and campaigns to reach the trending in a specific region or generate propaganda on a certain topic. Recently, it is attentively observed that the state-backed trolls have been implemented excessively in many political crises especially prior to democratic elections as occurred in the USA, Turkey, Russia, and Iraq [4, 13, 14, 19].

This paper presents an empirical study of using various sets of features and classifiers to accurately identify state-backed Twitter trolls. Various combinations from several categories of features were evaluated and compared to discover the most robust and relevant set of features. During the design and implementation of our algorithm, a set of points were taken into account to overcome the problems and limitations of other methods as will be discussed in Sect. 2. The originality of this paper is summarized in the following points:

- We have achieved a high rate of classification accuracy for the task of detecting trolls on state-supported Twitter (a success rate of more than 98% in detecting trolls accounts).
- A large dataset is gathered from diverse countries to train the classifiers in our detection system. To the best of our knowledge, this is the biggest dataset that gathered for the state-backed troll's detection task (we gathered more than 143 million tweets from around 37 thousand Twitter users).
- We propose a set of new effective and powerful features that can classify state-backed Twitter trolls with high precision.
- The analyzing time is reduced by replacing the features which usually require sentiment analysis techniques and take higher time to be calculated by other faster features.

The paper structure is organized as follows. Section 2 presents a brief review of trolls' accounts detection techniques from Twitter. Our proposed state-backed trolls' detection algorithm and its steps are presented in Sect. 3. The data-collecting steps are described in Sect. 4. In Sect. 5, we present the results of our experimental analyses to assess the performance of the proposed algorithm and analyze the impact of various categories of features. Finally, Sect. 6 concludes this paper.

2 Related Works

Since the topic of state-backed trolls or state-sponsored trolls is new, there are a few research papers that have been published to discuss this issue. Therefore, we will summarize the leading researches dedicated to this and other related topics such as detecting other types of trolls, detecting abusive behavior on Twitter, and detecting propaganda on Twitter. It is significant to mention that this short review is focused on Twitter only since it is the most used social media in political and public opinion issues.

A recent research work on detecting state-backed trolls is the work performed by Alhazbi [15] where he proposed a system that constitutes from a machine learning model and uses a set of behavioral features for Twitter users' activities. The used classifier is trained and tested using a set of trolls' accounts from Saudi Arabia that closed and published by Twitter in 2019 for research purposes. According to Alhazbi [15], the classification accuracy of this system was around 94%. In Sahmoud and Safi [4], another work has been done to analyze the propaganda that occurred during the Iraqi people protest in 2019. The authors studied the activities of Twitter users and identify if any suspicious influence exists. The results inferred that there was a set of external groups of trolls that try to amplify the propaganda and affect the political situation.

One of the most interesting and comprehensive frameworks to detect Twitter trolls is the TrollPacifier system [9]. It is a holistic system that utilizes a large number and different features to detect trolls. In this system, six different groups of features (style of writing, sentiment analysis, behaviors features, user interactions, linking information, and publication time features) are utilized to train a neural network to be able to classify all types of digital trolls. According to the presented results in the paper, the system achieved high classification accuracy (95.5%) and demonstrated excellent reliability in managing online anti-social behaviors.

The 2016 US Presidential Election is one of the most studied and investigated cases in the domain of state-backed trolls on Twitter [7, 20]. For example, in Badawy et al [21], the authors proposed a method to characterize the 2016 Russian trolls who tried to influence the US election. They used a label propagation method to identify Twitter users' ideology by focusing on the sources of the news they share. The proposed mechanism was able to classify around 84% of the troll accounts that participated in the campaigns. In another paper [22], the authors developed an adaptive framework to analyze Twitter user identity using a social sequence analysis method presented

in Abbott and Tsay [23]. Accordingly, the trolls' detection task transforms into a data clustering problem. As a case study, the data of Russian trolls that attempted to affect the general opinion in the 2016 US Presidential Election were employed to evaluate the performance of this method. The main idea of this paper is to design a new time-sensitive semantic metric called t-SED by mapping the text distances of tweets. For more investigation on USA Presidential Election and to check if a similar influence on elections can be detected in Europe, Kellner et al. [24] inspected the propaganda on Twitter during the German federal election in 2017. They found that a group of 79 trolls from the US election campaign have also tried to influence the German federal election. Furthermore, a research paper [25] accuses Russian trolls of amplifying the vaccine debate and uses it as weaponized health communication. Analyzing the texts of the tweets is not the only technique used to detect Russian trolls, Zannettou et al. [26] tried to characterize the use of images by state-backed troll accounts during the US election on Twitter. The main objective of the paper is to investigate how trolls employ imagery to achieve their goals and influence public opinion.

3 Proposed Approach for Troll Detection

By analyzing the existing techniques for detecting state-backed troll accounts in the previous section, we concluded that there are several issues that still exist in most algorithms. We can summarize these problems in the following points:

- While there are different types of digital trolls created for different objectives, most of the existing techniques try to handle all of them as one type which affects the performance of classifiers and degrades the performance of the detection framework. Therefore, our objective in this paper is to handle this problem independently and more specifically to enhance the detection accuracy.
- Most of the algorithms that deal with this problem use features from one category or a small number of features [4, 15], which affects the efficiency of the algorithm and restricts its detection ability. As shown in Fig. 1, our algorithm utilizes a large set of features from four categories to accurately hunt the state-backed Twitter accounts.
- After inspecting the datasets used in existing research papers to train the classifiers, it is found that those datasets suffer from many issues such as the small number of considered accounts [22], troll accounts are collected from only one region or country [4, 23, 25], and the manual classification of troll accounts which may cause many false classification samples [9, 27]. In this paper, we overcome these issues by gathering a large dataset from different regions and using the classification done by Twitter for the state-backed troll accounts (it is published for research work on Twitter) as will be described later in this paper.

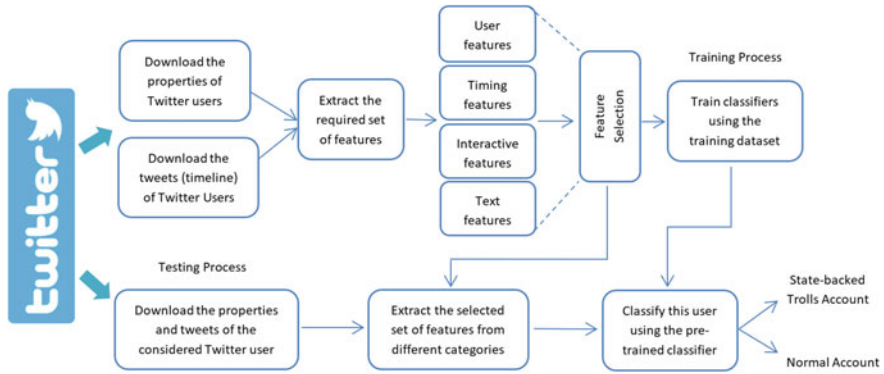


Fig. 1 The steps of the training and testing processes in our proposed state-backed troll detection

Table 1 The best 10 generated features in our system according to SFBS feature selection algorithm

| | Feature | Feature category |
|----|---|--------------------------|
| 1 | Average number of hashtags in retweets | Text contents |
| 2 | Mention rate in retweets | Interactions with others |
| 3 | Mention rate in replies | Interactions with others |
| 4 | Average number of emoji's in retweet | Text contents |
| 5 | Followers count | General behavior |
| 6 | Average time between tweets for a week | Time statistics |
| 7 | Average weekly like rate for retweets | Time statistics |
| 8 | Maximum number of tweets in one hour | General behavior |
| 9 | Account Age | General behavior |
| 10 | Average number of tweeting days in a week | Time statistics |

3.1 General Methodology

Our algorithm aims to detect the state-backed troll accounts on Twitter based on their tweets and account information. As shown in Fig. 1, we have two essential processes in our detection framework, the training process, and the testing process. In the training process, we used the large collected dataset for both troll and non-troll Twitter accounts to generate a set of features from four categories called user, timing, interactive, and text style features (more description is provided in Sect. 3.2). After that, a feature selection algorithm is employed to select the most relevant set of features that

can be used to classify the state-backed troll accounts. In the feature selection step, the Sequential Floating Backward Selection (SFBS) algorithm was utilized to select the most appropriate set of features for the classification task [28]. The SFBS is a commonly used feature selection technique that is considered as an extension to the simpler SBS algorithm. It starts from the full set of features and after each backward step; it performs forward steps as long as the objective function increases. Among the five tested classifiers to predict the labels of Twitter accounts, the Random Forest algorithm (RF) was employed. To apply the proposed method practically, firstly, the RF classifier is build and prepared using our training set. After that, the data of the considered Twitter account is downloaded; experimentally it is preferred to be more than 100 tweets. Finally, the RF classifier is applied to classify the considered Twitter account if it is a normal account or it belongs to a group of Twitter trolls.

3.2 *Classification Features*

Generating and selecting the most relevant feature set is a highly significant step in classification scenarios. As we have discussed in Sect. 2, researchers have used a wide variety of feature sets to detect trolls accounts. Accordingly, we have chosen a set of features from previously proposed algorithms and a new set of features from variant approaches is introduced in this paper. Before start describing our features, we clarify the difference between post types in Twitter to simplify understanding the generated features.

Tweet: A message posted to Twitter containing text, photos, GIF, and/or video.

Retweet: A re-posting of a tweet. Twitter's Retweet feature helps people quickly share that tweet with their followers.

Reply: When a Twitter user responds to another user's tweet. **Quote tweet:** It is a retweet that has been made with a comment. It is different from a regular retweet because it starts a new thread that people can retweet and like separate from the original tweet.

In particular, for training our state-backed trolls, we have identified four categories of features, as listed in the following paragraphs. Table 1 shows a sample of 10 selected features from our overall set of features.

User Features and General Behavior: This category constitutes of two types of features (1) general Twitter account features (2) features to describe the behavior of the user. Features in the first type are obtained directly from the account properties downloaded by Twitter API such as number of followers, number of followings, account age, and the description of the account. The second type of features is specially generated by carefully examining the behavior of state-backed troll accounts. Examples of these features include the highest number of tweets posted by the considered user in one hour, number of active days (days which a user posted at least one post) in a month, number of active days in a week, and the day of the week with

the highest number of tweets. According to our results, these features show a high ability to classify state-backed trolls as will be shown later.

Time Statistics: The timing features of different activities for Twitter users are widely used in many classification and clustering tasks [9]. In this category of features, we derived a group of features to correctly describe the timing habits of the considered Twitter account for all Twitter activities such as tweets, replies, retweets, and quote tweets. Examples of generated features in this category include daily tweets rate, tweets rate monthly, the average time between tweets, time of the first tweet in a day, time of the last tweet in a day, the average number of tweets in first 8 h of the day, and an average number of tweets in last 8 h of the day.

Interactions with Others: This category includes a set of features to describe integration between the considered Twitter user and other users in the Twitter community. Since most of the state-backed accounts try to post a higher number of posts without a real community that may react with these posts, it is expected to have fewer interactions than normal accounts. Mainly, we focused on four major activities that Twitter users frequently use to express their feeling regarding a tweet or user (1) like a tweet (2) retweet a tweet (3) mention a user (4) reply to a tweet. By mixing these four activities, we generated 28 new features such as average user mentions in tweets, average user mentions in retweets, average user mentions in replies, average user mentions in quote tweets, the average number of likes for tweets, the average number of likes for retweets, and the average number of likes for quote tweets. Moreover, we generated different versions of these features by merging them with the time dimension. As an example, for the average number of like rates in tweets, we produced three versions based on daily, weekly, and monthly rates.

Tweet Text Contents: This category includes a set of 36 features to describe the contents and style of the tweets, replies, and quote tweets of the considered user. The reason for including this category of features is that the state-backed accounts like other troll accounts usually write short tweets and follow a certain style in posts that can be detected and utilized to distinguish these accounts. On Twitter, we can find special content such as hashtags and web addresses (URLs), and we can find general content such as words, punctuation marks, and special characters. From these two types of contents, we generated a set of features that can detect the language kind of tweets. As we did in previous features' categories, we also merge the time dimension in the features of this category to discover all possible relevant features. Furthermore, we differentiated between a tweet, reply, and quote tweet when deriving this set of features. Examples of features of this category include URL average in tweets, URL average in replies, hashtags average in tweets, hashtags average in replies, average word count in tweets, average word count in replies, average word length in tweets, average word length in replies, average daily hashtags, hashtags rate in replies, hashtags rate in tweets, punctuation characters average in tweets, punctuation characters average in replies, emotion characters in tweets and emotion characters in replies. To better highlight the effectiveness of the generated and proposed set of features, Fig. 2 gives a scatter plot for the values of six sample features after evaluating

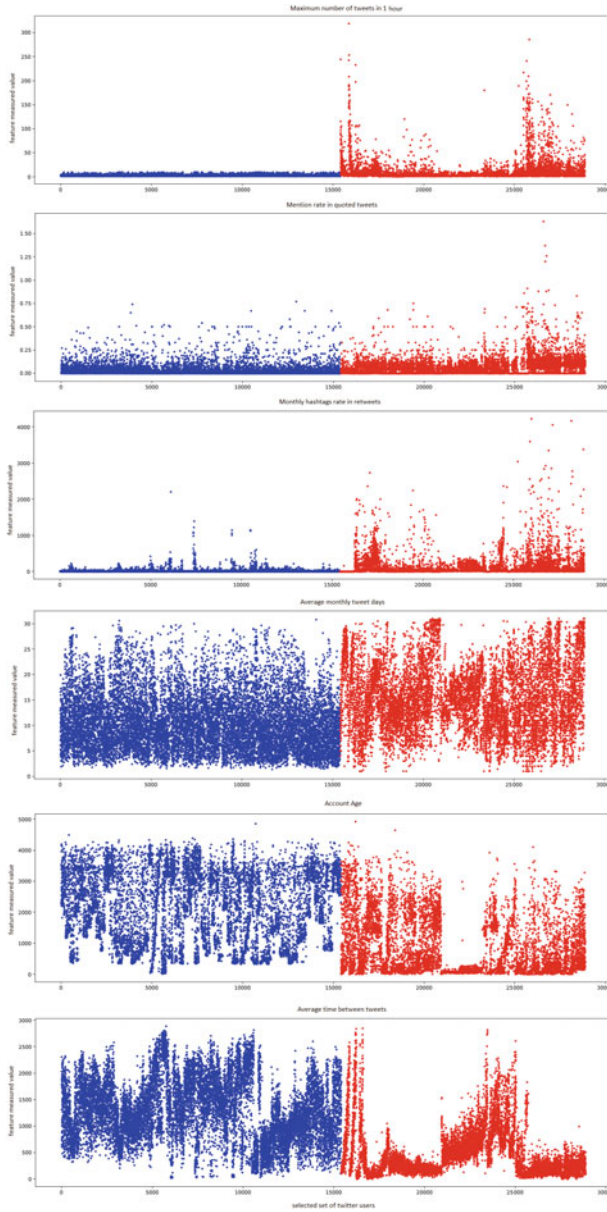


Fig. 2 Scatter plot for the calculated values of 6 sample features from our selected set of features. The blue dots represent the non-troll accounts where red dots represent the feature values of troll accounts

them using our gathered set of Twitter accounts. In Fig. 2, red points represent the troll accounts and blue points represent the normal or non-troll accounts. This figure demonstrates the ability of generated features in discriminating and characterizing the state-backed troll accounts. The first feature given in the first row represents the maximum number of tweets posted in one hour from the considered Twitter account. Although this feature represents a newly generated feature and to the best of our knowledge it was not used in any other troll detection algorithm before, it can significantly detect and identify the troll accounts as shown in Fig. 2. The third-row feature called “monthly hashtag rate in retweets” gives identical encouraging results. The last three rows represent some of the other good features used in our algorithm where the values of troll and non-troll accounts are close but there is still a difference between in behaviors of these two classes. The second row shows a sample feature where the values of trolls and non-trolls accounts have nearly similar behavior. Therefore, we expect that the feature selection step is crucially important to exclude the irrelevant and redundant features as possible. Our algorithm selects only 50 features from 110 generated features to identify the troll accounts.

4 Data Gathering

Data acquisition remains a crucial step in analyzing the behavior of troll users through social media networks. One of the most critical challenges facing data collection is the applied privacy laws regarding the use and share of human data in public social networks. Consequently, it becomes extremely rare to find a public dataset for Twitter users. This means that the only method to gather the targeted dataset is to download the data directly from Twitter using Twitter API functions. Fortunately, for transparency and to enhance the research on public inauthentic campaigns, Twitter started to make state-backed information archives publicly available. Until December 2020, Twitter disclosed around 83,673 accounts linked to state-backed information operation and put their data publicly including user’s information, tweets, and media. In our paper, we have two datasets, the first one for trolls accounts and the second for normal accounts. To create the first dataset, we gathered seven different sets from the public Twitter archives to be used as trolls. Table 2 shows the number of Twitter users, the number of tweets, and the release date for each dataset. After analyzing the datasets, we noted that many accounts do not have any tweets in their datasets. Therefore, we included another column in Table 2 to present the number of Twitter users that have tweeted. By excluding those zero-tweets users, we gathered 119,310,294 tweets from 23,738 Twitter users in our dataset. All these accounts are labeled as trolls in both training and testing processes. We selected different datasets from various countries which include the highest number of Twitter accounts that classified as state-backed trolls by Twitter operation information.

The second dataset is created by manually gathering the data of a large number of normal Twitter accounts. To accomplish this task, we employed two data downloading tools which are Tweepy and KNIME. Tweepy [29] is an open-source Python

Table 2 Description of the troll accounts dataset used in our experiments

| Countries | Number of users | Number of tweets | Release date |
|---------------|-----------------|------------------|--------------|
| KSA-Egypt-UAE | 5350 | 36,523,980 | Feb. 2020 |
| Egypt | 2541 | 7,935,329 | Feb. 2020 |
| UAE | 4248 | 1,325,530 | Mar. 2019 |
| UAE-Egypt | 271 | 214,898 | Apr. 2019 |
| Saudi Arabia | 5929 | 3,2054,257 | Oct. 2019 |
| Turkey | 7340 | 36,899,376 | May 2020 |
| Russia | 1382 | 4,356,924 | 2019–2020 |

Table 3 Description of the non-troll accounts dataset used in our experiments

| Country | Number of users | Number of tweets | Downloading date | Data size (GB) |
|--------------|-----------------|------------------|------------------|----------------|
| Saudi Arabia | 1296 | 3,494,800 | Oct. 2020 | 1.1 |
| Egypt | 2925 | 5,593,232 | Oct. 2020 | 1.5 |
| UAE | 3424 | 2,744,508 | Oct. 2020 | 1.3 |
| Turkey | 5700 | 11,297,261 | May 2020 | 4.61 |
| Russia | 491 | 1,368,174 | Oct. 2020 | 0.67 |

package that gives researchers a very simple and fast way to access the Twitter API with Python. KNIME [30] is an open-source software to create and produce data science projects using one easy and intuitive environment, and it includes a special package with different drag-and-drop components dedicated to download and analyze Twitter Data. Table 3 shows the details of the gathered data from Twitter. A total of 24,497,975 tweets from 13,836 Twitter users were collected from the same five considered countries of the first dataset. In order to obtain a more accurate dataset, we downloaded these data by searching on some general hashtags that are not related to any political attitudes or parties. Additionally, we applied different filtering steps to ensure that these accounts belong to normal persons and not trolls such as:

- Do not download any Twitter account with a high number of followers
- Do not download any Twitter account with a high number of friends
- Do not download any Twitter account with a high number of tweets relative to its age
- Do not download any Twitter account with a number of tweets less than 100

5 Results and Discussion

In this section, we present the experimental results of our proposed algorithm for detecting state-backed Twitter trolls and we show the importance of different feature categories in boosting the detection process. We first describe the performance

metrics used to evaluate the performance of considered algorithms and the used classifiers in our comparison framework. After that, the results and its discussion are given to assess the performance of our algorithm. In our experimental study, six machine learning algorithms (classifiers) are utilized to select the best classifier for our proposed algorithm: Logistic Regression (LR), Decision Tree (DT), Neural Networks (NN), Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). All algorithms are implemented and executed using sklearn library in Python. To assess the performance of compared algorithms, three metrics were used: (a) The classification accuracy: it is the most common performance measure and it is simply computing the average classification accuracy as a ratio of correctly predicted observation to the total observations. (b) F1 score: it is the weighted average of Precision and Recall, and it takes both false positives and false negatives into account. (c) Recall measure: it is the ratio of correctly predicted positive observations to all observations in the actual class.

The experimental results of this section demonstrate the effectiveness of the proposed and selected set of features for the classification of Twitter troll accounts. Table 4 shows the results of comparing the six considered classification algorithms by selecting a different number of features in every run. For all experiments, the 10-folds cross-validation was performed. The classification algorithms are compared using four different sets of features 30, 40, 50, and all features. The three sets of features (30, 40, 50) are obtained by utilizing the SFBS feature selection algorithm as explained in Sect. 3.1. In general, the results of Table 4 prove how our selected set of features can precisely discriminate the state-backed troll Twitter accounts where all classification algorithms achieve accuracy larger than 80%. From another direction, some tested algorithms such as RF and NN work much better than others such as SVM and KNN. For all experiment instances, the RF algorithm fulfills the best results with an accuracy of approximately 99%. One reason for this is the additional randomness that exists in the RF model during its tree construction. RF algorithm searches for the best feature among a random subset of features rather than searching for the most important feature while splitting a node. This aspect generally results in better classification accuracy. Moreover, it is noted that selecting

Table 4 Accuracy, F-measure, and recall for different number of features, using different classification algorithms

| Metric | Accuracy (%) | | | | F1-Measure | | | | Recall | | | |
|---------------|--------------|------|------|------|------------|------|------|------|--------|------|------|------|
| | 30 | 40 | 50 | All | 30 | 40 | 50 | All | 30 | 40 | 50 | All |
| # of Features | | | | | | | | | | | | |
| LR | 93.8 | 93.7 | 94.2 | 96.7 | 0.91 | 0.90 | 0.89 | 0.92 | 0.90 | 0.90 | 0.89 | 0.92 |
| KNN | 94.5 | 94.5 | 94.9 | 96.5 | 0.90 | 0.90 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.89 |
| CART | 99.5 | 99.3 | 99.5 | 99.5 | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 | 0.98 | 0.98 | 0.98 |
| SVM | 87.4 | 87.4 | 88.1 | 90.2 | 0.81 | 0.81 | 0.82 | 0.84 | 0.80 | 0.80 | 0.82 | 0.84 |
| RF | 99.7 | 99.8 | 99.8 | 99.8 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| NN | 98.6 | 98.6 | 97.1 | 99.1 | 0.96 | 0.94 | 0.96 | 0.96 | 0.96 | 0.95 | 0.96 | 0.96 |

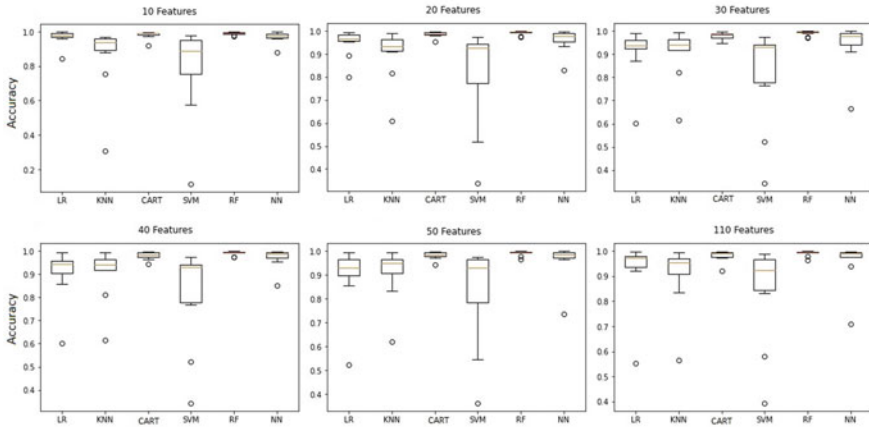


Fig. 3 Boxplot of the cross-validation process for 6 different sets of features

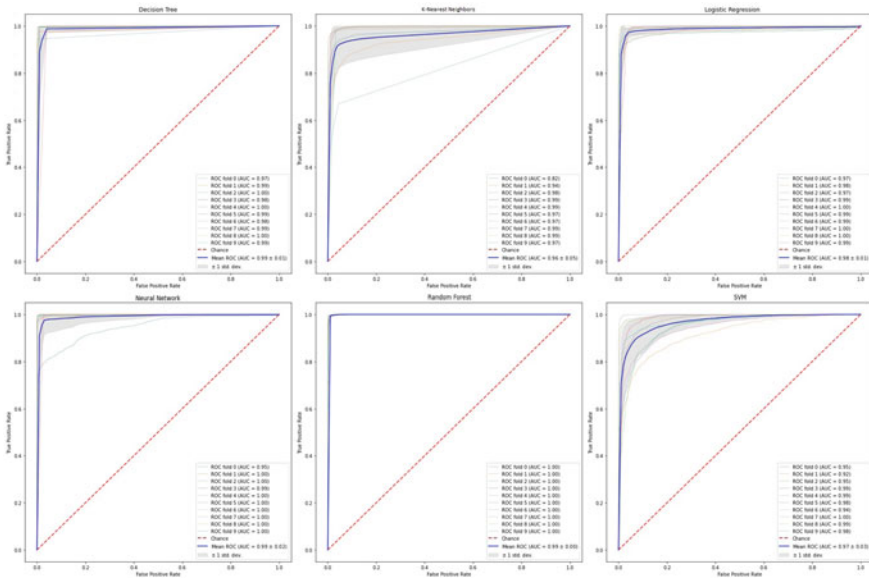


Fig. 4 ROC curves of the cross-validation process for 6 different sets of features

the best 40 and 50 features produces the best classification result. This means that some of our features are irrelevant or provide a similar contribution as other features and can be excluded from the set of selected features. Therefore, in our proposed state-backed troll detection algorithm, only 50 features are selected using the SFBS feature selection technique.

To better investigate the performance of each classification algorithm, Figs. 3 and 4 show the boxplot and ROC curves for the cross-validation process of the six

Table 5 Accuracy comparison between algorithms used in classifying trolls accounts on Twitter

| Algorithm | Trolls type | Accuracy (%) |
|---------------------|------------------------|--------------|
| Still out there [3] | Russian troll accounts | 98.9 |
| TrollBus [11] | All types of trolls | 90.0 |
| Behavior-based [15] | State-backed trolls | 94.4 |
| TrollPacifier [9] | All types of Trolls | 95.5 |
| Proposed algorithm | State-backed trolls | 99.0 |

considered classification algorithms. The results of these two figures are consistent with the results of Table 4. RF algorithm demonstrates the best performance in both highest accuracy and lowest variance. In the second place, the CART algorithm shows good results but with a relatively high variance in the results comparing to the RF algorithm. The SVM algorithm obtained the worst results among the six compared algorithms in the cross-validation process. Figure 4 demonstrates again the robustness of using the RF algorithm with our selected set of features.

At last, Table 5 compares the results of our algorithm with the results of other algorithms used to detect Twitter state-backed troll accounts. Because of the limited number of research papers that consider state-backed trolls, we include the results of other algorithms that consider other types of Twitter trolls. Our algorithm and the work given in Im et al. [3] achieve the highest accuracy in classifying state-backed troll accounts. By analyzing the compared algorithms, it is found that the algorithms that deal with all types of troll accounts as one group obtain less accuracy than the algorithms that deal with specific types of troll accounts such as in our case (state-backed trolls) and in Im et al. [3] (Russian Troll Accounts). Regarding the behavior-based algorithm [15], although it concentrates on state-backed trolls, it obtained a low accuracy compared to our work. One reason for these results is the lesser number of considered features in Alhazbi [15] (only eight features). On the other hand, the proposed algorithm works faster than the other algorithms as a result of avoid using any sentiment analysis feature. Indeed, sentiment analysis features require higher processing time than other mathematical and text processing features. Since the time information was not provided for the compared algorithms, we executed another experiment to show the reduction in time obtained after excluding the sentiment analysis features. Examples of sentiment analysis features includes measuring the levels of positive, negative and neutral sentiments for tweets, and the emotion detection from tweets text. The results show that excluding sentiment analysis features can make a reduction of more than 35% on the processing time when applying the proposed algorithm on the data of one twitter user.

6 Conclusion

In this paper, we presented a robust and fast framework to accurately detect and classify state-backed troll accounts on Twitter using various sets of features and classifiers. In addition to using the previously proposed set of features, we proposed a set of new effective and powerful features that can classify state-backed Twitter trolls with high precision. During features generation, various combinations from several categories of features were utilized. A large dataset was gathered from diverse countries to train the classifiers. Since we inhibited the use of any language-specific features, the proposed algorithm has the ability to detect the state-backed trolls regardless of the account language or the account country.

Moreover, we reduced the required analyzing time during the classification process by excluding the features that typically require sentiment analysis techniques and take a long time to be calculated. The proposed technique achieved the highest known classification accuracy for the state-backed troll detection task on Twitter with around 99%.

References

1. Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on world wide web, pp 591–600
2. Wolfsfeld G, Segev E, Shefer T (2013) Social media and the Arab spring: politics comes first. *Int J Press Politics* 18(2):115–137
3. Im J, Chandrasekharan E, Sargent J, Lighthammer P, Denby T, Bhargava A, Gilbert E (2020) Still out there: modeling and identifying russian troll accounts on twitter. In: 12th ACM conference on web Science, pp. 1–10
4. Sahmoud S, Safi H (2020) Detecting suspicious activities of digital trolls during the political crisis. In: 2020 IEEE international conference on informatics, IoT, and enabling technologies (ICIOT). IEEE, pp 532–537
5. Atanasov A, Morales GDF, Nakov P (2019) Predicting the role of political trolls in social media. [arXiv:1910.02001](https://arxiv.org/abs/1910.02001)
6. Enjolras B (2002) How politicians use Twitter and does it matter? The case of Norwegian national politicians. Institute for Social research, Oslo
7. Tomaiuolo M, Lombardo G, Mordonini M, Cagnoni S, Poggi A (2020) A survey on troll detection. *Future Internet* 12(2):31
8. Hardaker C (2010) Trolling in asynchronous computer-mediated communication: from user discussions to academic definitions
9. Fornacciari P, Mordonini M, Poggi A, Sani L, Tomaiuolo M (2018) A holistic system for troll detection on Twitter. *Comput Human Behav* 89:258–268
10. Coles BA, West M (2016) Trolling the trolls: online forum users constructions of the nature and properties of trolling. *Comput Human Behavior* 60:233–244
11. de Lacerda TNC (2020) TrollBus, An empirical study of features for troll detection
12. Nyst C, Monaco N (2018) How governments are deploying disinformation as part of broader digital harassment campaigns. Institute for the Future. <https://bit.ly/2Mi8DYm>
13. Salamanos N, Jensen MJ, He X, Chen Y, Sirivianos M (2019) On the influence of twitter trolls during the 2016 US Presidential Election. [arXiv:1910.00531](https://arxiv.org/abs/1910.00531)

14. Jachim P, Sharevski F, Pieroni E (2020) TrollHunter2020: real-time detection of trolling narratives on twitter during the 2020 US Elections. In: Proceedings of the 2021 ACM workshop on security and privacy analytics, pp 55–65
15. Alhazbi S (2020) Behavior-based machine learning approaches to identify state-sponsored trolls on twitter. *IEEE Access* 8:195132–195141
16. Engelin M, De Silva F (2016) Troll detection: a comparative study in detecting troll farms on Twitter using cluster analysis
17. Reynard LJ (2020) Troll farm: anonymity as a weapon for online character assassination. In: Developing safer online environments for children: tools and policies for combatting cyber aggression. IGI Global, pp 230–265
18. McCombie S, Uhlmann AJ, Morrison S (2020) The US 2016 presidential election and Russia's troll farms. *Intell National Security* 35(1):95–114
19. Klimburg A (2018) Trolling, hacking and the 2016 US presidential election. *Nature* 562(7726):188–190
20. Ghanem B, Buscaldi D, Rosso P (2019) TexTrolls: identifying Russian trolls on Twitter from a textual perspective. [arXiv:1910.01340](https://arxiv.org/abs/1910.01340)
21. Badawy A, Addawood A, Lerman K, Ferrara E (2019) Characterizing the 2016 Russian IRA influence campaign. *Social Netw Anal Mining* 9(1):1–11
22. Kim D, Graham T, Wan Z, Rizoio MA (2019) Analysing user identity via time-sensitive semantic edit distance (t-SED): a case study of Russian trolls on Twitter. *J Comput Social Sci* 2(2):331–351
23. Abbott A, Tsay A (2000) Sequence analysis and optimal matching methods in sociology: review and prospect. *Sociol Methods Res* 29(1):3–33
24. Kellner A, Wressnegger C, Rieck K (2020) What's all that noise: analysis and detection of propaganda on Twitter. In: Proceedings of the 13th European workshop on systems security, pp 25–30
25. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, Dredze M (2018) Weaponized health communication: twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health* 108(10):1378–1384
26. Zannettou S, Bradlyn B, De Cristofaro E, Stringhini G, Blackburn J (2019) Characterizing the use of images by state-sponsored troll accounts on Twitter. [arXiv:1901.05997](https://arxiv.org/abs/1901.05997)
27. García-Recuero Á (2016) Discouraging abusive behavior in privacy-preserving online social networking applications. In: Proceedings of the 25th international conference companion on world wide web, pp 305–309
28. Pudil P, Novovičová J, Kittler J (1994) Floating search methods in feature selection. *Pattern Recogn Lett* 15(11):1119–1125
29. Tweepy API (2021) Website: <https://docs.tweepy.org/en/latest/index.html>. Accessed 20 March 2021
30. Jagla B, Wiswedel B, Copp'ee JY (2011) Extending KNIME for next-generation sequencing data analysis. *Bioinformatics* 27(20):2907–2909

A Review on Artificial Intelligence Based E-Learning System



U. Arun Kumar, G. Mahendran, and S. Gobhinath

Abstract Today, the e-learning system is vital to the educational system. Technology integration in the classroom aids in the effective and efficient delivery of content-based education, hence increasing student confidence. Personalized educational systems concentrate on learning behaviour, interest, and course design based on learners' aptitude and fundamental knowledge. It is a versatile teaching style that may be tailored to match the needs of individual pupils. The individualised learning strategy caters to the specific demands of each student. Understanding learners and developing a strategy that meets individual learning requirements and student interests is required for an efficient education system. An smart Tutor system is an expertise way to monitor the performance of the pupils in order to deliver tailored tutoring. Computer-based education, web-based acquiring knowledge, crowdsourcing, and virtual classrooms are examples of e-learning applications. AI may be used to automate learning processes such as building teaching materials, curriculum, training, evaluating student performance, and employing current teaching technique. Artificial intelligence is the most recent e-learning trend in higher education and business. AI aids in the provision of individual decisions through data analytics, which leads to improved education for tailored teaching and the streamlining of the educational process.

Keywords Artificial intelligence (AI) · E-Learning · MOOC · Genetic algorithm (GA) · Artificial Neural Network (ANN)

1 Introduction

Education is a continuous process that is influenced by the learner's interest, attitudes, prior subject knowledge, and ability to learn. The e-learning content is sent through

U. Arun Kumar (✉) · G. Mahendran
Department of EEE, Kathir College of Engineering, Coimbatore 641062, India
e-mail: arun.udayakumarn@gmail.com

S. Gobhinath
Department of EEE, N.G.P. Institute of Technology, Coimbatore 641048, India

the internet in the form of audio, video, presentations, text, forum discussions, webinars, and so on. E-learning materials assist learners in gaining information and skills that are relevant to their needs. The tailored e-learning strategy is learner-centric, which aids in providing learners with a suitable learning route. Although there is a wealth of e-learning content available on the internet, it can be difficult for the student to choose appropriate e-material that meets his or her learning objectives. In today's academic world, creating a tailored e-learning environment is a major difficulty. The review of literature highlights some of the key studies conducted in a tailored e-learning system. Researchers present a variety of ways for creating and developing personalised e-learning systems, such as Case-Based reasoning, Fuzzy logic—based systems, Artificial Neural Network, Genetic Algorithm, Data mining algorithms, and so on. This review aids in identifying the issues and future prospects of a current e-learning system.

2 Literature Review

Using a case-based reasoning (CBR) and genetic algorithm, Huang and Chen [1] suggested an e-learning system to build an appropriate learning route for individual learners (GA). Personalized curricular sequencing, final assessment, and assessment analysis are all possible with these tools. In a web-based setting, a researcher employs scientific investigation to generate relevant course content for learners based on their needs and learn more successfully.

Villaverde et al. [2] described the operation of a feed-forward neural network for recognising individuals' styles of learning. Felder and Silverman's approach is used to categorise pupils based on their vision & skill development. Backpropagation Artificial Neural Network (ANN) architecture uses Felder & Silverman modeling techniques to learn the relationship between students' activities in an educational environment.

Data mining techniques for e-learning optimization are depicted by Castro et al. [3]. E-learning optimization requires the use of neural networks, genetic algorithms, clustering, probabilistic reasoning, inductive learning, and visualisation techniques. E-learning concerns and pupil categorisation learning—based effectiveness can be solved using data mining techniques, according to the researcher.

Alami et al. [4] developed a proactive e-learning monitoring system which makes use of a virtual learning environment. This model employs a dynamic rule-based expert system to assess user engagement and behaviour in e-learning. The user may get information, recommendations, and tips through this system at any time. Additionally, the researchers advised that the present web-based system be enhanced by including intelligent agents.

According to Li et al. [5], tailored e-learning systems aid in the recommendation of educational materials for the student. The proposed model for a personalised web-based learning system is recommended based on its features for selecting user interest modules based on user characteristics and teaching resources. The Vector matrix is

used in the development of a user interest module. The adaptive filtering algorithm, which is based on the vector space model, is used to filter teaching resources in order to provide the learner with a personalised learning experience.

Erla Morales et al. [6] proposed an object model for an e-learning system, which aids in the development of an interactive learning system. The object hierarchy is elaborated by researchers in four distinct aggregate levels; the first level includes objects such as example, strategy, practise activity, and evaluation activity. The second level contains data, concepts, procedures, and processes, as well as content, summaries, cognitive levels, and objective and overview objects. It integrates several learning modules and activities at the third level. The e-learning system's module-level structure contributes to a learner-centric approach.

Kacalak and Majewski [7] discuss the many intelligent components of an e-learning system in detail, including voice recognition, biometric authentication, phrase meaning analysis, word and sentence recognition, and user response evaluation. The researchers are particularly interested in issues relating to the assessment of spoken language phrases. Additionally, they recommended including a hybrid neural network for problem solving in an intelligent e-learning system.

Baylari and Montazer [8] has developed a test for learners to determine their capacities; training material is produced based on the learners' knowledge, and the adaptive exam is administered. Learners' recommendations for further adaptation of learning material are gathered from review tests. A backpropagation network is used to perform supervised learning on the dataset. The system's output is compared to the outcome of the learning style index technique. According to the study, a tailored e-learning system based on ANNs is an excellent technique for learners to learn according to their abilities.

Kacalak and Majewski [7] suggested an interactive e-learning system based on natural language recognition. In this system, the user and the e-learning system may communicate verbally. Researchers used a fuzzy neural network to recognise words and sentences and a hamming neural network to recognise patterns. An effective approach for assessing learners' knowledge for interactive e-learning systems.

The Intelligent Tutoring System (ITS) was introduced by Pipatsarun and Jiracha [9]. It is an expert system that monitors learner performance and tailors teachings to the learner's learning style. For computer-based instructions, both ITS and Adaptive Hypermedia System (AH) are employed. Researchers present component-based models that incorporate expert, instructional, and learner systems for improving learner performance.

Susnea [10] expanded on the significance of ANN in improving the performance of the elearning system. The results of the online questionnaire are used to categorise the student dataset. For data analysis, ANN approaches such as Radical Based Function (RBF) and Multilevel Perceptron (MLP) are applied. The error rate for raw and processed data is computed, and a comparison analysis is done for the output produced by the RBF and MLP networks. The study discovered that the RBF network has a lower error rate than the MLP network, and the error rate is higher owing to the large number of classes. The researcher has proposed the use of ANN to overcome this issue.

Tung-Cheng et al. [11] used e-learning to construct a fuzzy inference system for learner profile analysis and selecting appropriate English language articles based on learner needs, interest, and learning capacity. The adaptive learning strategy proposed in this research aids in the improvement of learners' English learning abilities as well as their learning interest. Item response theory's recursive method is used to calculate learners' current English vocabulary. Fuzzy logic and memory cycle theories were used to choose relevant articles from the article's vast datasets for each learner. The trapezoid membership function is used to display a learner's article difficulty levels for each linguistic term, namely low, medium, and high. Membership functions are used to express the degree of membership function. In four phases, researchers construct a fuzzy inference method to determine the optimal difficulty level of an article for the learner. The linguistic characteristics of each report are specified in the first phase. The degree of membership for the linguistic quality is then calculated using a trapezoid membership function in the following step. In the Inference stage, use the AND/OR operator to create a fuzzy inference rule using three linguistic words and five fuzzy input variables [42]. The researcher use the discrete Center of Area (COA) computing approach in the defuzzifier stage to determine how appropriate an article is for a certain learner based on a quantized value between 0 and 1. The greater the value, the more relevant the topic for the learner. The Analytic Hierarchy Process is used to analyse a matrix for three criteria for each learner's performance based on the experimental and control groups. The pre-test and post-test findings for both groups are generated for assessment using the statistical t-test and z-test methods.

Judy et al. [12] describe the several goals of Massive Open Online Courses (MOOC): to offer open learning resources, cohesive learning content, and a formative learning environment. MOOCs have the significant advantage of increasing social interaction among learners by encouraging online discussions and assessing work based on peer reviews. MOOC WG12 technologies deliver dynamically composed tailored lessons for each individual student. It emphasises on video lectures, quizzes, and providing learners with quick feedback. Researchers provide information on the scalability of MOOCs in teaching, learning, and model evaluation. It works on Artificial Intelligence in Education to solve this problem (AIED). This system aids in the integration of pedagogical interfaces in order to improve the performance of e-learning systems [41].

Yathongchai et al. [13] used the K-means clustering technique to predict individuals' learning behaviour based on their profiles. Using log files in the "Modular Object-Oriented Dynamic Learning Environment" (Moodle) Learning Management System, the data mining approach is applied to extract the implicit pattern and develop the model for learner categorization (LMS). The C4.5 decision tree technique is used to classify students based on their use of e-resources in the LMS. Using free source Waikato Environment for Knowledge Analysis (WEKA) data mining tools, data is preprocessed and algorithms are applied. The learners' classification model is evaluated using a tenfold cross-validation model. According to the researcher, the proposed model assists teachers in determining student performance in LMS.

Richa [14] employs adaptive content sequencing strategies to address challenges in adaptive e-learning. The knowledge-based framework was created to address

concerns with social e-learning. The research process is separated into many stages: During the Socialization phase, the Nave Bayes classifier is utilised to categorise people based on their expertise. To handle the user's suggestions, the hierarchical analysis and fuzzy modelling approach are employed during the externalisation step. During the combination phase, the researcher created a Stigmergy-based framework to analyse the behaviour of learners. The ant-based algorithm is used to create personalised econtent for the learner. During the internalisation phase, a Context-aware Multi-agent Knowledge Sharing System based on Mobile Ad Hoc Network (MANET) was employed to share learners' knowledge. The study work is given to highlight the significance and possible advantages of incorporating social opinion into e-content creation. The goal of the research is to assist e-learners embrace e-learning approaches and to raise social awareness among e-learners.

Jain et al. [15] created an artificial intelligence-based technique for assessing students' comprehension via the use of a concept map. They used an eXtensible Markup Language (XML)-based parsing approach to evaluate student learning. The idea map method is used to assess students' understanding of a certain subject. The data is analysed by researchers by comparing the expert's knowledge domain to the students' comprehension. According to research, the customised learning system promotes adaptive learning and assists in identifying students' knowledge gaps. Researchers devised a step-by-step method in which experts/instructors constructed idea maps that were then used as a reference for assessing concept maps made by students. This map is then transformed into XML-based documents. The idea and relations are retrieved from the XML file using an XML parser. The Markov Chain decision-making model is used to forecast a student's concept map based on their degree of comprehension. This tool (AISLE) was created using Java and an XML parser to extract the necessary information from the idea map.

Salami et al. [16] proposes an intelligent fuzzy evaluation system for a successful e-learning system. This approach aids in the development of a learning profile for each student. The triangular membership function is used to categorise questions based on their level of difficulty. The Gaussian membership function is used to categorise replies based on the outcome. Six fuzzy sets were created to organise students' learning levels. Using the various operators, fuzzy rules are employed to link the fuzzy input variable to the output variable. These guidelines are used to measure learner knowledge. Defuzzification works by determining the centre of gravity in order to translate information into numerical magnitude. Fuzzy-based evaluation, according to researchers, is an effective method for determining students' knowledge level and performance in e-learning.

According to Chrysafiadi and Virvou [17], supplying e-material depending on student characteristics is a big difficulty in a web-based e-learning system. Researchers shed light on the issues with web-based e-learning systems. In an adaptive learning system, they propose employing student models to recognise learners' knowledge and attributes. To accommodate uncertainty in students' models, the researcher created a nonlinear-fuzzy Knowledge State Definer system for learning computer programming technologies. The Fuzzy Cognitive Map technique is used to determine the relationship between domain concepts and the domain knowledge

of learners. Researchers created various stereotypes based on their knowledge of specific concepts. Data was gathered for each stereotype from the learner. Because this data is imprecise, fuzzy membership functions are used. Understanding fuzzy sets combined with users' stereotypes and overlay models from remote adaptively and personalising web-based e-learning applications is described in this paper.

Limongelli and Sciarrone [18] suggested developing a fine-grained student model for the adaptive educational Hypermedia system. During e-learning, fuzzy logic is used to detect ambiguity in student behaviours, cognitive capacities, and knowledge. The Felder-Silverman Model was suggested by the researchers as a method for identifying learning styles. In this study, researchers created concept/topic graphs and questionnaires to assess participants' prior knowledge of certain ideas. The system offers a collection of topics that the particular learner should study based on the test result. The researchers compared student performance to that of a standard e-learning system. It has been discovered that this technique is effective in informing teachers about how many students comprehend a subject and in providing students with suitable comments during e-learning, resulting in student happiness.

Abeer and Thakaa [19] created an online learning method for slow learners. Researchers used the WEKA data mining technique to categorise pupils based on their learning habits. The REP (Reduced Error Pruning) tree is used for classification, utilising a tenfold cross-validation approach. A researcher-created approach for teaching English grammar to delayed learners is shown here. The outcomes of traditional learning were compared to those of e-learning, and it was discovered that e-learning is a superior teaching style for slow learners to absorb topics in a more participatory way.

Mata-Garcaa et al. [20] discuss the difficulties of customising the learning process to meet the specific needs, circumstances, and characteristics of individual students in an e-learning environment. A fuzzy logic framework has been created by the researcher to address these e-learning challenges. The fuzzy inference system is utilised in a collaborative learning environment to examine student behaviour. Fuzzy rules of association are used to delineate the ties between various forms of learning activity. Students' learning scores are predicted using Fuzzy Inductive Reasoning.

Students' knowledge and cognitive abilities were assessed using fuzzy logic, which was developed by Priya and Keerthy [21]. For the classification of pupils based on their attributes, a fixed weight neural network is utilised. Neuro-fuzzy models are utilised to gather input from students and the Information Technology Services department (ITS). Using the student's knowledge and skills, the word "fuzzification" was created. There are four steps to the model's development: fuzzifier, fuzzy relational system, fuzzy aggregation network, and finally fuzzified. Students' knowledge levels are classified using a back propagation neural network.

Learning style, attitude toward e-learning technology, and their behaviour while using an e-learning system were evaluated by Afzaal H. Seyal, Mohd Noah, and a number of other researchers [22]. They used VAKT models (visual, auditory, kinesthetic, and tactile) to figure out the best way to teach each individual student's strengths and weaknesses. Data was analysed using SPSS, a statistical software package.

An adaptive learning system for a customised e-learning system is shown by Bernard et al. [23] for a personalised e-learning system. The learner's learning styles are classified using ANN in this case. The Felder-Silverman learning style model was used to produce four ANN dimensions: Active/Reflective, Sensing/Intuitive, Visual/Verbal, and Sequencing/Global. The learning rate is calculated after the neural network result is mapped with Index Learning Style questionnaires. To boost the growth of ANN, the stratification approach is applied. Finally, findings are separated into an independent dataset using a tenfold cross-validation procedure.

Pandey and Singh [24] created a multi-agent recommender system for a tailored e-learning system using fuzzy logic. The researcher suggested a multi-agent system with several components in this paper. The student and the e-learning content communicate via the interface agent. Task Agent is in charge of handling user requests and resolving conflicts amongst them. A task agent asks the information Agent questions, and the information Agent responds with learning material. Recommender Agent suggests fresh learning materials based on the comments and requirements of the user. Database Management Agent is a database that stores course materials, user account information, feedback reports, and other information. Fuzzy logic is employed to create recommender agents in this case. Defuzzification strategies include Center of Area, Center of Maxima, and Center of Minima. The Neuro-Fuzzy system was also used to offer a Fuzzy cognitive mapping approach for guided learning.

Sarasu and Thyagarajan [25] describe how they used an adaptive Neuro-Fuzzy inference system to provide e-learning material tailored to the learner's knowledge and learning needs. They employed an ontology-based e-learning system to generate ideas and apply a Fuzzy Cognitive Map approach to detect relationships between concepts. The researcher used the ANFIS model to create a fuzzy decision tree that classified data according to the learner's knowledge. Learners' knowledge levels are determined using fuzzy inference methods. Linguistic characteristics are used to categorise learners based on their degree of expertise. When fuzzifying input, a trapezoidal membership function is used to determine the student's knowledge level. For parameter calculation, ANFIS employs back propagation methods. In a fuzzy system, the Gradient Descent Function is used to change parameters.

A fuzzy rule-based individualised e-learning system was proposed by Priya and Keerthy [21]. The researcher hypothesised that a fuzzy Rule-based system can offer a course to a student based on their competence and domain knowledge. The study reveals that adaptive e-learning increases learners' performance.

In South Africa, Jugoo and Mudaly [26] depicted students' difficulties in learning computer programming. Through action research, they are investigating the interaction between students and teachers in order to understand the blended learning environment. For qualitative analysis, the researcher creates a dynamic action research model.

Permphan and Nicha [27] utilised RBF, MLP, SVM, and PNN approaches for constructing learning algorithm. Predictive modelling using data mining, decision trees, and regression. This algorithm measures medical students' learning performance. The results of ANN and regression are compared for accuracy. Researchers

say ANN is utilised to find predictor factors and implicit relationships between predictor variables and target variables.

The researchers Bhattacharya et al. [28] suggested an ANN-based intelligent recognizer to detect the learner's learning status based on test results. For the categorization of a multidimensional pattern vector, a multilayer feed-forward network is utilised. The algorithm's output is used to identify learner performance and is represented as a confusion matrix. This system is used to determine the gap between the learning aim and the cognitive state of the learner.

For the Intelligent Tutoring System, Mandal [29], designed a new architectural framework. Use of Bloom's taxonomy may help the tutor create a domain model and course material for the student. Static and dynamic learning methods of the learner were combined to create an adaptive learning system. Learners' feelings and comments are taken into account by the system, which then recommends appropriate learning material. A java-based system for individualised e-learning was built by the researcher utilising neuro-fuzzy architecture. Learners' learning styles may be assessed using the FSML (Federal Silver Model of Learning). The Kort spiral learning model is used by researchers to determine the amount of learning of a learner based on their emotions. Artificial Intelligence (AI) is used to build the whole system.

In Beulah's case [30], A customised e-learning framework presented by aims to detect a learner's learning patterns, objects, styles, and pathways. Learners' learning styles are identified through association rule mining in FSML. The difficulty of the learning object and the degree of knowledge of the learners are both identified using a prior algorithm. Learning paths may be optimised using the genetic algorithm approach. Recommending a learning route makes use of a variety of techniques, including content-based filtering, collaborative filtering, and a hybrid method. The study's author argued that e-learning performance can be assessed by focusing on the amount of time students spend learning.

Artificial Intelligence (AI) has the potential to revolutionise e-learning, according to Kothari [31]. As an aid to knowledge provision, Artificial Intelligence in e-learning allows for real-time querying. Gamification-based e-learning relies heavily on artificial intelligence (AI) to pique students' attention and keep them engaged. Artificial intelligence (AI) is good enough at mimicking emotion to make learning more enjoyable.

The Dragonfly Neural Network method, developed by Veera et al. [32], may be used to identify students' progress in a tailored e-learning system. From the dataset, this model predicts students' mark scoring tendencies.

Math instruction for slow learners may benefit from an e-learning technique, according to Shivannathan [33]. Synchronous, asynchronous, and web-based e-learning are all terms used by the researcher to describe distinct e-learning approaches. When it comes to teaching slow learners arithmetic, the study found that e-learning strategies were more successful than conventional classroom techniques.

A hybrid fuzzy-based mapping and recommendation system developed by Appalla et al. [34], was used to examine student profiles and learning activities. Using a collaborative sequential map filtering technique for sequential mapping, students with similar learning interests may be identified and collaborative filtering

can be applied. Using a matching algorithm, e-learning content may be recommended to students based on their specific needs. Each learner's behaviour, study time, and knowledge level are stored in a cloud-based database. Learning activities were organised into a tree-structured fuzzy-based learning activity model based on the learner's profile. Learning resources such as pdf, audio, video and text may be used with the suggested technique and the fuzzy model provides better results than a knowledge-based proposed system.

This system was developed by Karthika et al. [35]. It provides e-content, the topic's sequence based on the knowledge of each learner. Intelligent tutoring systems employ the Fuzzy Cognitive Map (FCM) approach to describe and reason about knowledge in an intelligent manner (ITS). For C programming students, a mechanism has been built to dynamically detect a student's expertise level. E-learning content may benefit from the usage of FCM to find the connections between topics. The suggested system monitors e-learning performance and offers e-learners with adaptive training.

Learning Object Management (LOM) standard metadata is described by the researcher in an ontology-based eLearning paradigm that allows adaptive content management. Teacher-centeredness is the hallmark of conventional educational practise. Adaptive learning relies heavily on an inductive method. Student domain knowledge and task ontology are shown in models. Ontology-based models for the teaching domain are provided by these models [36].

In an e-learning system, data mining on the web is critical. Online use mining is a data mining approach used to uncover patterns in web data in order to meet the needs of websites and web-based applications. E-learning aids in the development of a culture of self-directed learning among students as well as the acquisition of new skills. Learning techniques may be used in any location and at any time thanks to web service-enabled technology that provides a wide range of options. A service-oriented architecture for a tailored e-learning system was proposed by the researcher. Service-oriented protocols make it simple to customise the user interface to suit the preferences of the individual. Learning materials are suggested, and students are given assistance in adapting to the teaching model and its suggestions thanks to the system under consideration. Categorization, preprocessing web information and determining what actions are required for tailored systems are all part of this approach. The HITS (Hyperlink Induced Topic Search) technique was used by the researcher to identify online communities. K-means, Suffix Tree, and LINGO clustering algorithms are utilised to improve the search engine's performance. Web use and web content mining may be used to customise e-learning courses, according to an academic study [37].

In classroom teaching, it is difficult for the instructor to adapt instructional methods to the specific requirements and interests of each student. Images, charts, graphs, audio, video, and text are just a few of the many forms in which e-learning materials may be found. It might be difficult to determine which e-material is most suited for each student [39]. VARK is a methodology for identifying a learner's preferred method of learning. Four categories are used to classify learning methods, such as visual and auditory, read/write, and kinesthetic. Learning styles may be identified using VARK online surveys [40]. Data mining methods, such as Naive Bayes

Table 1 Different algorithms and techniques for adaptive education system

| S.no | Technique used | Usage |
|------|---|---|
| 1 | Genetic algorithm, case based reasoning [1] | Create optimal learning path |
| 2 | Feed forward neural network [2] | Identify learning style of learner |
| 3 | Back propagation ANN [3] | Handle association in learning environment, provide learning recommendation |
| 4 | Inductive learning, clustering [4] | E-learning optimization |
| 5 | Adaptive filtering algorithm [5] | To provide teaching resources |
| 6 | Hybrid neural network [7] | Develop intelligent e-learning system |
| 7 | Fuzzy inference system [20, 25] | Students behavior in collaborative learning environment |
| 8 | Fold cross validation technique [13] | Separate result into independent datasets |
| 9 | Hamming neural network [7] | Pattern recognition |
| 10 | Gradient descent function [25] | To adjust parameters in fuzzy inference system |
| 11 | Analytic hierarchy process [11] | Learner performance evaluation |
| 12 | Markov chain model [15] | Decision making |
| 13 | Dragonfly neural network algorithm [32] | Student progress identification |

and decision trees, are used in this investigation. VARK-based adaptive eLearning and mentoring systems were suggested by the researcher utilising the data mining method [38]. According to literature review researchers suggested different algorithms and techniques for adaptive education system is listed in Table 1.

3 Recommended Approach of E-Learning

Figure 1 shows the suggested personalized E-Learning approach to overcome the research gap mentioned below. Table 2 shows the approach and functions of the recommended personalized E-learning System.

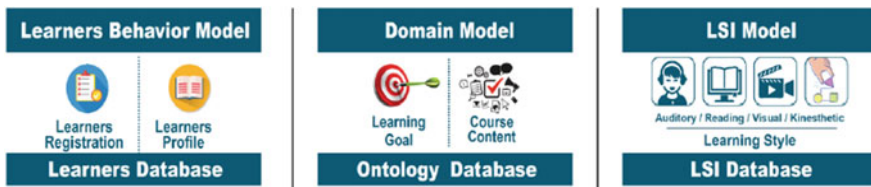


Fig. 1 Suggested AI based personalized e-learning system to overcome the research gap

Table 2 Recommended personalized e-learning approach

| S.no | Technique used | Usage |
|------|---|--|
| 1 | Behavior model | To identify the learning behavior of the individual learner |
| 2 | Learning style identification (LSI) model | To identify the learning style of the individual learner |
| 3 | Domain model | To predict existing domain knowledge of learner |
| 4 | Adaptive model | To determine the learning path and recommendation of suitable E-material to the individual learner |

4 Conclusion and Research Gap

Theoretically, a customised e-learning system may help students enhance their performance. Use of expert systems for individualised instruction is common. There is no interactive system established to educate students according to their abilities and interests, despite the fact that machine learning algorithms are used to determine a learner’s aptitude. The development of an adaptive e-learning system relies on data mining techniques, which provide an average level of accuracy. A study of the literature shows that an AI system is capable of adapting to the specific learning needs of a pupil. It can help to develop a system that provides meaningful learning experiences to the students. Intelligent tutoring systems in the e-learning environment are capable of determining the learning style and knowledge level of students.

In conventional or classroom instruction, most educationalists concentrate on identifying students’ preferred learning styles. Only a small percentage of students pay attention to the online course. E-learning systems that are tailored to each student’s unique learning style are critical to their success. These models, FSLM and VARK, encourage the identification of the learner’s learning behaviour. The learner’s static and dynamic characteristics influence their preferred learning method. To determine a student’s learning behaviour, it is important to take into account the student’s aptitude, academic achievement, interest, skill, and learning attitude. There are several advantages to creating an e-learning system that is tailored to each student’s learning style, degree of knowledge, and capacity for new information intake and retention. It is possible to increase the effectiveness and accuracy of an e-learning system by offering an ideal learning route and proposing appropriate e-learning materials for the learner, The development of an adaptive e-learning system necessitates the assessment of student performance.

References

1. Huang M-J, Chen M (2007) Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach. *Expert Syst Appl* 33(3):551–564
2. Villaverde JE, Godoy D, Amandi A (2006) Learning styles recognition in e-learning environments with feed-forward neural networks. *J Comput Assist Learn* 22(3):197–206
3. Castro F, Vellido A, Nebot A, Mugica F (2007) Applying data mining techniques to e-learning problems. *Springer Evol Teach Learn Parad Intell Environ Stud Comput Intell* 62:183–221
4. Alami ME, Casel N, Zampunieris D (2007) An architecture for e-learning system with computational intelligence. In: *Springer lecture notes in computer science knowledge-based intelligent information and engineering systems*, pp 58–65
5. Li X, Luo Q, Yuan J (2007) Personalized recommendation service system in e-learning using web intelligence. In: *Springer, computational science—ICCS, lecture notes in computer science*, pp 531–538
6. Morales Erla M, Garcia F, Barron J, Moreira AT (2008) Learning objects for e-learning systems. *Springer, CCIS*, vol 19, pp 153–162
7. Wojciech K, Maciej M (2009) E-learning systems with artificial intelligence in engineering. In: Huang D-S et al (ed), *Springer-Verlag Berlin Heidelberg, ICIC, LNCS 5754*, pp 918–927
8. Baylari A, Montazer A (2009) Design a personalized e-learning system based on item response theory and artificial neural network approach. *J Expert Syst Appl Elsevier* 36(4):8013–8021
9. Pipatsarun P, Jiracha V (2010) Adaptive intelligent tutoring systems for elearning systems. *Procedia Soc Behav Sci Elsevier* 2:4064–4069
10. Elena S (2010) Using artificial neural network in e-learning system. *U.P.B Sci Bull Series C Series C Elect Eng* 72(4). ISSN 1454–234x
11. Tung-Cheng H et al (2012) A fuzzy logic-based personalized learning system for supporting adaptive english learning. *Int Forum Educ Technol Soc*. ISSN 1436–4522 (online) and 1176–3647
12. Judy K, Reimann P et al (2013) MOOCs: so many learners, so much potential. *IEEE Intell Syst. IS-28–03-Education*1541–1672/13
13. Yathongchai C, Yathongchai W et al (2013) Learner classification based on learning behavior and performance. In: *IEEE conference on open systems (ICOS)*, Sarawak, Malaysia. 978–1–4799–0285–9/13
14. Richa S (2013) Adaptive content sequencing in cooperating social opinion in an e-learning environment. *University of Delhi, India*
15. Jain PG, Gurupur V et al (2014) Artificial intelligence-based student learning evaluation: a concept map-based approach for analyzing a students understanding of a topic. *IEEE Trans Learn Technol* 7(3)
16. Salami K, Magrez H, Ziyat A (2014) Fuzzy expert system in evaluation of e-learning system. In: *Third IEEE international colloquium in information science and technology (CIST)*. Tetouan, Morocco, *IEEE Xplore Digital Library*
17. Chrysafiadi K, Virvou M (2014) Fuzzy logic for adaptive instruction. In: *An e-learning environment for computer programming*. *IEEE Transactions on Fuzzy Systems*
18. Limongelli C, Sciarone F (2014) Fuzzy student modeling for personalization of e-learning courses. *Springer International Publishing Switzerland, LCT 2014, Part I, LNCS 8523*, pp 292–301
19. Thakaa M, Abeer M (2014) Teaching the slow learner students with and without e-learning environment. *Int J Emer Trends Technol Comput Sci (IJETTCS)* 3(4). ISSN 2278–6856
20. Mata Garciaa M, Guijarroa M et al (2015) A comparative study of the use of fuzzy logic in e-learning systems. *J Intell Fuzzy Syst*
21. Priya R, Keerthy G (2015) Rule-based fuzzy logic for automatic learning process in an e-learning environment. *Int J Adv Res Comput Commun Eng* 4(7). E-ISSN: 2278–1021
22. Afzaal H, Mohd Noah A (2015) understanding learning styles, attitudes and intentions in using e-learning system: evidence from brunei. *World J Educ* 5(1):61–72. ISSN: 1925–0754

23. Bernard J, Chang T et al (2015) Using artificial neural networks to identify learning styles. Springer, International Publishing Switzerland, AIED-2015, LNAI 9112, pp 541–544
24. Pandey H, Singh V (2015) A fuzzy logic based recommender system for e-learning system with multi-agent framework. *Int J Comput Appl* 122(7). ISSN-0975 –8887
25. Sarasu R, Thyagarajan K (2015) Retrieval of e-learning materials using adaptive neuro-fuzzy inference system. *Adv Eng Technol*
26. Jugoo V, Mudaly V (2016) The use of action research in a computer programming module taught using a blended learning environment. *Int Scient Res J* 72(9)
27. Permphan D, Nicha K (2016) Application of artificial neural networks for prediction of learning performances. In: 12th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD), 978–1–5090- 4093–3/IEEE
28. Bhattacharya S, Roy S, Chowdhury S (2016) A neural network-based intelligent cognitive state recognizer for confidence-based e-learning system. Springer, The Natural Computing Applications Forum
29. Mandal L (2016) An architecture for an intelligent tutoring system in a learning environment with distributed knowledge. Jadhavpur University, Kolkata, India
30. Beulah C (2016) A learner centric e-learning framework for web based personalized e-content delivery. Karunya University, Coimbatore
31. Kothari N (2017) Future of artificial intelligence in e-learning. <https://netbramha.com/blog/future-artificial-intelligence-e-learning>
32. Veera M et al (2017) Research study on applications of artificial neural network and e-learning personalization. *Int J Civil Eng Technol* 8(8):1422–1432
33. Shivannathan R (2017) Effectiveness of e-learning strategy on achievement of slow learners in mathamatic. *Int J Teacher Educ Res* 6:1–9
34. Appalla P, Selvaraj R et al (2018) Hybrid fuzzy recommendation system for enhanced e-learning. *Advances in systems, control and automation*. Springer Nature, Singapore P. Ltd
35. Karthika R et al (2019) Intelligent e-learning system based on fuzzy logic. *Neural computing and applications*. Springer-Verlag London Ltd, Springer Nature
36. Bouarab Dahmani F, Comparot C (2015) Ontology based teaching domain knowledge management for e-learning by doing systems. *Electr J Knowl Manag* 13(2)
37. Lian Y (2011) An online adaptive tutoring system for design centric courses. *IEEE Int Symp Circuits Syst* 1191–1195
38. Arguedas M, Xhafa F (2015) An ontology about emotion awareness and affective feedback in e-learning. In: *International conference on intelligent networking and collaborative systems*, pp158–163
39. Veeramanickam MRM, Mohana Sundaram N, Raja L, Kale SA, Mithapalli UP (2018) ‘i-Campus’: internet of things based learning technologies for e-learning. In: *International conference on intelligent data communication technologies and internet of things*. Springer, Cham, pp 1225–1232
40. Patel V, Kapadia D, Ghevariya D, Pappu S (2020) All india grievance redressal app. *J Infor Technol Digital World* 2(2):91–99
41. Mohammadi SO, Kalhor A (2021) Smart fashion: a review of ai applications in virtual try-on & fashion synthesis. *J Artif Intell* 3(4):284
42. Kumar UA, Ravichandran CS (2021) Upgrading the quality of power using TVSS device and PFC converter fed SBLDC motor. *Arab J Sci Eng*. <https://doi.org/10.1007/s13369-021-05600-z>

Text to Image Generation Using Gan



Rajni Jindal, V. Sriram, Vishesh Aggarwal, and Vishesh Jain

Abstract Text to image synthesis, one of the most fascinating applications of GANs, is one of the hottest topics in all of machine learning and artificial intelligence. This paper comprises techniques for training a GAN to synthesise human faces and images of flowers from text descriptions. In this paper, we are proposing to train the GAN progressively as proposed in the ProGAN architecture and along with that trying to improve its results by proposing a custom update rule for alpha which controls the fading rate during the progressive growth of the architecture. With experimental testing using the Oxford102 and LFW datasets, our proposed architecture and training process ensures fast learning and smooth transitions between each trained generation.

Keywords Generative adversarial networks · Text to image generation · Progressive GAN · stackGAN · Image generation · Nearest neighbour interpolation · Generator · Discriminator · Wasserstein loss · Equalised learning rate · Mini-batch standard deviation

1 Introduction

In the past few years, researchers and computer scientists have been trying to develop AI models that can closely mimic the behaviour of the human brain. Several tasks, once thought to be impossible, have now been achieved. Recently, they have achieved several milestones in the areas of images modelling, processing and generation like performing super-resolution, and this makes even restoration of antique images now

R. Jindal · V. Sriram (✉) · V. Aggarwal · V. Jain
Delhi Technological University, New Delhi, India
e-mail: vsriram_2k18co380@dtu.ac.in

R. Jindal
e-mail: rajnijindal@dce.ac.in

V. Aggarwal
e-mail: visheshaggarwal_2k18co388@dtu.ac.in

V. Jain
e-mail: visheshjain_2k18co391@dtu.ac.in

possible. Just like Image Captioning, synthesising images from textual data helps explore the visual semantic behaviour of the human brain. Text to image synthesis, besides being a budding research topic, has several important applications. Many machine learning solutions require a huge amount of training data in order to provide satisfactory results. Such a dataset could be obtained using a well-trained text to image generator model. Further, loading entire datasets into memory is not always feasible. A highly pre-trained text-to-face generator could generate data as per the need at runtime.

In this paper, we propose to develop a generative model that produces high-quality images from given text data like generating an image of a face by describing the facial features of a person, or picture of a flower from its characteristics. Our model is based on the concept of progressive growth of GANs, from generating lower resolution images in earlier stages, to gradually increasing resolutions, as the training progresses.

2 Related Work

The idea of generating images from text was first proposed by Reed in 2016 and is still a very hot topic among researchers. New models and techniques are proposed very frequently, each yielding unbelievable results which prove that there is still scope for a lot of improvement in this field. Several techniques like Pixel Recurrent Neural Networks and Regressive Autoencoders have been proposed in the past, but in comparison to these techniques, GANs produce better results.

The paper Text-Conditional C-GAN [1] divided the main task into two major subproblems which are learning a feature representation model and the second is to generate an image from these features. A simple DC-GAN architecture has been proposed which uses a feature vector generated using an RNN model as conditioning. Several algorithms such as GAN, GAN-CLS, GAN-INT and GAN-INT-CLS have been used to generate the images, and the results compared for each of them. The images generated using these techniques seem to have the required physical traits as described in the text, but they are not photo-realistic! A similar model architecture is used in [2] where the base model is taken as a DC-GAN [1] but has been trained progressively by using the training process as proposed in [3].

StackGAN [4] and StackGAN++ [5] emphasises on the fact that other available models performing this task tend to ignore minute details provided in the description. To tackle this problem, the authors of [4] divided the major problem into subproblems and proposed a two-staged architecture. The first stage is responsible for building the physical structure of the image while stage 2 enhances the output of the first stage to introduce tiny details provided in the description. The authors of [5] have extended this approach to multiple stages, each stage improving upon the results obtained from the previous. The same problem of focusing on intricate details has been addressed by the authors of [6] using an Attentional Generative Adversarial Network. It is a multi-level refinement model consisting of two main components—the first uses

the global sentence vector for generating the image, and the second enhances every attribute using a word vector for each.

One of the notable disadvantages of training the models as proposed in [4–6] is that while training such a complex architecture, there is a high chance that the generator is not properly trained because of low or no effective feedback from the discriminator. This problem can be solved as proposed in [7] where multiple discriminators are trained against a single generator so that the chances of the generator obtaining effective feedback increases. This problem can also be tackled by training the GAN progressively, similar to the proposed training methodology.

3 Experimental Design

3.1 Dataset

LFW (Labelled Faces in the Wild) is a dataset from Kaggle consisting of over 13,000 photographs of about 5700 different people. For this paper, we have used the Face2Text v1.0 dataset [8] which consists of text descriptions for 400 select images from LFW. For each image, an average of 10 text descriptions are used making our entire training dataset effectively of 4000 samples.

The Oxford102 Flowers dataset consists of images of 102 different categories of flowers, with each class containing between 40 and 258 images. We have carefully selected 16 categories from this dataset to include a wide range of flowers in terms of shape and colour. We then selected 40 images from each category and 10 descriptions corresponding to each image. This makes the training dataset effectively of over 6000 samples.

3.2 Data Preprocessing

The input to the adversarial network is a combination of textual data and image data. The descriptions and images cannot be fed directly to the neural network. Several steps of pre-processing are required to finally vectorize the data.

Text Preprocessing.

- Punctuations (e.g., ! \$ () * % @) are removed from the data as they are meaningless to describe a face.
- Stop words like a, an, the, etc. are filtered out to make the description more accurate to feed into the network.
- Each word in the description is converted to only lowercase to avoid repetition of the same words just because of uppercase characters.
- Each sentence is split into a list of tokens.

After preprocessing, a global word dictionary is created, containing each unique word from face and flower descriptions separately. This dictionary is then used to perform Count Vectorization to transform the face and flower descriptions into a vector, based on the frequency of words present in each description.

Image Preprocessing—The facial images from the dataset are not close-up shots. A good proportion of these images have complex background details which might impact the rate at which the model learns. Besides, they do not have much significance as far as the descriptions are concerned.

For the LFW dataset, we have used pre-trained OpenCV models to extract faces from the images and for the Oxford102 dataset, we have used the bounded boxes data provided in the dataset itself for cropping the images.

3.3 Architecture

The proposed architecture closely resembles the architecture of the traditional Text Conditional C-GAN—consisting of a generator that generates images of certain dimensions, and a discriminator or a critic, that classifies images of the same dimensions into real and fake. All the difference responsible for the ground-breaking results lies in the process of its training, mostly arising due to the progressive growth shown by the two adversarial networks (Fig. 1).

Conditioning Augmentation—To store the features given in the descriptions of each image in a mathematical feature space, we used an LSTM network for generating embeddings from the feature vector. The generator requires a combination of noise and a set of text descriptions in order to generate a distribution. This data cannot be passed on directly to the generator model as only vectorized data is accepted by the neural network. The conditioning augmentation layer samples a random set of latent variables from the descriptions and produces a bigger number of image-description pairs for the model. This is similar to image augmentation in the sense that this process gives rise to variations in the vectorised text input for a given image.

For conditioning augmentation, embedding is split into σ_o and μ_o

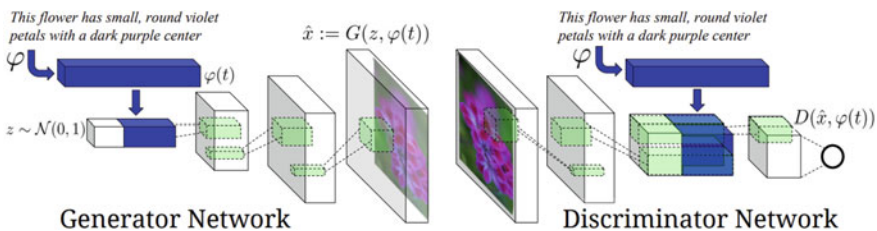


Fig. 1 Text conditional C-GAN: architecture (Source [1])

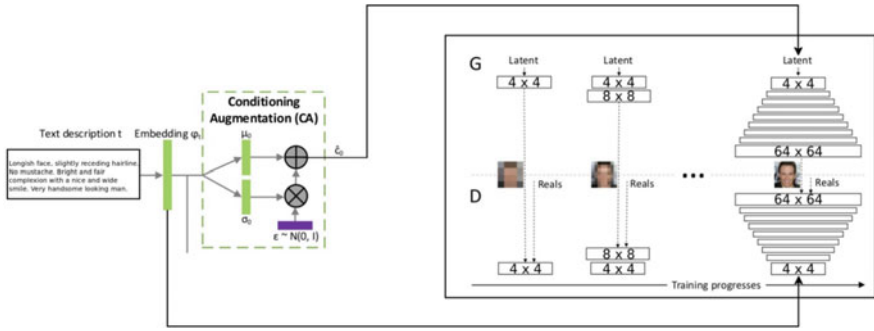


Fig. 2 Progressively growing GAN: architecture (Source [3])

$$\widehat{C}_o = \mu_o + (\sigma_o * \epsilon_o) \tag{1}$$

where ϵ_o represents random noise.

Layers in GAN—Instead of having n number of layers from the start of the training process, what we are proposing is to begin with a single layer in generator and discriminator of dimensions 4 × 4, and then to add new layers of higher dimensions as the training progresses. Figure 2 shows the resulting architecture of our proposed model.

The progressive model gradually raises the quality and dimensionality of the produced images, in our case doubling the size with each generation. The generator upsamples the input data to the required dimensions. On the other hand, the discriminator does the opposite and downsamples the image it receives from the generator.

Initially, the network starts by producing images of low resolution from the descriptions that are fed. The model then proceeds to increase the resolution progressively. This is achieved by appending new layers to both the generator and the discriminator networks. As a result, the generator and discriminator models are always the mirror images of each other. During training in the initial generations, the model focuses primarily on developing an overall structure of the image. This focus eventually shifts towards generating the finer details mentioned in the accompanying text descriptions.

Process of Fading in New Layers and Transition Between Generations—The change in the generations cannot be abrupt, as a sudden addition of layers to the neural networks might negatively affect the already trained model. Thus the new layers must be added carefully and smoothly. To demonstrate this process of fading in, consider the transition from the third generation to the fourth, i.e., from a 16 × 16 resolution to a 32 × 32 resolution, as shown in Fig. 3.

After a certain level of training in the third generation, a new layer is added to the existing model. In the generator, the image is upsampled using nearest neighbour

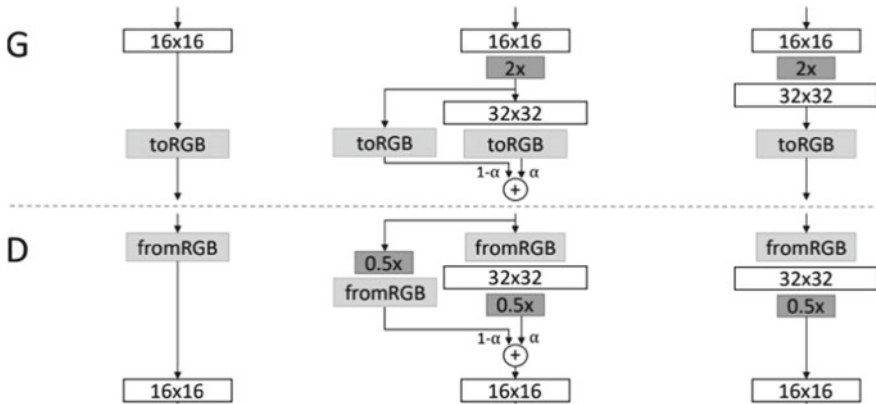


Fig. 3 Fading in new layers (Source [9])

interpolation to 32×32 , while the image is downsampled to 16×16 through an average pooling layer in the discriminator.

To increase the network depth, a block of 32×32 dimension is gradually added to the main network. This additional block contains two convolution layers and one upsampling layer in the generator model and in the discriminator, it has two convolution layers with one down sampling layer. In case of the generator, this new block is assigned some weightage, say α , and the previous layer a weightage of $(1-\alpha)$ as given in Eq. (2). The value of α gradually increases from 0 to 1. This prevents hampering of the weights trained during the previous generation due to the addition of a new layer with random weights.

$$M' = (1 - \alpha)L + \alpha M \tag{2}$$

where L denotes the penultimate layer, M denotes the newly added layer and M' denotes the updated M layer.

The weightage parameter α that denotes the division of importance given to the two largest, outermost layers in the generator and the two largest, innermost layers in the discriminator, needs to be updated as training in a generation proceeds. Initially, more weightage is given to the penultimate layer, and eventually, the ultimate layer gets complete weightage.

Proposed Update Rule for α —While training our GAN progressively, we observed that generation transition had a drastic impact on the weights of the newly added layer, thus impacting the training results. For making this process smoother and more seamless, and also allowing the new added layer to fade in without any sudden changes in the weights of the model, we propose a quadratic update rule for the purpose.

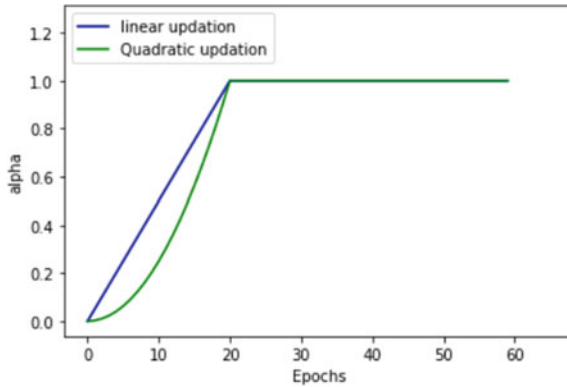


Fig. 4 Linear and quadratic update rules for α

The novelty in our proposed rule lies in the updation mechanism of this parameter. Through our comprehensive literature review, it became evident that several researchers, like in [3] and [2], use a linear update rule for the parameter.

But as α varies in the range $[0,1]$, the gradient of a quadratic update is lower than that of a linear update rule, as evident in Fig. 4. This means that, in case of a quadratic update rule, the value of α increases slower than in case of a linear update equation. Consequently, the generational shift is more gradual and is expected to yield better results.

Loss Function—The loss function is an important factor on which the training of both the generator model and the discriminator model depends. Our implementation has used the Wasserstein loss function, which is viable when the discriminator is a critic. Instead of classifying images into real or fake, a “critic” discriminator assigns a score to the generated batch. The aim of the discriminator is to simply assign a score as large as possible to real images, and a score as small as possible to fake ones.

The discriminator thus tries to maximise the value of the critic loss, given in Eq. (3)

$$Loss = D(x) - D(G(z)) \tag{3}$$

while the generator tries to maximise the generator loss

- D(x): score assigned to a real instance by the discriminator
- G(z): generator output for a random sample z
- D(G(z)): score assigned to a fake instance by the discriminator

As training a GAN is a very difficult task, we have also used certain techniques like mini-batch standard deviation for avoiding mode collapse, equalised learning rate for keeping the training process stable and feature vector normalisation for scaling the weights.

Implementation Details—The maximum length of the captions that we allowed in our implementation is of length 50. The hidden size of the LSTM network is 512, number of layers is 3 and the embedding size is taken as 128. The output size after performing the conditioning augmentation is of length 256. For progressive GAN, the learning rate is 0.001, normalising factors i.e. beta_1 is 0 and beta_2 is 0.99, eps is 0.00000001 and drift is 0.001.

4 Results and Analysis

The proposed model, after training in the fifth generation, began producing satisfactory results. Figure 5 shows the generated images belonging to the various generations throughout the training process, thus demonstrating the progressive evolution of the generator model.

Generation 0 generates a 4×4 patch-like image that is incapable of representing most of the features described in the text, and only contains a few colours. The next two generations produce 8×8 and 16×16 images, with more distinction in the colours. The model in the third generation begins producing 32×32 images, with more colours and features becoming prominent. After the fourth generation, the model began production of 64×64 images, where the features gradually began resembling those being described in the accompanying text. The model after five generations becomes capable of producing 128×128 images, where even the finer details become eminent.

For evaluating our models, we used inception score as an evaluation metric. Inception score for the model which has been trained using the linear update rule comes out to 1.881 and for the model which has been trained by our proposed update rule comes out to be 1.878. Although this metric is not appropriate for our field of study as mentioned in [4] because it cannot judge how well the images have picturised the text descriptions. But on manual observation, we can see that the images produced using our proposed update rule are smoother and more aligned towards the given captions as input.

Figure 6 depicts a few sample images generated by the model, corresponding to the text descriptions that were provided. The produced images, in the case of both facial data and flower data, match the corresponding descriptions to a large extent. The image also depicts the difference arising due to different rules used to update the parameter α .

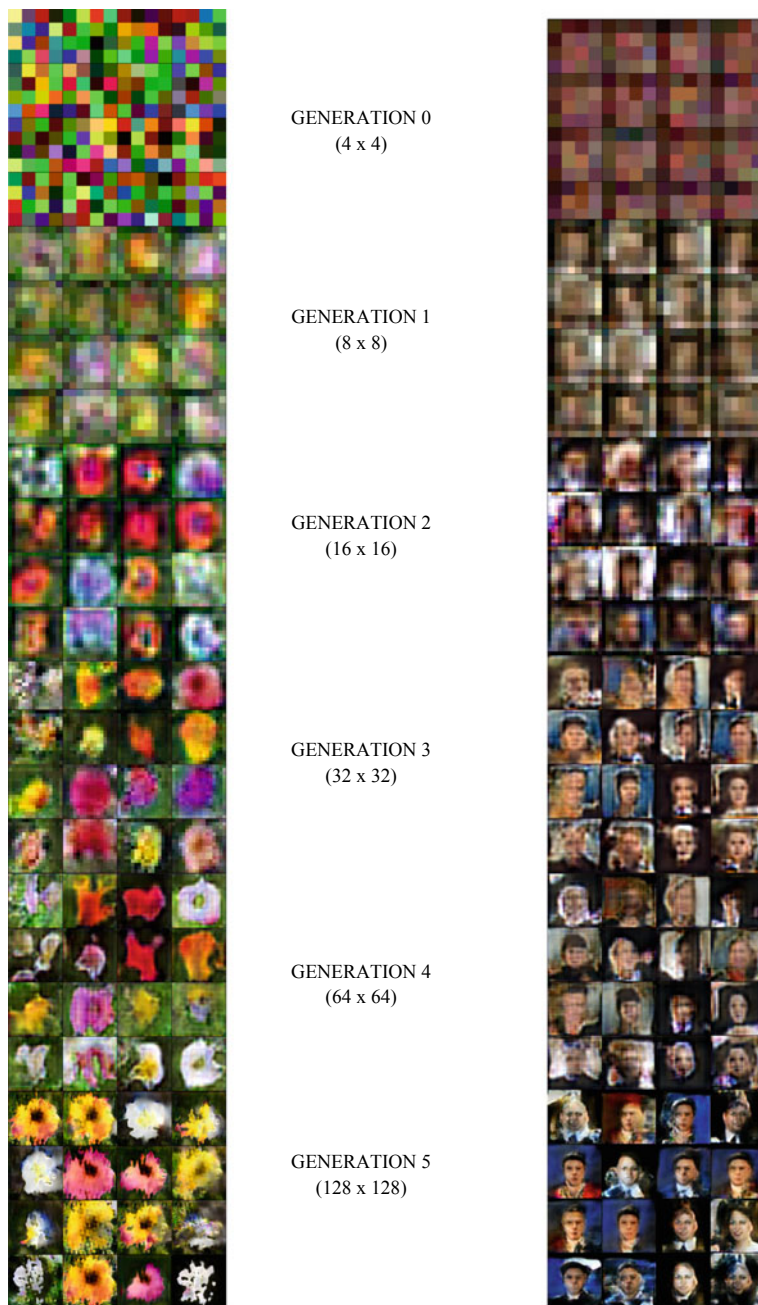


Fig. 5 Results



Fig. 6 Comparison of images generated based on given text description, through **a** linear update rule **b** proposed quadratic update rule

5 Conclusion and Future Work

In this paper, we have proposed the synthesis of images from text descriptions using a progressively growing GAN, where the generator and discriminator continuously evolve to improve the quality of the produced images. We have concluded that this model begins producing satisfactorily good results, faster than most of the proposed techniques.

Besides this, we have also proposed a novel update rule for the weightage parameter α . When α is updated using a quadratic update function rather than a linear one, the processing of fading in becomes further smoother, thus producing even better results.

In future, we plan to train the model on the entire dataset for higher generations producing high-resolution images (512×512 and 1024×1024). Images with a resolution of 128×128 have a lot of scope for improvement by training the model for further generations. Another way around this would be to enhance the particular feature quality of the images using word vectors and low-resolution images to train the network.

Appending a super-resolution GAN model to the existing architecture could produce high definition quality images. The network currently uses nearest neighbour interpolation to upsample the images during the transition between the generations. Another alternative is using bilinear interpolation instead in the hope of better results.

References

1. Reed S, Akata Z, Yan X, Logeswaran L (2016) Generative adversarial text to image synthesis. Proceedings of the 33rd international conference on machine learning, New York, NY, USA. In: JMLR, W&CP, vol 48
2. Sawant R, Shaikh A, Sabat S, Bhole V (2021) Text to image generation using GAN. Proceedings of the international conference on internet based control networks & intelligent systems—ICICNIS 2021, Available at SSRN: <https://ssrn.com/abstract=3882570> or <https://doi.org/10.2139/ssrn.3882570>
3. Karras T, Aila T, Laine S, Lehtinen J (2018) Progressive growing of gans for improved quality, stability, and variation. Published as a conference paper at ICLR 2, 6th international conference on learning representations, Canada
4. Zhang H et al (2017) StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks. Published in 2017 IEEE international conference on computer vision (ICCV), pp 5908–5916. <https://doi.org/10.1109/ICCV.2017.629>
5. Zhang H et al (2019) StackGAN++: realistic image synthesis with stacked generative adversarial networks. In IEEE Trans Pattern Anal Mach Intell 41(8):1947–1962. <https://doi.org/10.1109/TPAMI.2018.2856256>
6. Xu T et al (2018) AttnGAN: fine-grained text to image generation with attentional generative adversarial networks. Published in 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 1316–1324. <https://doi.org/10.1109/CVPR.2018.00143>
7. Durugkar IP, Gemp I, Mahadevan S (2017) Generative multi-adversarial networks. In: ICLR
8. Huang GB, Ramesh M, Berg T, Learned-Miller E (2019) Labelled faces in the wild: a database for studying face recognition in unconstrained environments. vis.cs.umass.edu
9. An Introduction To The Progressive Growing of GANs, TheAILearner. <https://theailearner.com/tag/progressive-gans/>
10. Chen X, Qing L, He X, Luo X, Xu Y (2019) FTGAN: a fully-trained generative adversarial networks for text to face generation. Published as a conference paper at ICLR [arXiv:1904.05729v1](https://arxiv.org/abs/1904.05729v1)
11. van den Oord A, Kalchbrenner N, Kavukcuoglu K (2016) Pixel recurrent neural networks. Proceedings of the 33rd international conference on machine learning, New York, NY, USA. In: JMLR, W&CP, vol 48
12. van den Oord A, Kalchbrenner N, Vinyals O, Espeholt L, Graves A, Kavukcuoglu K (2016) Conditional image generation with PixelCNN decoders. In: 30th conference on neural information processing systems, Barcelona, Spain
13. Li B, Qi X, Torr PHS, Lukasiewicz T (2020) Lightweight generative adversarial networks for text-guided image manipulation. Published in 34th conference on neural information processing systems, Vancouver, Canada
14. Che T, Li Y, Zhang R, Hjelm RD, Li W, Song Y, Bengio Y (2017) Maximum-likelihood augmented discrete generative adversarial networks. [arXiv:1702.07983](https://arxiv.org/abs/1702.07983)
15. Gauthier J (2015) Conditional generative adversarial networks for convolutional face generation. Technical Report
16. Arjovsky M, Bottou L (2017) Towards principled methods for training generative adversarial networks. In: ICLR
17. Huang X, Li Y, Poursaeed O, Hopcroft J, Belongie S (2017) Stacked generative adversarial networks. In: CVPR
18. Reed S, Akata Z, Yan X, Logeswaran L, Schiele B, Lee H (2016) Generative adversarial text-to-image synthesis. In: ICML
19. Manoharan J (2021) Samuel: capsule network algorithm for performance optimization of text classification. J Soft Comput Paradigm (JSCP) 3(01):1–9

20. Hamdan Y (2021) Babiker: construction of statistical SVM based recognition model for handwritten character recognition. *J Inf Technol* 3(02):92–107
21. Vijayakumar T, Vinothkanna R (2020) Capsule network on font style classification. *J Artif Intell* 2(02):64–76
22. Sindu S, Kousalya R (2016) Recurrent neural network for content based image retrieval using image captioning model. In: international conference on computational vision and bio inspired computing, pp 1067–1077. Springer, Cham

SARS-CoV-2: Transmission Predictive Tool Based on Policy Measures Adopted by Countries Using Basic Statistics



Charles Roberto Telles and Archisman Roy

Abstract This study starts with a problem concerning “S” (susceptible) and “R” (removed) compartments of the S.I.R model equation, due to confounding scenarios of variables that sustain it for policies’ predictive analysis. Considering “I” (infectious) term as the most stable point of analysis to achieve proper predictive methods globally for policies measures. To observe this assumption, statistical mode analysis was performed to extract a numerical indicator for the “I” compartment. This tool could be used with precise data about when policies in a given country, region, or spatial sample started and it is possible to track the type and duration of policies with the patterns outputs for each analysis. One other advantage of the method is that there is no need to describe the event in its full behavior (very long-time lengths) as well as wait for longer time series to be able to make predictions about how policies influence the event.

Keywords COVID-19 · Mathematical modeling of infectious diseases · Predictive methods · Modeling of processes and systems

1 Introduction

Your contribution should be prepared in Microsoft Word. This research points to the asymptotic instability of the S.I.R. (susceptible, infectious, and removed) model to predict the behavior of SARS-CoV-2 infection dissemination patterns over the effect of policies adopted by countries. Mainly for the (S) and (R) terms of the equation, the predictive results can fail due to the confounding environment of variables that sustain the virus contagion within the population complex network basis of analysis.

C. R. Telles (✉)

Secretary of State for Education and Sport of Paraná, Curitiba, Paraná, Brazil
e-mail: charlestelles@seed.pr.gov.br

A. Roy

Department of Physics, Institute of Sciences, Banaras Hindu University, Varanasi, India

By removing (S) and (R) compartments from the equation, it leads to the asymptotic feature of the (I) compartment as the most stable point of analysis to achieve proper predictive methods and results globally for policies and COVID-19 reduction.

The starting point to the limitation of predicting future rates of transmission was importantly and deeply noticed by Manzo and Roberts et al. [1, 2] during the pandemic dissemination, and also briefly described by Merchant and other authors [3–6]. These authors presented described analysis of SIR models and variants failures to achieve a good to fit predictability and sensitivity of how SARS-CoV-2 infectious spreading patterns might occur over time (also this limitation being live verifiable by automated tools databases available to perform it [7] as exemplified in "Fig. S1"). Many pieces of research based on SIR models or their variants [8], just to mention very few examples [7, 9–18] are confronted with barriers found mainly based on (S) and (R) compartments of the SIR model due to pandemic complex scenario where distinct spreading and dissemination patterns of infection can be listed in the "S2 figure".

Despite new formulations based on agent-based modeling approaches [19, 20] for SARS-CoV-2 are being deeply described [21–23], the lack of fixed point orientation and asymptotic instability for (S) and (R) are still visible when we consider that an agent-based model can be under or overestimated by the type and duration of policymaking worldwide [24–34]. This confounding environment of study presents itself for COVID-19 epidemiological models as an impossible methodology to be used in terms of prediction since embedded components between policies range, individual behavior compliance, and spreading patterns are presented and act as a limiting factor for the machine learning analytics for (S) and (R) as well as the pre assumptions of a global homomorphism basis for the individual scale of an infected person and SARS-CoV-2 behavior to population worldwide.

Recent researches have been pointing to the robust analysis of convergence orientation within policies adoption by countries (non-pharmaceutical interventions (NPIs)) [24–34]. This can be understood in this sense, since each policy for each country, diverges itself, generating distinct scenarios for (S) and (R) as non-global solutions. Considering all countries' epidemic behavior in convergence to the policies adopted by each country using a SIR model without adaptations to explain it globally can be erroneous. (I) compartment presents a good performance of indicator of analysis behavior to infer with numerical real data results of daily new cases, without the noise [1–7] caused by some compartment components of SIR models approaches of predictability and with true correlation concerning policies adopted globally worldwide.

Concerning statistics, the fixed point stable parameter of (I) can create a confident region of statistical analysis [24–34] in terms of observation of the exponential growth of virus along days, and therefore it could be more conclusive to many mathematical infectious disease models (SIR stochastic or deterministic approaches) that were created since the beginning of epidemics and later pandemics dissemination as a global solution worldwide.

To avoid inputs of data that can create a confounding environment of study, where the mathematical simulation of SIR model finds limitations due to these non-homologous data (no global solutions) [1–7] and consequently heteroscedasticity form for data results; in this sense, it is justifiable that the analysis for (I) must be correlated directly on policies adopted by countries as the most reliable, at the moment [24–35], the form of reducing COVID-19 cases while no vaccine or drugs present consistent and effective use for treating the disease or stop virus propagation, that could influence directly the (S) and (R) components, thus, performing a possible robust predictive analysis. Therefore, the daily new cases (I) behavior is attributed briefly to the policies adopted by countries in a straightaway.

2 Materials and Methods

2.1 Related Work

The confounding environment of variables involved in the pandemic forms of spreading and dissemination patterns assume distinct values for the exponential behavior of infection over time not only for the same sample of analysis (country) [7] but samples compared to each other [36, 37].

To identify a possible fixed point orientation with not homologous data from the pandemic evolution, concerning the system as a global homomorphism over time, no stability pattern was found for (S) and (R) in recent researches [7–23] except the (I) [24–33] in its simple form of daily new cases data and how policies have a strong influence on it [24–34]. This stability feature for (S) and (R) is not commonly observed in the empirical terms of the analysis for short periods of observation while these data presented time-varying aspects and unresolved partial differential equations based on epidemics evolution and it still applies for long periods as well [1, 2, 7, 38]. While (S) and (R) periodic fixed point that could uncover a stability global pattern formation needed for a deterministic SIR modeling is not obtained (vaccines/drugs/immunization) [7–23] and it generates probabilistic distributions permeating the system for short time lengths (local homomorphism) and longtime lengths (global homomorphism) generating distinct distributions shapes and scales for each country. This set of variables that form the system assumes different behaviors over periods and it is not possible to determine universally that the infection dissemination system have a persistent not embedding behavior in the way it was constantly observed in many articles of research using SIR and its variants models as global tools [7–23].

The compartment (I) resembles the growth of infection over days that depending on the policies adopted by a country expresses higher or lowers exponential growth behavior. If considering the exponential growth as how efficient are the policies adopted to reduce transmission patterns, then, we have an indicator of analysis that is pre-assumed on fixed point orientation of infection dissemination patterns, and

hence, a valid indicator of countermeasures, monitoring aspects and expected results in a wide domain of policies to be adopted.

2.2 Proposed Work

For this analysis, consider the infected population samples $Y = (Y_1, \dots, Y_n)$ with exponential behavior of infection $f(Y; \lambda) = \lambda e^{-\lambda Y}$, where samples are taken from zero cases until the observed maximum exponential growth reached per population over time $Y(t)$, for each country [36, 37]. Starting with a Weibull shape behavior $k = 1$ or $k > 1$ of irregular distributions of SARS-CoV-2 defined in the original form as $\frac{dI}{dt} = \beta \frac{SI}{N} - gI$, it can be understood as the general distribution for $I = \left(\frac{\omega}{\lambda}\right)^k$ Eq. (1), where the infected I is influenced by the unpredictable scale of infection λ (N) with inconsistent transition rate (βSI) defined as ω spreading patterns (no global solution) and is not assumed for gI in the original form of R, that there is a normal distribution output for this virus spreading and dissemination patterns. For this reason, the proposed method of analysis considers only the aspect of the evolution of cases rather than pre assuming full immunity or susceptibility. Therefore, the unpredictable shape k (close to reality shapes), mainly defining this shape caused λ and ω asymptotic instabilities generated by S and R compartments over time as a global proposition. This equation represents the presence of confounding variables environment ω with an unknown predictive scale of $exp\lambda$ or maximum likelihood estimator for λ due to nonlinear inputs for S and R, and therefore generating nonlinear local outputs k (asymptotic instability) [39] by the virus local dissemination I over Y (population).

To separate data from time series analysis of daily new cases, concerning Eq. (1), the outputs with heteroscedasticity form for k and λ need to be removed to achieve the $\kappa < 1$ Weibull parameterization aspect [40] of best distribution with the most reliable space of analysis (attractive orientation of non-homologous data). To achieve this, it was modified from the Eq. (1) to $I = \left(\frac{Y(t)}{T}\right)^{k < 1} - \left(\frac{\omega}{\lambda}\right)^k$ Eq. (2), hence with the global proposition to $I = I' - (S + R)$, where I is asymptotic to I' when infection spreading and dissemination patterns are under policy measures influence [24–34].

Considering a new scope of analysis of time series data given in the last paragraph, it is needed now to uncover the regions of the graph in which confounding scenarios can be dismantled with a more robust relation of cause and effect according to Eq. (2). To achieve this, Eq. (2) is found as a persistence diagram existence [41] by mapping each adjacent pair to the point $(f(Y), f(t))$ local minimum and maximum values of the daily new cases in a given sample, due to worldwide behavior of epidemic growth and subtle reduction due to policies measures. Therefore, resulting in critical points of Y function over time t as y (data samples range) in not adjacent form for all the epidemic linear curvature, thus expressing random critical values defined by

$I' = \left(\frac{Y(t)}{T}\right)^{k < 1}$, where for some regions y_n of times series data, the behavior of variables in terms of cause and effect vary enough to infer an analysis of it linear.

To achieve data filtering of the region of stability defined as $I' = \left(\frac{Y(t)}{T}\right)^{k < 1}$, it is needed to filter $f(Y_n) - f(t_n) = y_n$ unstable critical points (oscillatory instability of S and R) to an attractive behavior for the distribution of data samples as $\lim_{n \rightarrow \infty} F_{S_n}(Y, t) + F_{R_n}(Y, t) = F_{I_n}(Y, t)$ for strictly determined points, and therefore, the redefinition of samples $y_n = (Y_1, \dots, Y_n)$ to be collected and mean μ over time of samples behavior is a necessary condition to extract objective numerical solutions of the problem.

At this point, it is possible to reject unstable critical points generated with non-normality form of $\kappa < 1$ for policies and its outcomes in the time series as $\pi < y_n < \frac{\pi}{2}$, thus, obtaining a local maximum and minimum points of the event as an average region where solutions for predictive analysis (mode) can be obtained. By having y_n with a limited number of samples considering policies influence range, the nonlinear oscillations of epidemic behavior, limited between local maximum growth defined by π by its half curvature oscillations $\frac{\pi}{2}$ as a local minimum, assume therefore a non-periodic (2π) condition due to $\kappa < 1$ as shown in Fig. 1.

Therefore, the desired y_n of data assumes a range of values at a region condition to normality like $\pi < y_n < \frac{\pi}{2}$ where persistent homology of cause and effect can be found for policies under $t \therefore \kappa < 1$. The oscillations pairing region of $\sin(\pi) = 1$ and $\cos(\pi) = 0$ for a new periodic T desired coordinates of persistent fluctuations in $(f(Y_n), f(t_{k < 1})) = y_n$ of stability can differ itself a long time, where policies range of influence is no longer stable (weak boundaries points of persistence), and therefore assuming $t + 1$ discrete form, defined as $y_n = f(f(Y_n), f(t_{k < 1})) \int_{\frac{\pi}{2}}^{\pi} \mu \sum(Y_0, \dots, Y_n) dx$.

In this simple and overall form, redefining the mean shape for the infection spreading and dissemination patterns of I' approximately as $y_n = 1/\lambda$ under $\kappa < 1$, the numerical representation of the scale considering a global homomorphism of the

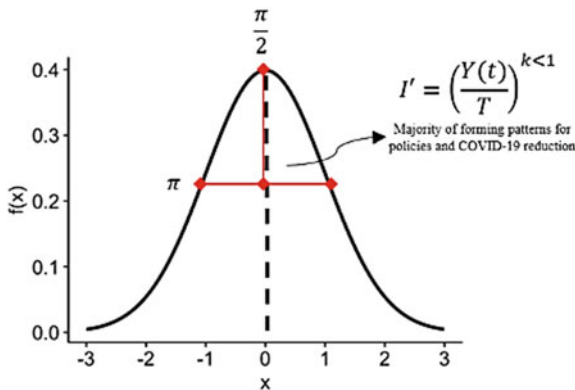


Fig. 1 The region where policies made worldwide to combat COVID-19 spreading and dissemination patterns, present the most notable influence

persistent homology [41] observed within a region of the graph, that can be with the mode statistics in the form like $y_n = Y_n(t_{k < 1})$.

Concerning mobile apps and person-to-person surveys to detect SARS-CoV-2 features, many statistical results and predictive methods present possibly false information, the subjectivity of data, and limited scope of analysis due to ethical or policy implications. For these limiting points, the data sets used to estimate values for S and R can challenge serious flaws and inconsistency [42–46].

3 Result Analysis

The results were based on a period of analysis of the COVID-19 outbreaks, where no vaccines or other strong influencing factor was existent. It was considered this period as a proper one to obtain epidemics tracking tool based on policy analysis.

Table 1 presents the growth of infection per amount of population $Y(t)$ observed for a given sample (country) by the empirical observations, therefore, in the proposition, $I = I' - (S + R)$, it gives the values of the exponential growth mean reached in the I' shape in average days peaks of infection since the outbreak. The data collection for “y” and “t” were retrieved from the Worldometer website in the given dates of analysis during 2020.

- 7 days’ sensitivity and prediction: from the epidemics outbreaks until days considered as of 7 and 13 April. The behavior of infection was obtained for prediction and sensitivity, indicated by arrows. In the 25 samples analyzed, 12 presented different behavior of infection dissemination patterns while 13 expressed a constant evolution of infection that indicates also a positive analysis for predictive statistics even with a low amount of time considered. While in 7 days of predicting analysis, several countries implemented new policies [47, 48] measures to reduce SARS-CoV-2 dissemination patterns and the prediction was affected highly by this measure.
- 19 days’ sensitivity and prediction: from the epidemics outbreaks until days considered as of 13 April to 01 May, counting for a higher period of analysis. For this new predictive result based on policy influences adopted, for 19 days in the future, the behavior of infection obtained for prediction on 13 May, indicated by arrows, was modified on 01 May. In this period, the same as happening before, new policy measures were adopted or in the course by countries [47, 48], thus influencing the predictive/sensitivity analysis. In the 25 samples analyzed, 9 presented different behavior of infection dissemination patterns while 16 expressed a constant evolution of infection that indicates also a positive analysis for predictive statistics even with a low amount of time considered.
- 33 day’s sensitivity and prediction: now from 01 May to 02 June, in the 25 samples analyzed, 11 presented different behavior of infection dissemination while 14 expressed a constant evolution of infection that indicates also a positive analysis

Table 1 The sensitivity of time and exponential dissemination patterns confronted with policy measures adopted by countries. \uparrow , \downarrow and $=$ symbols represent the transmission patterns. Data source Worldometer

| Country | R 01 May | R 7 April | R 13 April | R 01 May | R 02 June |
|-------------|---------------------|--------------------|---------------------|---------------------|---------------------|
| Worldwide | 606,06 \downarrow | 675,67 \uparrow | 886,07 \uparrow | 606,06 \downarrow | 572,51 \downarrow |
| Europe | 329,67 \downarrow | 410,95 \uparrow | 379,74 \downarrow | 329,67 \downarrow | 162,6 \downarrow |
| Italy | 78,43 $=$ | 90,90 \downarrow | 78,43 \downarrow | 78,43 $=$ | 78,43 $=$ |
| South Korea | 12,5 $=$ | 12,5 \downarrow | 12,5 $=$ | 12,5 $=$ | 0,09 \downarrow |
| China | 53,57 $=$ | 53,57 \downarrow | 53,57 $=$ | 53,57 $=$ | 53,57 $=$ |
| Iran | 43,47 $=$ | 44,44 \uparrow | 43,47 \downarrow | 43,47 $=$ | 25,64 \downarrow |
| Spain | 90,90 \downarrow | 131,57 \uparrow | 121,95 \downarrow | 90,90 \downarrow | 90,90 $=$ |
| France | 88,88 \downarrow | 138,88 \uparrow | 100 \downarrow | 88,88 \downarrow | 3,89 \downarrow |
| US | 566,03 \uparrow | 625 \uparrow | 540,54 \downarrow | 566,03 \uparrow | 235,29 \downarrow |
| Brazil | 98,03 \uparrow | 25 \uparrow | 28,57 \uparrow | 98,03 \uparrow | 180,72 \uparrow |
| Germany | 97,56 \downarrow | 105,26 \uparrow | 100 \downarrow | 97,56 \downarrow | 2,73 \downarrow |
| Russia | 108,69 \uparrow | 15,62 \uparrow | 27,02 \uparrow | 108,69 \uparrow | 96,15 \downarrow |
| UK | 86,20 \uparrow | 64,51 \uparrow | 69,64 \uparrow | 86,20 \uparrow | 22,22 \downarrow |
| Singapore | 8,77 \uparrow | 0,96 \uparrow | 1,75 \uparrow | 8,77 \uparrow | 5,61 \downarrow |
| Portugal | 10,20 \downarrow | 15,15 \uparrow | 13,15 \downarrow | 10,20 \downarrow | 1,85 \downarrow |
| India | 17,24 \uparrow | 8,33 \uparrow | 9,43 \uparrow | 17,24 \uparrow | 64,44 \uparrow |
| Canada | 31,25 \uparrow | 27,77 \uparrow | 24,39 \downarrow | 31,25 \uparrow | 9,37 \downarrow |
| Japan | 7,46 \uparrow | 3,92 \uparrow | 7,14 \uparrow | 7,46 \uparrow | 0,30 \downarrow |
| Sweden | 8,62 \uparrow | 5,71 \uparrow | 0,65 \downarrow | 8,62 \uparrow | 6,66 \downarrow |
| Argentina | 1,75 \downarrow | 1,51 \uparrow | 2,56 \uparrow | 1,75 \downarrow | 6,17 \uparrow |
| Chile | 9,61 \uparrow | 10 \uparrow | 8,33 \downarrow | 9,61 \uparrow | 41,66 \uparrow |
| S. Arabia | 18,51 \uparrow | 3 \uparrow | 4,16 \uparrow | 18,51 \uparrow | 17,44 \downarrow |

for predictive statistics even with a high amount of time considered (33 day's sensitivity and prediction).

An important point for removing S and R terms of the SIR model equation is related to the table's data and random dates that refer to different epidemic phases of data collection for each country. These distinct phases are important to be considered together in any sample collection due to the need for a methodology that can extract behavior of the disease in the not optimal evolution (SIR deterministic approaches) of the virus infection and policies adopted by countries, hence, revealing in the complex scenario the disease dynamics under a confounding environment of data with more sensitivity and possible rapid response of countries policies as presented in Fig. 2.

Concerning study limitations and class imbalance of data retrieved, note that no specification of policy measures was provided. The objective of this research is not to determine which measure gives the best result, but just to provide the observation that policies do influence the COVID-19 spreading patterns. For further interests on

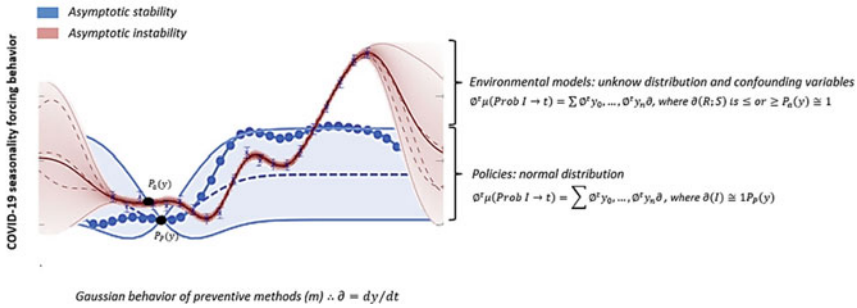


Fig. 2 Stability region of analysis for mode statistics used in this study can be viewed by this compiled algorithmic information

the type, duration, and efficacy of measures adopted by countries, refer to [24–34]. Also, data was collected during period where no vaccines and other technological and scientific knowledge was fully or partially developed, that is during first semester of 2020.

4 Conclusion

A method to be used as a qualitative and numerical indicator of local or global COVID-19 epidemics dissemination patterns is offered and it is based on daily new cases time-series data.

The results obtained showed that the indicator can be valid as a predictive method of analysis in terms of identifying how the epidemics are behaving over time. The method used is highly sensitive to policies adopted by countries having a confidence power relative to how a country responds to an epidemic in terms of these policies.

The importance of the research is about which strategic policy interventions can be proposed while events take place over time as well as evaluations of how policies already adopted can be failing to achieve the desired result.

Statement of Ethics: No humans or animals were involved in this study. Ethics approval was not required. No trial registration is applied for this study.

Data Availability

- (a) All data about COVID-19 daily cases over time are held in a public repository: <https://www.worldometers.info/coronavirus/>.
- (b) All data about COVID-19 policies adopted by countries are held in a public repository:
 1. <https://www.ilo.org/global/topics/coronavirus/country-responses/lang--en/index.htm>

2. <https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#A>

(c) All other relevant data are within the manuscript and its supplementary materials.

Supplementary Materials

Figure S1 is available online at www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html (Fig. S2).

National Forecast

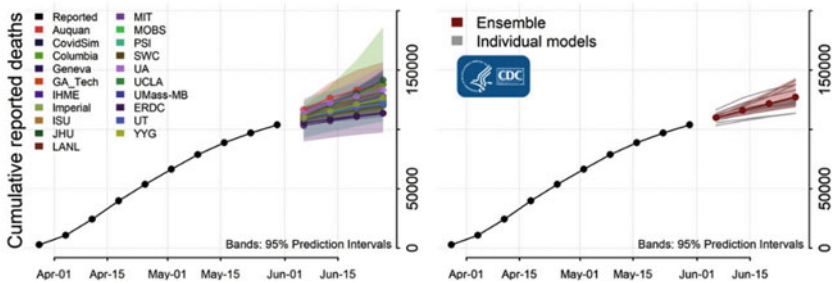


Fig. S1 Uncertainty of predictive methods with automated tools. *Source* Centers for Disease Control and Prevention (CDC). Retrieved on June 11, 2020. Most of the models present a shaded region where prediction might face an uncertain degree of confidence. S1 supporting information image from CDC is under license of Public Health Image Library (PHIL): royalty-free

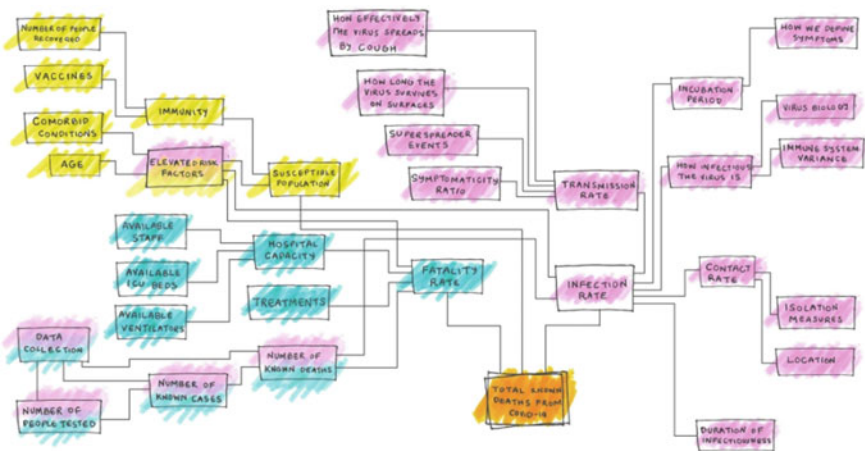


Fig. S2 Confounding environment of variables that influence COVID-19 infectious spreading patterns

S2 figure is available online at <https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/>.

Image rights - Authors: By Maggie Koerth, Laura Bronner, and Jasmine Mithani. Title of featured article: “Filed under Coronavirus”. Published Mar. 31, 2020. Source: ABC News Internet Ventures. All rights reserved – Five Thirty-Eight. Source: Centers for Disease Control and Prevention (CDC). Retrieved on June 11, 2020.

Code availability (software application or custom code): Not applicable.

Authors’ Contributions all authors contributed equally.

Funding “This research received no external funding”.

Conflicts of Interest “The authors declare no conflict of interest.”.

References

1. Manzo G (2020) Complex social networks are missing in the dominant COVID-19 epidemic models. *Sociological* 14(1):31–49
2. Roberts M, Andreasen V, Lloyd A, Pellis L (2015) Nine challenges for deterministic epidemic models. *Epidemics* 10:49–53
3. Merchant H (2020) CoVid-19 may not end as predicted by the SIR model. *The BMJ* 369:m1567-rr
4. Adam D (2020) The simulations driving the world’s response to COVID-19. How epidemiologists rushed to model the coronavirus pandemic. *Nature* 580:316–318. <https://doi.org/10.1038/d41586-020-01003-6>
5. Luo J (2020) Predictive Monitoring of COVID-19. SUTD Data-Driven Innovation Lab
6. Koerth M, Bronner L, Mithani J (2020) Why It’s so freaking hard to make a good COVID-19 model. *Abc News: FiveThirtyEight*. Retrieved from: <https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/>
7. Best R, Boice J (2020) Where the latest COVID-19 models think we’re headed—and why they disagree. *Abc News: FiveThirtyEight*. Retrieved from: <https://projects.fivethirtyeight.com/covid-forecasts/>
8. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J (2020) The reproductive number of COVID-19 is higher compared to the SARS coronavirus. *J Travel Med*
9. Wu JT, Leung K, Leung GM (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modeling study. *The Lancet* 395:689–697. [https://doi.org/10.1016/s0140-6736\(20\)30260-9](https://doi.org/10.1016/s0140-6736(20)30260-9). (2020)
10. Liu T, Hu J, Xiao J, He G, Kang M, Rong Z, Lin L, Zhong H, Huang Q, Deng A, Zeng W (2020) Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China. *bioRxiv*
11. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, Xing X (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England J Med*
12. Gu C, Jiang W, Zhao T, Zheng B (2020) Mathematical recommendations to fight against COVID-19. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3551006>
13. Wang L, Wu JT (2018) Characterizing the dynamics underlying the global spread of epidemics. *Nat Commun* 9. <https://doi.org/10.1038/s41467-017-02344-z>
14. Roques L, Klein EK, Papaix J, Sar A, Soubeyrand S (2020) Using early data to estimate the actual infection fatality ratio from COVID-19 in France. *Biology* 9(5):97

15. Tang B, Scarabel F, Bragazzi NL, McCarthy Z, Glazer M, Xiao Y, Heffernan JM, Asgary A, Ogdén NH, Wu J (2020) De-escalation by reversing the escalation with a stronger synergistic package of contact tracing, quarantine, isolation, and personal protection: feasibility of preventing a COVID-19 rebound in Ontario, Canada, as a case study. *Biology* (5):100
16. Griette Q, Zhihua L (2020) Estimating the last day for COVID-19 outbreak in mainland China. medRxiv
17. Liu Z, Magal P, Seydi O, Webb G (2020) A model to predict COVID-19 epidemics with applications to South Korea, Italy, and Spain. *SIAM News* (to appear)
18. Cotta RM, Naveira-Cotta CP, Magal P (2020) Modelling the COVID-19 epidemics in Brasil: Parametric identification and public health measures influence. medRxiv
19. Chen D (2014) Modeling the spread of infectious diseases: a review. *Analyzing and modeling spatial and temporal dynamics of infectious disease*, 19–42
20. Jackson MO, López-Pintado D (2013) Diffusion and contagion in networks with heterogeneous agents and homophily. *Netw Sci* (1):49–67
21. Chang SL, Harding N, Zachreson C, Cliff OM, Prokopenko M (2020) Modelling transmission and control of the COVID-19 pandemic in Australia. arXiv preprint [arXiv:2003.10218](https://arxiv.org/abs/2003.10218)
22. Rothan HA, Byrareddy SN (2020) The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* 26:102433
23. Vermeulen B, Pyka A, Müller M (2020) An agent-based policy laboratory for COVID-19 containment strategies
24. Tuite AR, Fisman DNR, Growth E, Numbers R, for the, (2019) Novel Coronavirus (2019-nCoV) Epidemic. *Ann Intern Med*. [https://doi.org/10.7326/m20-0358\(2020\)](https://doi.org/10.7326/m20-0358(2020))
25. Lipsitch M (2003) transmission dynamics and control of severe acute respiratory syndrome. *Science* 300:1966–1970. <https://doi.org/10.1126/science.1086616>
26. Ebrahim SH, Memish ZA (2020) COVID-19—the role of mass gatherings. *Travel Med Infectious Disease*, 101617. <https://doi.org/10.1016/j.tmaid.2020.101617> (2020)
27. Alves TH, de Souza TA, de Almeida S, Ramos NA, de Oliveira SV (2020) Underreporting of death by COVID-19 in Brazil's second most populous state. medRxiv
28. Casola AR, Kunes B, Cunningham A, Motley RJ (2021) Mask use during COVID-19: a social-ecological analysis. *Health Promot Pract* 22(2):152–155
29. Bai F, Brauer F (2021) The effect of face mask use on COVID-19 models. *Epidemiologia* 2(1):75–83
30. Block P, Hoffman M, Raabe IJ, Dowd JB, Rahal C, Kashyap R, Mills MC (2020) Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat Hum Behav* 4:1–9
31. Romano A, Sotis C, Dominioni G, Guidi S (2020) Covid-19 data: The logarithmic scale misinforms the public and affects policy preferences. Available at SSRN 3588511
32. Ferguson N, Laydon D, Nedjati GG, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunuba ZU, Cuomo-Dannenburg G, Dighe A (2020) Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand
33. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, El-harakeh A, Bognanni A, Lotfi T, Loeb M, Hajizadeh A (2020) Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet*
34. Konar A, Banerjee T, Roy A (2020) Detailed study of Covid-19 outbreak in India and West Bengal. *Res Rev Int J Multidisciplinary* 5(5). <https://doi.org/10.31305/rrijm.2020.v05.i05.010>
35. World Health Organization (WHO) Coronavirus disease 2019 (COVID-19) situation reports. Retrieved from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (2020)
36. Worldometer. COVID-19 Coronavirus pandemic. World Data. Retrieved from: <https://www.worldometers.info/coronavirus/> (2020)
37. John Hopkins University. Coronavirus Resource Center. Retrieved from: <https://coronavirus.jhu.edu/map.html> (2020)

38. Britton T, House T, Lloyd AL, Mollison D, Riley S, Trapman P (2015) Five challenges for stochastic epidemic models involving global transmission. *Epidemics* 1(10):54–57
39. Magal P, McCluskey CC, Webb GF (2010) Lyapunov functional and global asymptotic stability for an infection-age model. *Appl Anal* 89(7):1109–40. <https://doi.org/10.1080/00036810903208122>
40. Jiang R, Murthy DN (2011) A study of the Weibull shape parameter: properties and significance. *Reliab Eng Syst Saf* 96(12):1619–1626
41. Edelsbrunner H, Harer J (2008) Persistent homology—a survey. *Contemp Math* 453:257–282
42. Hua J, Shaw R (2020) Coronavirus (Covid-19) “infodemic” and emerging issues through a data lens: The case of China. *Int J Environ Res Public Health* 17(7):2309
43. Sungheetha A (2021) COVID-19 risk minimization decision making strategy using data-driven model. *J Inf Technol* 3(01):57–66. [https://doi.org/10.36548/jitdw.2021.1.006\(2021\)](https://doi.org/10.36548/jitdw.2021.1.006(2021))
44. Shakya S, Smys S (2021) Big data analytics for improved risk management and customer segregation in banking applications. *J ISMAC* 3(3):235–249. <https://doi.org/10.36548/jismac.2021.3.005>
45. Kumar R, Rajat K, Pinki K, Vishal K, Sanjay C, Sukhen D (2021) Prediction of protein–protein interaction as carcinogenic using deep learning techniques. In: *Proceedings of international conference on intelligent computing, information and control systems*, pp 461–475. Springer, Singapore. https://doi.org/10.1007/978-981-15-8443-5_39
46. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M (2020) Assessing the risks of “infodemics” in response to COVID-19 epidemics. arXiv preprint [arXiv:2004.03997](https://arxiv.org/abs/2004.03997)
47. International Labour Organization (ILO). COVID-19 and the world of work: Country policy responses. Geneva, Switzerland: ILO. Retrieved from <https://www.ilo.org/global/topics/coronavirus/country-responses/lang--en/index.htm> (2020)
48. International Monetary Fund (IMF) Policy Responses to COVID-19. Retrieved from <https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#A> (2020)

Performance Evaluation of an IoT Based Fetal Heart Monitoring Device



Olubunmi Ige, Adedotun Adetunla, Joshua Adewolu, and Adeyinka Adeoye

Abstract An important indicator to know the condition of fetal well-being is the fetal heart rate. The fetal heart rate variability signals during pregnancy is a biological index that must be well monitored, so as to assess the fetal well-being. This paper focuses on developing an IoT-based fetal heart monitoring device, there is accurate conformity between the newly developed device and the existing electronic stethoscope with <10% variation. The device was tested initially on the heart rate of an adult, the result was compared with the result obtained from using an exiting Doppler fetal monitor. From the result, it can be deduced that the relationship between the developed device and the standards used is linear. The device was then tested on pregnant women for detecting the heart rate of the fetus, the heart rate was monitored to determine the beats per minute. Hence, the result shows that the IoT-based monitoring device can detect the fetal heart rate as from 30 weeks upward (Third trimester).

Keywords IoT · Fetal heart rate · Embedded system · Biomedical device

1 Introduction

Cardiovascular disease (CVD) complicates 1–4% of pregnancies, with congenital heart disease (CHD) being the most common pre-existing condition. The incidence of maternal CVD appears to be growing, likely due to increasing maternal age, cardiovascular risk factors (i.e., obesity, diabetes, and hypertension) and lifespan of patients with CHD. Heart and circulatory system is one of earliest organ developed

O. Ige · A. Adetunla (✉) · J. Adewolu · A. Adeoye
Department of Mechanical and Mechatronics Engineering, Afe Babalola University, Ado-Ekiti,
Nigeria
e-mail: dotunadetunla@yahoo.com

O. Ige
e-mail: ige.olubunmi@abuad.edu.ng

A. Adeoye
e-mail: aadeoye@abuad.edu.ng

in the fetus which the fetal experienced the first heart beat by the 3rd week of life [1]. There are differences in term anatomy and physiology of fetal heart and newborn heart. During pregnancy, fetal heart circulation is different compare to newborn heart circulation. Oxygen is supplied to the fetal through placenta, so the heart only functioning to pump the oxygenated blood throughout the body including the lungs [2]. However, the lungs will supply oxygen to the newborn heart same as adult.

Fetal heart rate should be monitor during pregnancy which serves as an indicator for the well-being of the fetal [3]. Also another indicator to monitor the fetal well-being during pregnancy is by the monitoring the movement of the fetal [4]. However, there are a lot of shortcomings to monitor fetal movements. It is important to monitor because prevention can be done if there is a detection of abnormalities in fetal heart rate which lead to pre-maturity and miscarriage [5]. When a heart beats, it pumps oxygenated blood throughout the fetal body. Adequacy of fetal oxygenation is important to prevent hypoxia that affecting the whole fetal body. If hypoxia occurs, there might be a decrease in fetal cerebral blood flow. Therefore, fetal heart rate monitoring able to recognize fetal asphyxia [6]. Fetal asphyxia is severe enough that can lead to neurological damage or even fetal death. In addition, fetal heart rate monitoring can solve two problems, which are, it serve as a screening test for severe asphyxia and it allows recognition of early asphyxia so that timely obstetric intervention could avoid asphyxia-induced brain damage or death in the newborn [7].

Several studies have reported that numerous mortalities have occurred over the last five decades which is attributed to increasing Cardiovascular disease [8]. The World Health Organization has reported in 2004 that over 133 million births were recorded worldwide, and about 3.7 million died about four weeks or less after birth, while over 3 million were stillborn. Their report suggested that the highest death rates happened in sub-Saharan African countries, which have almost 60 deaths on every 1000 births, against the four deaths in 1000 births seen with the developed nations [9, 10]. A fetoscope has been chosen as a standard to monitor the fetal heart rate in most African countries, which has limited the assessment of the heart rate to one person at one point in time. This fetoscope has various challenges such as user complexity, inaccurate observation of heart rate, and misjudgment in cases where the heart rate cannot be heard. The use of the fetoscope also requires the observer to be positioned in an awkward and uncomfortable way while trying to observe the fetal heartbeat over the maternal heartbeat and other environmental noises as shown in Fig. 1.

Another created device is the Doppler Fetal Monitor which is chosen as a standard in developed nations, this device is very expensive and unaffordable for most hospitals in the developing nation, due to its use of ultrasound technology and there are concerns about the effects it may have on the foetus because of the waves it sends into the womb [11, 12]. It does not also have the functionalities of displaying graph waveform of the heart readings as well as the ability to view results from a remote location.

Fetal heart rate monitoring is an important parameter to determine the anomalies or distress of the fetal so that urgent attention can be given to any abnormalities observed during the pregnancy period. There is a need for a device that can effectively monitor

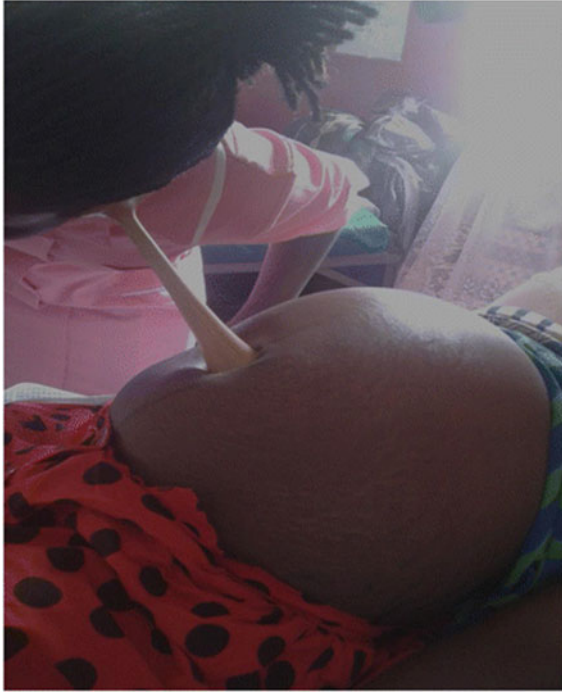


Fig. 1 Midwife using a Pinard stethoscope to monitor the heart rate of a fetus

the fetal heart rate, capture the data and store it online using cloud-based technology so the data can be accessed from a remote location. The aim of this work is to develop a fetal heart online monitoring device which will capture the heart rate data, store it online and the graph form display on a mobile app which can be accessed by the doctor or specified user in a remote location in real time as well as give out sound output. This schematic for this developed device will be present and explained in detailed.

2 Methodology

2.1 Material Selection

The materials used for the development of the IoT based fetal heartrate monitoring device are;condenser microphone, which serves as a sound sensor. Some resistors, which helps to control the flow of current, IN4001 that helps to prevent backflow of current, Hollow cone for the acoustic amplifier, the Audio amplifier used is TBA800, NodeMCU that serves as the microcontroller with wifi connectivity, some capacitors

that helps to store energy. The display of the output is done on an Android phone, another material used is the switch that helps to open and close the circuit. An LED was employed as an indicator, while a battery is used to supply voltage and current to the circuit [13]. The assembly is shown in Fig. 2.

Figure 3 shows the block diagram of the method adopted in the design and development of the device. The overall circuit required between 4.5 and 5 V so a 9 V was stepped down to 4.5 V using voltage divider. From the condenser microphone to the Audio amplifier and to the RC filter, which is subsequently sent to the NodeMCU that serves as the microcontroller with wifi connectivity, some capacitors that helps to store energy.

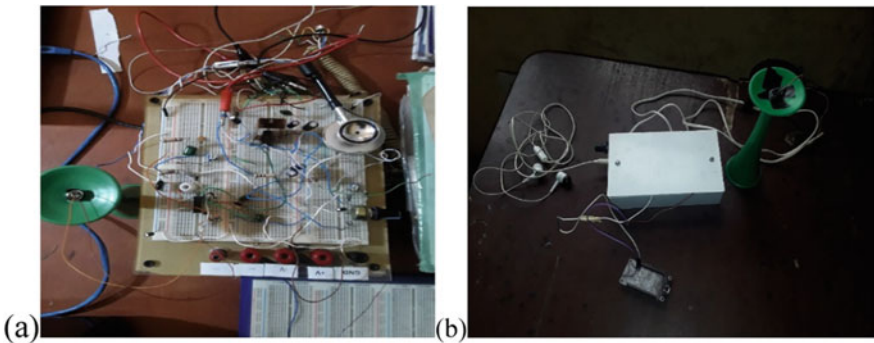


Fig. 2 Assembly of the IoT-based fetal heartrate monitoring device

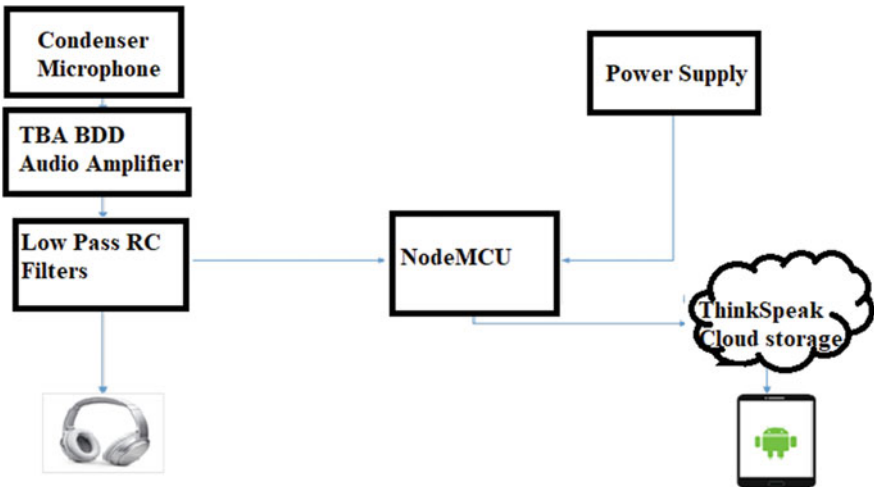


Fig. 3 Block diagram of the design

2.2 Signal Acquisition

The fetal heart rate (fHR) signal is extracted by fetal phonocardiography (fPCG). This is a cheap and non-invasive method of extracting fHR. The heart’s mechanical activity is accompanied by the generation of a variety of characteristic sounds. These sounds are associated with changes in the speed of blood flow, as well as with the opening and closing of heart valves and could provide diagnostic information, accordingly [14, 15]. For detecting the fetal heartbeat, the condenser Microphone was used as shown in Fig. 4.

The condenser microphone will require an acoustic amplifier to help identify and isolate the faint sounding heartbeat, the Hollow cone was chosen for the acoustic amplification. For amplifying the electric signal generated from the condenser microphone, TBA800 audio amplifier was selected due to its use as a low frequency Class B amplifier. A TBA800 IC and its pin configuration are shown in Fig. 5a and b. Pin 1 is the Supply Voltage Pin, Pin 4 is the Voltage Feedback Pin, Pins 5, 7, 9, 10, 11 are grounded, while Pin 8 is the Input Signal Pin, and Pin 12 is the Output Signal Pin.



Fig. 4 Condenser microphone

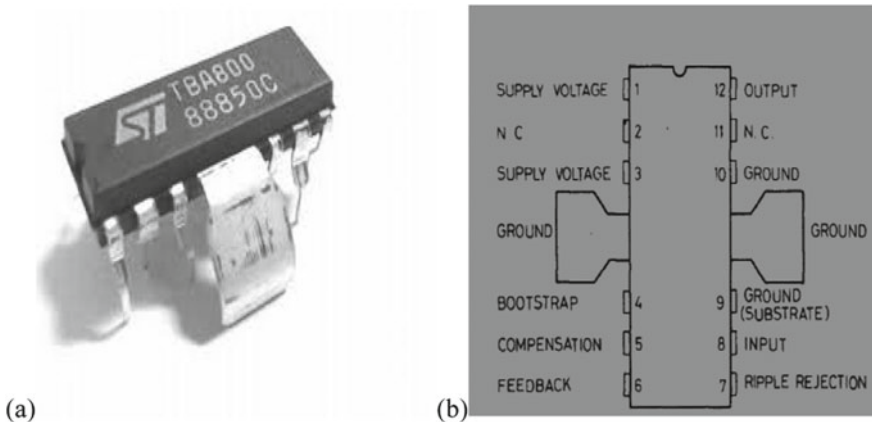


Fig. 5 a TBA800 IC, b TBA800 IC pin configuration

The amplified signal from the TBA800 is filtered using a low pass RC filter. The capacitor required to filter out background noises was determined using the formula

$$\text{Frequency Cut - off, } f = \frac{1}{2\pi RC}$$

Where, R is the resistor value in ohms, C is the capacitance value in Farads The frequency range for the fetal heartbeat is 35–200 Hz, therefore

Given,

$$f_c = 200 \text{ Hz}$$

$$R = 1500 \Omega$$

$$C = \frac{1}{2\pi f_c R}$$

$$C = \frac{1}{2\pi} * 1500 * 200$$

$$C = 531 \text{ nF}$$

Hence the value of capacitor required with a 1.5 k Ω resistor to attenuate signals above 200 Hz is 531nF but due to unavailability of the capacitor, a 470nF capacitor was used.

2.3 Digital Analysis

For digital analysis of the signal, the signal cable from the circuit is connected to analog pin A0 on the NodeMCU while the ground cable is connected to Ground pin GND. The processed data is sent to the specified Thing-Speak API through the ESP 8266 12E wifi connectivity on the NodeMCU.

The waveform of the readings of the heartrate was viewed using Keysight U8903A Audio Analyzer as shown in Fig. 6, and the waveform on the audio analyser is shown in Fig. 7.

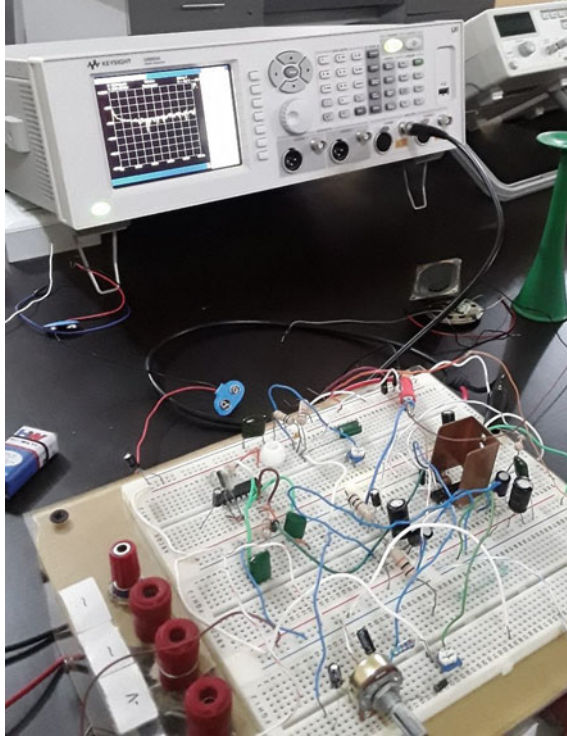


Fig. 6 Waveform display on keysight U8903A audio analyser

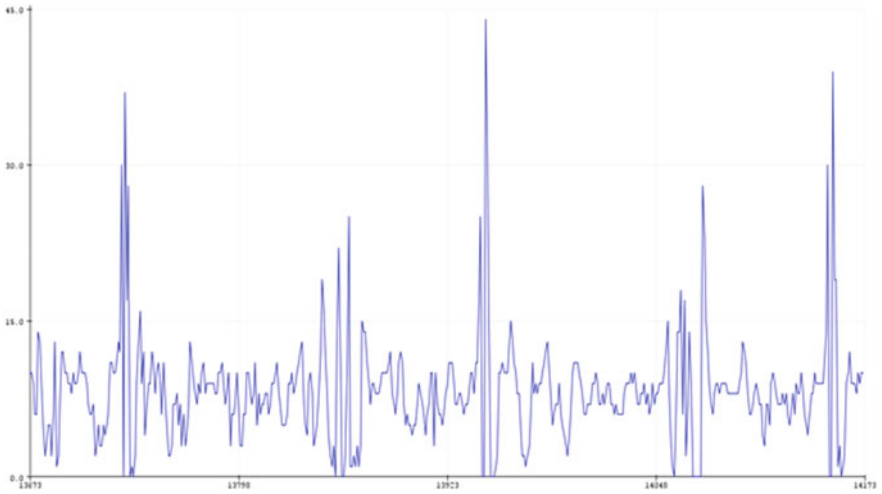


Fig. 7 Waveform display on Arduino IDE serial plotter

3 Results and Findings

The results obtained after the successful development of the IoT-based fetal monitoring device are discussed as follow.

3.1 Detection of an Adult Heart Rate

The device was tested initially for detecting the heartrate of an adult. The heart rate was monitored to determine the beats per minute. The result was compared with the result obtained from using S-Health app on Samsung Galaxy S6. Table 1 presents the values obtained from determination of the beats per minute of an Adult using both devices.

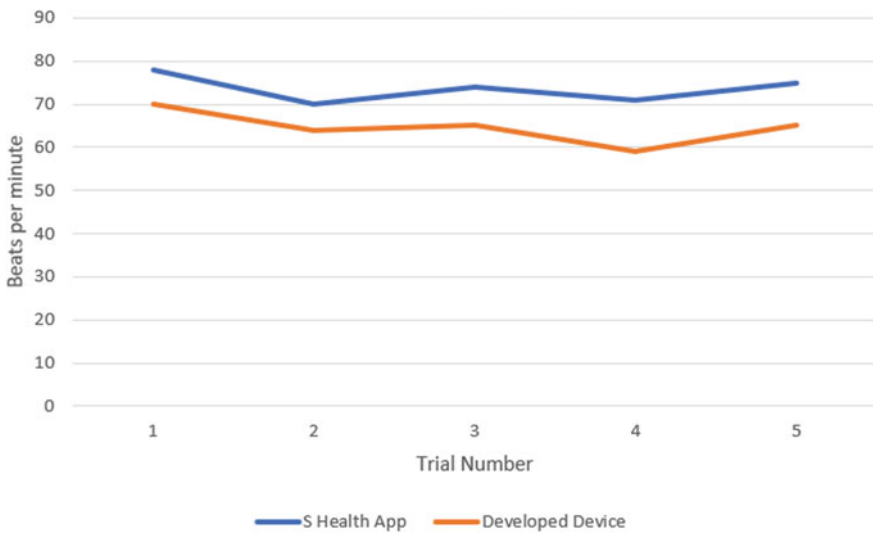


Fig. 8 Graph of adult heartrate readings

Table 1 Comparison of results obtained from the S-Health app and the developed device

| Trial number | S-Health app (bpm) | Developed device (bpm) |
|--------------|--------------------|------------------------|
| 1 | 78 | 70 |
| 2 | 70 | 64 |
| 3 | 74 | 65 |
| 4 | 71 | 59 |
| 5 | 75 | 65 |

Table 2 Comparison of between the Doppler fetal monitor and the developed device

| Trial number | Doppler fetal monitor (bpm) | Developed device (bpm) |
|--------------|-----------------------------|------------------------|
| 1 | 130 | 155 |
| 2 | 135 | 164 |
| 3 | 155 | 200 |

In Fig. 8, a graph of the readings gotten from the S-Health app on Samsung Galaxy S6 and the developed device was plotted for visual analysis. It can be deduced that the relationship between the developed device and the standards used are linear.

Mean Fetal Heartrate of the Developed Device

$$= (155 + 164 + 200)/3 = 173 \text{ bpm}$$

Mean Fetal Heartrate of existing device = $\frac{130 + 135 + 200}{3}$

$$= 140 \text{ bpm}$$

Average Difference = $\pm(\text{mean of developed device} - \text{mean of existing device})/3$

$$= (173 - 140)/3 = \pm 33 \text{ bpm}$$

%Error = $\pm(\text{true value} - \text{measured value})/\text{true value} * 100$

$$= \pm(140 - 173)/140 * 100 = 23.57\%$$

3.2 Detection of Fetal Heart Rate

The device was then tested on pregnant women for detecting the heartrate of the fetus. The heart rate was monitored to determine the beats per minute. The results were compared with the results obtained from using a Doppler fetal monitor [16]. Table 2 presents the values obtained from determination of the beats per minute of the fetus using both devices.

A graph of the readings gotten from the Doppler fetal monitor and the developed device was plotted for visual analysis as shown in Fig. 9. Similar observation was seen with the fetal heart rate monitoring device. It was observed that the developed device gave out a slightly better sound output of the fetal heart than a pinard stethoscope. Also, the developed device could only detect fetal heart rates in advanced pregnancy stages (as from 30 weeks upward) but couldn't detect early pregnancy stages.

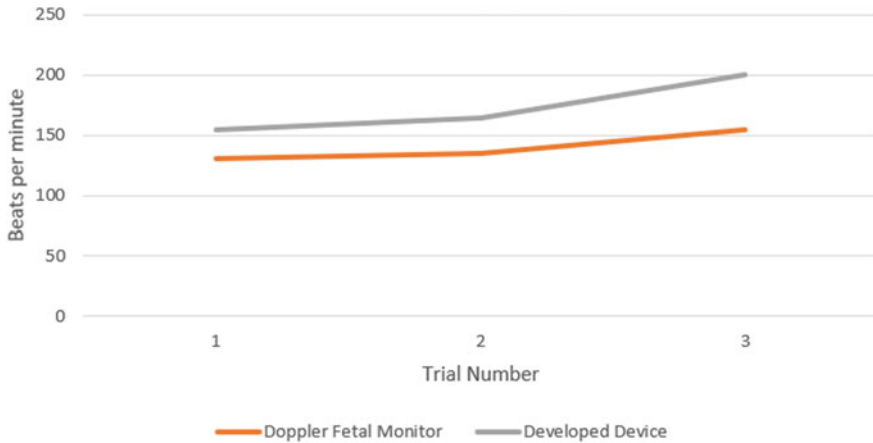


Fig. 9 Graph of fetal heartrate readings

4 Conclusion

The developed device shows a non-invasive fetal heartrate monitor that can efficiently measure the heartrate of the fetus in real time. The device gives out sound output as well as store the captured data online making it accessible from a remote location. It can be seen that the relationship between the output voltages from condenser microphone is directly proportional to the heart rate of the fetus. When the fetal heart rate increases, the output voltage from the condenser microphone also increases. The result shows that IoT-based monitoring device can detect the fetal heart rate in advanced pregnancy stages (as from 30 weeks upward).

References

1. Seshadri DR, Tyson K (2019) Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *npj Digit Med* 2(1)
2. Liu C, Liu F, Zhang L, Su Y, Murray A (2017) Smart wearables in healthcare: signal processing, device development, and clinical applications. *J Healthc Eng* 2018(September):2018
3. Zhang T (2016) Biomechanical design and control of an anthropomorphic artificial hand for prosthetic applications. *Robotica* 34(10):2291–2308
4. Xing D, Yeh CI, Shapley RM (2009) Spatial spread of the local field potential and its laminar variation in visual cortex. *J Neurosci* 29(37):11540–11549
5. Deshpande AD et al (2013) Mechanisms of the anatomically correct testbed hand. *IEEE/ASME Trans Mechatron* 18(1):238–250
6. Bibbo D et al (2020) A new approach for testing fetal heart rate monitors. *Sensors (Switzerland)* 20(15):1–15
7. Macones GA (2009) Intrapartum fetal heart rate monitoring: nomenclature, interpretation, and general management principles. *Obstet Gynecol* 114(1):192–202

8. Roedel K, Roland R (2020) Mechanical ventilation and mortality among 223 critically ill patients with coronavirus disease 2019: a multicentric study in Germany. *Aust Crit Care* 34(2):167–175
9. Datta S, Kodali BS, Segal S (2010) *Obstetric anesthesia handbook*. *Obstet Anesth Handb*, June 2014
10. Akande S, Adetunla A, Olanrewaju T, Adeoye A (2021) UAV and its approach in oil and gas pipeline leakage detection, vol 2021
11. Jones LC, Topoleski LDT, Tsao AK (2017) *Biomaterials in orthopaedic implants*. Elsevier Ltd.
12. Adetunla A, Akinlabi E (2019) Improving the biocompatibility and corrosion resistance of AZ31 Mg alloy for biomedical applications. *MS T 2019—Mater Sci Technol* 2019:291–297
13. Sathesh A (2020) Computer vision on IOT based patient preference management system, vol 2, no 2, pp 68–77
14. Balasubramaniam V (2020) IoT based biotelemetry for smart health care monitoring system, vol 2, no 3, pp 183–190
15. Aheleroff S et al (2020) Advanced engineering informatics IoT-enabled smart appliances under industry 4.0: a case study, vol 43, no February, 2020
16. Vijayakumar T (2021) Fusion based feature extraction analysis of ECG signal interpretation—a systematic approach, vol 3, no 1, pp 1–16

A Novel Approach for Handling Imbalanced Data in Breast Cancer Dataset



Nagateja Banothu and M. Prabu

Abstract Breast cancer is a deadliest disease which is the second most leading cause of death in the world. There are 5 types of commonly occurring cancers which include Ductal carcinoma in situ (DCIS), Invasive ductal carcinoma (IDC), Tubular carcinoma, Medullary carcinoma, and Invasive lobular carcinoma (ILC). Early detection of breast cancer creates 93% or more chances of survival. So in order to detect breast cancer, one of the important morphological feature is mitotic count—showing the rate of division of cell. In a dataset, if one class have less number of examples compared to number of examples in the other class, then the dataset is imbalanced. Since there is much abundance of non-mitotic cells over mitotic cells, the dataset has class imbalance problem. This class skewness leads to inefficiency in classification task. At present there are many models which uses data sampling methods in order to decrease the class imbalance problem. But if there is a feature set that provides the exact decision boundary, then there is no need of sampling methods for solving the class imbalance issue. Therefore in this work a model is proposed which uses autoencoders to learn a feature set which has better classification capabilities of majority class and minority class to address the class imbalance problem. In this paper, Dual stacked autoencoder is applied on imbalanced mitotic and non-mitotic dataset, which showed improvement in f-measure and also reduces the overhead of smapling the dataset. The proposed dual stack autoencoder shown f-measure of 0.82, and outperformed many of the traditional classification methods.

Keywords Breast cancer · Early detection · Mitotic count · Biopsy · Class imbalance

N. Banothu · M. Prabu (✉)

Intelligent Computing Lab, Department of Computer Science and Engineering, National Institute of Technology Calicut, Kozhikode 673601, Kerala, India

e-mail: prabum@nitc.ac.in

N. Banothu

e-mail: banothu_m200288cs@nitc.ac.in

1 Introduction

Cancer is a major issue worldwide and is a serious public health concern [18]. Every year around the world 1.68 million women are affected by breast cancer and are diagnosed by it and nearly 0.5 million die due to it [7]. In 2020, 2.3 million are diagnosed and 0.68 million deaths are reported globally. Diagnostic tools like ultrasound, biopsy and MRI are used to detect Cancer [23]. In biopsy, affected tissues from the tumour area are collected and a small slice is extracted, and these slices are treated with Hematoxylin(H) and Eosin(E) where nuclei are turned to blue because of Hematoxylin and entire cytoplasm is turned to red by Eosin [7, 20]. This biopsy is used by pathologists to diagnose disease. Mitosis count in the tumour area is one of the most significant morphological feature to detect the aggressivity of breast cancer [7]. Since the rare occurrence of mitotic cells over non-mitotic cells, there will be huge skewness in the dataset which creates Class imbalance problem [4, 7].

Class Imbalance is most complicating factor for the machine learning classification methods. Most machine learning methods will perform badly when the dataset is unevenly distributed and require modifications in order to avoid predicting majority class in each and every case [10].

Class Imbalance occurs when there is uneven instances in class distribution [1]. Class which has less number of instances is the positive (minority) class, and the class with more number of instances is negative (majority) class. More often the minority is one that is of more interest and more importance [11].

When rare examples which do not frequently occur, they will be predicted as rare occurrences which makes a minority class. Most real world dataset consists of many normal instances and very small fraction of abnormal instances in it. Class imbalance distribution can arise in two situations: (1) imbalance occurs naturally (intrinsic), due to few occurrence of some events. In case of credit card fraud detection, the fraudulent transactions are <1% of total transactions, similarly in disease detection very less data available of infected animal than healthy animal [4, 7, 11]. (2) When the dataset is balanced naturally, every class can have equal distribution, but in some cases it is very difficult to get number of instances for minority class due to confidentiality, cost and huge effort to find well distributed dataset [11].

Ratio between the instances of minority class and majority class defines the degree of imbalance. In recent years there is a lot of research in class imbalance classification. Class imbalance problem present in many real life scenarios like fault diagnosis, anomaly detection, medical domain, detection of oil spillage using satellite captured images, text classification [7].

Most existing classification methods perform better when the number of instances in each class are nearly equal. But the problem arises when instances in one class far exceed the other class. Main concern with class imbalance is that most algorithms measure their performance using accuracy. There are many real life applications with class imbalance distribution and lot of interest kept in studying minority class. However, in the case of class imbalance, classification accuracy reveals little about the minority class [11]. Because of many classification methods don't take dataset

distribution into consideration results in inaccurate learning model and deterioration of performance [4, 7, 9].

For almost every disease, a medical laboratory has less number of patients who are having disease than those who are not having the disease. As the number of instances collected from diseased patients is far less than the number of instances which are from normal patients, this creates imbalance in medical domain [1, 1].

The rest of the paper is organised as follows, Sect. 2 describes problem definition, Sect. 3 discusses literature review, Sect. 4 explains proposed methodology, Sect. 5 discusses the result obtained and finally conclusions are made in Sect. 6.

2 Problem Definition

Computer aided diagnosis is an emerging research area in medical imaging, where the main focus is to draw a distinction between malignant and normal. After the emergence of machine learning and deep learning models, medical image classification has shown promising progress. Usually, deep learning models require a huge dataset for better performance.

Data collection in the case of the medical domain results in imbalance in the classes due to rare occurrence of diseased over non-diseased. So, a lot of medical datasets are biased towards the class which has more number of examples, resulting in an imbalanced dataset.

In the case of a balanced dataset many models show promising results. Performance of the model degrades in the case of an imbalanced dataset. Now, we propose a model named two phase learning where phase one is used for balancing the data and phase-2 for classification on balanced data. Our problem is to build a model to mitigate the class imbalance problem while classifying mitotic and non-mitotic nuclei in given histopathology images of breast cancer.

In order to classify mitotic cells from non-mitotic cells first Segment given image of tissue into separate images of nuclei. This will produce an imbalanced dataset of mitotic and non-mitotic nuclei. Imbalance in dataset will degrade the performance of any machine learning model. But to cope up with this class imbalance problem currently So, Reduce the imbalance in the dataset. This will give modified dataset which is nearly balanced. Build and train a model with balanced dataset. Classify the nuclei either as mitotic or non-mitotic using the trained model.

3 Literature Survey

To deal with imbalanced dataset, several strategies have been introduced. They can be classified into three categories:

- (i) Data Based Methods
- (ii) Model Based Methods
- (iii) Hybrid Methods

- (i) **Data Based Methods:** Methods in Data Based approaches modifies the training set to achieve balanced distribution among the classes. Data Based Methods uses pre-processing level to rebalance the data. Data Based methods are also known as sampling methods. Data Based techniques uses two approaches in order to achieve balanced dataset: (1) Oversampling, (2) Undersampling. This can be achieved by undersampling or oversampling or both [6, 7].

Oversampling is applied in order to increase the examples in minority class by creating affine transformation copies or by generating new images. Synthetic Minority Oversampling Technique (SMOTE) is widely used to generate examples that are similar to the original examples. On medical images SMOTE cannot be applied because of sensitive nature of medical images [11, 21].

Undersampling decreases the number of examples from the majority class by removing the examples which don't contain important information. Randomly removing the instances from the dataset may result in information loss, therefore some other approaches are used such as EasyEnsemble, and Clustering Methods. Compared to random sampling, clustering based approaches performed well [11].

- (ii) **Model Based methods:** Approches in model based methods tweak the learning mechanism of the neural network. There are some successful methods in this types. Two significant types are adaboost and Ensemblance methods. Adaboost (adaptive boosting), in which different weights to classes to change the weight update mechanism [8, 9]. Ensemblance, in which many classifiers are trained on training data and combined in order to produce a optimal model [11].
- (iii) **Hybrid methods:** Here both Data based method and Model based method are combined. Sampling is used along with a model based approach to compensate for the imbalanced dataset. SMOTEBoost and databoost are some significant methods [7–9, 22].

Mitotic and Non-mitotic Cells

Mitosis is the cells that are in dividing phase. One cell divides into two identical daughter cells. Even though cells in every human divides but the tumor tissue divides more rapidly [19]. In Fig. 1 top row images shows Mitotic Images, Bottom row images are non-mitotic images.

Where as non mitotic cells are the cells which in normal form with out any division.

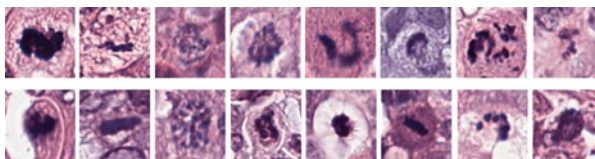


Fig. 1 Mitoses (top row), non-mitoses (bottom row)

Almost all works on handling the class imbalance problem use one of above three approaches or combining them, each of them have their own advantages and disadvantages [13]. In Sampling methods, overfitting problem may arise. Model based methods may need some preprocessing, to find hard examples and concentrate more on them. In case of hybrid method, deciding the good design by combining two or more models is hectic task [11, 14].

Gao et al. [1] uses sampling methods to balance the dataset initially and the applying CNN on it. Similar to above model Krawczyk et al. [10] have applied sampling techniques, for oversampling, flipping and rotating the minority class examples and for undersampling they used k-means clustering algorithm model given in Reza et al. [4].

Wahab et al. [7] uses SMOTEBoost to generate minority class samples in the dataset to balance it. After balancing there can be a problem of overfitting. In order to solve above issue Ng et al. [2] used Biased SVM in addition to SMOTEBoost technique, this have shown improvement in the performance.

Hartono et al. [5] used one class classification using SVM, in one class classification only one class is trained and all other classes are considered as outliers.

Similar to above model Bader-El-Den et al. [8] and Chawla et al. [9] instead of SVM the author used CNN and autoencoder respectively. Out of all above approaches Autoencoder performed as they are good feature extractors (Table 1).

Authors mention recent works have applied mostly hybrid methods which include ADASYN [1, 6], SMOTEBoost [2, 7, 10], and model based methods which includes OCSVM [5] (One class support vector machine), Autoencoders [9], OCC [8], Biased Random forest [3]. ADASYN [6] method generates the synthetic data of minority class and advantages of is that adasyn concentrates much on generating examples which are harder to learn, and also in order to learn hard examples it adaptively shifts its boundary [6]. SMOTEBoost also generates synthetic examples which results in updating the weights for minority classes [2]. In One class Classification (OCC) the model learns only the examples which belongs to majority class and take the examples of minority class as outliers [1]. In OCSVM [5] (One class support vector machine), SVM obtain the one class learning by creating a closed boundary hypersphere enclosing all the examples of majority class [16]. When OCSVM comes across the minority class examples then it will classify them as outliers [10, 12].

Autoencoders learns majority class and its feature set provides clear decision boundary between majority and minority classes [2]. Biased Random forest identifies critical areas in the dataset using nearest neighbour algorithm, then this critical areas are fed with more random trees. Author [1] in ADASYN model mentioned that there will be lose in information due to randomly undersampling which will cause degradation in the performance. In case of OCC and OCSVM results are better but it also has limitations when the dataset is perdominatly consists of outliers [4]. Stacked autoencoders have shown good results but it is difficult to decide number of autoencoders should be kept per stack [2].

Author [3] mentioned that Biased Random forest shown better results when f-measure is used as performance metric. But when G-mean is used as performance metric Random forest performed better than Biased Random forest method [8].

Table 1 Comparing all approaches

| Sl no. | References | ML approach | Dataset used | Classification | Limitations |
|--------|-------------------------|------------------------------|--|--|--|
| 1 | Reza and Ma [4] | CNN, SMOTE, ADASYN | 162 Breast cancer WSI scanned at 40x | Undersampling Majority class and oversampling minority class and then apply CNN for classification | Undersampling the majority dataset may lose important information |
| 2 | Hartono et al. [5] | SMOTEBoost, Biased SVM | Iris dataset | Biased SVM and generating synthetic images using smoteBoost | Less accuracy compared to normal SVM |
| 3 | Bader-El-Den et al. [8] | Biased Random Forest | KEEL datasets (286 examples) | More decision trees are applied to learn difficult areas in the training dataset | Using F-measure as performance metric BARF performed better, but while using G-mean as performance metric random forest performed better than BARF |
| 4 | Lin et al. [3] | k-means clustering algorithm | KDD cup 2008 breast cancer dataset | Clustering-based undersampling on the majority dataset to reduce the class skewness | Due to undersampling we may lose important data |
| 5 | Krawczyk et al. [10] | OCSVM | 10 imbalanced benchmark datasets | One class classification using SVM is applied on minority class | Proposed method has difficulties when the dataset contains outliers |
| 6 | He [6] | ADASYN | pima Indian diabetes dataset | Reduces the class skewness and adaptively shifts the boundary to harder to learn examples | Cannot deal with multi-class imbalance problem |
| 7 | Chawla [9] | SMOTEBoost | KDDCup-99 Intrusion, Mammography dataset | Reduces class skewness by generating synthetic images of minority class | Performance of SMOTEBoost degrades in the presence of noise |

(continued)

Table 1 (continued)

| Sl no. | References | ML approach | Dataset used | Classification | Limitations |
|--------|------------------|---------------|-------------------------------|--|--|
| 8 | Gao et al. [1] | ICOCC | MRI dataset–1946 tumor images | Concept of image-complexity is used to optimally learn the single class | Images with high resolution containing local anomalies cannot be detected |
| 9 | Ng et al. [2] | Autoencoder | 14 UCI datasets | Two stacked autoencoders are used to learn different type of features using different activation functions on the training dataset | Deciding number of autoencoders per stack is new challenge |
| 10 | Wahab et al. [7] | Two phase CNN | MITOS12 | Phase-1 is to balance the imbalanced dataset and then Phase-2 is to train the balanced dataset | It may experience an overfitting problem and while undersampling then there may be chances of losing important information and oversampling may repeat the same data |

Author [10] presented a model named Two phase Training to mitigate the class imbalance and classify the mitotic and non-mitotic nuclei. This method contains to phases. The main aim of Phase-1 training is to Categorize the non-mitotic data into easy, medium and hard using the classifier’s confidence (Output score). After categorizing the the non mitotic data, easy examples are reduced using Blueratio histogram based k-means clustering, hard are augmented by flipping and rotating and medium category examples remains same. On the other hand, the mitotic data is oversampled by rotation and flipping. Using this information a modified and relatively balanced training set was prepared for Phase-2.

4 Proposed Methodology

The two phase learning model achieved a F-score of 0.72 whereas Stacked Dual autoencoder given good results and two phase learning may experience an overfitting problem and while Undersampling then there may be chances of losing important

information and oversampling may repeat same data. Many classification models were applied after balancing the dataset, using sampling methods. But sampling methods also have their own disadvantages.

Sampling techniques are not applied in Stacked Dual autoencoders [2] which reduces the problem of losing data or creating redundant data.

4.1 Classification Using Autoencoder

Class Imbalance problem in a dataset occurs in a rare event problem. Which means, number of examples in one class is very less than the number of examples in another class. Class having less number of examples is positive class, and the class with more number of examples is negative class.

Any medical laboratory, getting the samples of diseased person is rare event as it occurs in less number of instances than those who are not having the disease. So this rare event causes the class imbalance which is observed largely in medical datasets.

Most of the machine learning classification methods will degrades their performance when the classes unevenly distributed. Classifying the events that are occurring rarely is challenging.

Effectiveness of Autoencoders: Autoencoders are powerful feature extractor. Autoencoders are unsupervised learning algorithm, which are capable of learning efficient representation of input data, this representation is also called as codings.

Codings is a lower dimensionality representation of input data, which makes autoencoders useful for dimensionality reduction. More than one autoencoder stacked together to get deeper insights of the features which helps to get better representation of patterns. Stacked autoencoder are connected in such a way that result of the first encoding part of an encoder is provided to next stacked autoencoder as their input [15, 17].

Framework of Dual Autoencoding Features (DAF): In this framework, features that are learned from two stacked autoencoder are combined to form dual autoencoding features. Two stacked encoders are trained using different activation functions. One autoencoder uses sigmoid activation function and other autoencoder uses tanh activation function as shown in Figs. 2 and 3.

Standard backpropagation algorithm is used for training. All the weights and biases are made to zero at the beginning. Input data is then fed to the autoencoder, by using decoding layer gradients for the weights and bias gets updated.

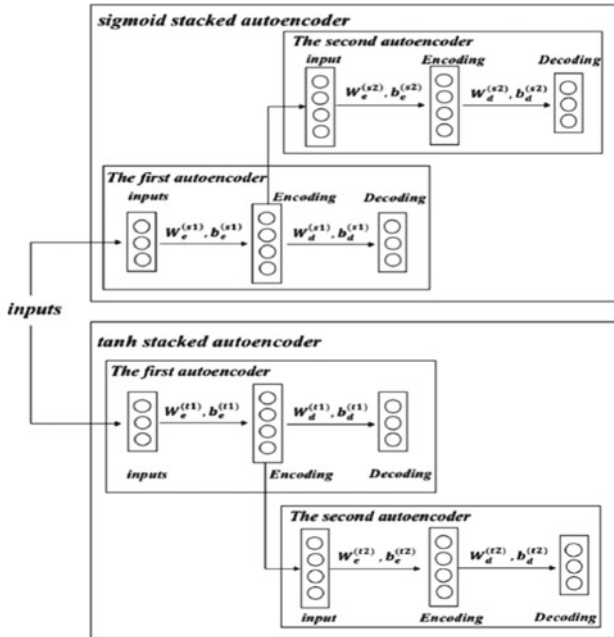


Fig. 2 Feature learning phase

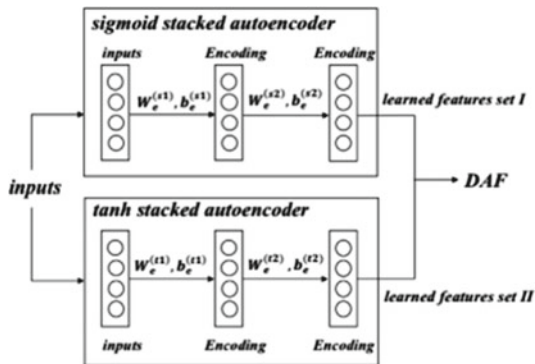


Fig. 3 Feature encoding phase

4.2 Dataset

In this work MITOS12 dataset which is provided by ICPR2012 is used. MITOS12 dataset [7] composed of 5 slices of biopsy slides from 5 patients were collected where 10 HPFs of 40X magnification were made for each slide and each slide is being analyzed by the professional pathologist to point out the mitotic cell and they have

Table 2 Candidates detection after segmentation

| Dataset | Total blobs detected | Mitoses | Non-mitoses |
|----------|----------------------|---------|-------------|
| Training | 16,182 | 223 | 15,959 |
| Testing | 6897 | 100 | 6797 |

counted a total of more than 300 mitotic cells. So 70% of the segmented dataset is considered as training data and 30% of the segmented dataset is considered as testing and each HPF is of size $512 \text{ um} \times 512 \text{ um}$ and these 50 images are collected from 3 scanners H, scanner A and multi-spectral microscope. We use Scanner A dataset for this problem which produces 2048×2048 pixel image from given HPFs.

Scanner A dataset contains RGB images of 2084×2084 pixels.

Scanner H dataset contains RGB images of 2252×2250 pixels.

Training Dataset: Each HPF image is provided with CSV file, this file indicates the single mitosis region. In the CSV file pixel coordinates of are provided (x, y) format. One line in csv file provides coordinates of all pixel values belonging same mitosis. Pixel origin starts from top of the image (Table 2).

4.3 Design

In this work the dataset ICPR where it contains about 50 images of histology slides with ground truths of only the regions of mitotic cells are annotated. First while preprocessing each individual cell from the histology image is being segmented and Then preprocessed the data for the classification and finally train with dual autoencoder to get latent feature extraction for classification.

Data Preprocessing: In the dataset, they provided HPF's (magnified patches in H&E slices) and the x, y coordinates of mitotic cells in respective csv files. To detect whether a particular nuclei is mitotic or not, both mitotic and non mitotic nuclei are needed, so each every nuclei from image should be segmented and made into separate images for further analysis. The Segmentation process is done in the following process:

- A blue-ratio histogram of the image is created. This histogram highlights the nuclei as the nuclei have high blue ratio.
- Then the image is modified using binary global thresholding to segment out individual nuclei.
- Then blobs are detected in the image using blob detection techniques, and the centroids of the blobs are found. And we get patches of 80×80 pixels.
- The dataset created by segmentation is then labelled by labelling the patches closest to the given coordinates of mitotic cells as mitotic class, and the rest as non-mitotic.

Segmenting and labelling patches produces a new dataset which is highly imbalanced due to very less number of mitotic cells (1%) compared to non mitotic cells (99%). Due to the imbalanced dataset the performance of any machine learning classifier degrades.

The data set has 50 images in total and each image has a corresponding csv file of annotated mitotic cells. First Segment out the nuclei from the image using following methods.

Blob Detection: The Histology Images collected from the biopsy were treated with the H & E chemicals where Hematoxylin makes the nuclei blue due to interaction with the nuclei. So blue ratio histogram is calculated using the below formula. Applying Blue Ratio Histogram on cancer slide produces the image as shown in Fig. 4.

$$BR = (100 * B)/(1 + R + G) * 256/(1 + R + G + B) \quad (1)$$

The image needs to be enhanced hence Otsu thresholding is used to threshold the images.

Otsu's Method for Global Binary Thresholding

By Using Binary Thresholding algorithm pixels of nuclei is separated from background pixels.

Global Binary thresholding transforms the images in which if the intensity of pixels is higher than a threshold T then the pixels are changed to white, if the intensity is less than the threshold T then the pixels are changed to black.

Otsu's thresholding method is used to optimally get the threshold value T. This threshold value divides foreground and background in this case it is cytoplasm and Nucleus. Applying Otsu's thresholding on blueratio image produce the image shown in Fig. 5.

Dual Autoencoders Training: After getting the dataset of nuclei which is extremely imbalanced it is fed as input to the two stacked autoencoders. In this model two

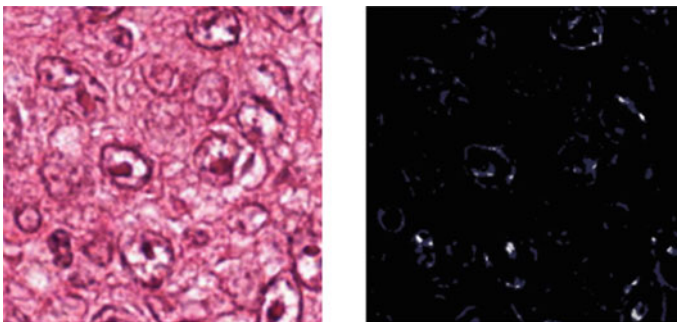


Fig. 4 Left image: original image, Right image: After apply Blue Ratio Histogram on original image

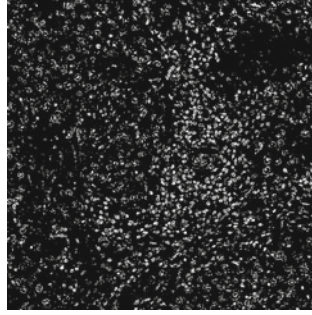


Fig. 5 Global binary thresholding

stacked autoencoders with different activation functions are used to learn two different set features and these features sets are combined to form a Dual Autoencoding Features. Then the samples are classified using this new feature space instead of original input space.

5 Results and Discussion

In this model, each autoencoder is applied on different activation function, (i.e. sigmoid and Tanh) this two activation functions helps in learning different features, and both of the feature sets are combined together to form Dual Autoencoding Features (DAF) as shown in Fig. 3.

If accuracy is used as an evaluation metric then the result can be easily fooled. As this model produced 90% accuracy, it classifies everything as non-mitotic cells. So instead of accuracy, here F-measure metric is used.

$$f\text{-measure} = \frac{2 * \textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}} \quad (2)$$

F-Measure is harmonic mean of recall and precision where precision is:

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

and Recall is

$$\textit{Recall} = \frac{TP}{TP + FN} \quad (4)$$

TP (true positive) = Total positives correctly predicted as positives

Table 3 Results

| Method | TP | FP | FN | Precision | Recall | F-score |
|-----------------------|----|----|----|-----------|--------|---------|
| Baseline-CNN [7] | 59 | 42 | 41 | 0.58 | 0.59 | 0.59 |
| Random-CNN [7] | 67 | 27 | 33 | 0.77 | 0.67 | 0.69 |
| Balanced-CNN [7] | 76 | 16 | 24 | 0.83 | 0.76 | 0.79 |
| DAF [proposed method] | 78 | 14 | 24 | 0.89 | 0.78 | 0.82 |

FP (false positive) = Total negatives wrongly predicted as positives

FN (false negative) = Total positives wrongly predicted as negatives

TN (true negative) = Total negatives correctly predicted as negatives

Here TP is total mitoses predicted as mitoses. Count is done only when the predicted centroids found within 5 mn from given ground truth labelled centroids.

FP is the total number of non-mitotic predicted as mitotic by the model. FN is the total number of missed mitotic count. TN is the total number of non-mitoses predicted as non-mitotic by model.

Dual stacked autoencoder has shown improvement in the result compared to other traditional methods based on CNN and Balanced-CNN [7]. Baseline CNN and Random CNN has very low performance where as Balanced-CNN has f-measure of 0.79, But to apply CNN on the dataset, Balanced dataset is required. For creating balanced dataset, sampling methods are applied on imbalanced dataset which is overhead to the model. In case of dual stacked autoencoder, dataset created after segmenting and labelling mitotic and non-mitotic images, is directly used for training without using any sampling methods for balancing the dataset (Table 3).

6 Conclusion

This project proposes a model which mitigate the class imbalance problem and improve the classifiers performance. This paper mainly focuses on class imbalance problem that is caused due to mitotic and non-mitotic nuclei in the breast cancer histopathological images. By effectively detecting mitotic cells help the pathologists to decide the degree of cancer. In many real life scenario there are class imbalance problem by solving this it can be applied to any other domain.

To detect the mitoses in the histopathology images of breast cancer, intially blue ratio histogram is applied to the origianl breast cancer images, and to separate nucleus from the whole image, global binary thresholding is applied and then some blob detections techniques are applied to get all the nuclues [7]. Nucleus are labelled either mitotic or non-mitotic by using the ground truths. This dataset is then trained on Dual stacked autoencoder.

References

1. Gao L, Zhang L, Liu C, Wu S (2020) Handling imbalanced medical image data: a deep-learning-based one-class classification approach. *Artif Intell Med* 108, Article 101935
2. Ng WWY, Zeng G, Zhang J, Yeung DS, Pedrycz W (2016) Dual autoencoders features for imbalance classification problem. *Pattern Recogn* 60:875–889
3. Lin W-C, Tsai C-F, Hu Y-H, Jhang J-S (2017) Clustering based undersampling in class imbalanced data. *Inform Sci* 409–410:17–26
4. Reza MS, Ma J (2018) Imbalanced histopathological breast cancer image classification with convolutional neural network. In *Proceedings of ICSP2018*, vol 2018, pp 619–624
5. Hartono, Sitompul OS, Tulus T, Nababan EB (2018) Biased support vector machine and weighted-smote in handling class imbalance problem. *Int J Adv Intell Inform* 4(1):21–27
6. He H, Bai Y, Garcia EA, Li S (2008) Adasyn: adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence), pp 1322–1328
7. Wahab N, Khan A, Lee YS (2017) Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection. *Comput Biol Med* 85
8. Bader-El-Den M, Teitei E, Perry T (2018) Biased random forest for dealing with the class imbalance problem. *IEEE Trans Neural Netw Learn Syst* 30(7):2163–2172
9. Chawla NV, Lazarevic A, Hall LO, Bowyer KW (2003) Smoteboost: improving prediction of the minority class in boosting. In: *Knowledge discovery in databases: PKDD 2003*, pp 107–119
10. Krawczyk B, Wozniak M, Herrera F (2014) Weighted one-class classification for different types of minority class examples in imbalanced data. In: 2014 IEEE symposium on computational intelligence and data mining (CIDM), pp 337–344
11. Ali A, Shamsuddin SM, Ralescu AL (2015) Classification with class imbalance problem: a review. *Int J Adv Soft Comput Its Appl* 7
12. Iranmehr A, Masnadi-Shirazi H, Vasconcelos N (2019) Cost-sensitive support vector machines. *Neurocomputing* 343
13. Galar M, Fernandez A, Barrenechea E, Bustince H, and Herrera F (2012) A review on ensembles for the class imbalance problem: Bagging, boosting, and hybrid-based approaches. *IEEE Trans Syst Man Cybern Part C: Appl Rev* 42(4):463–484
14. He H, Garcia EA (2009) Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 21(9):1263–1284
15. Baldi P (2011) Autoencoders, unsupervised learning, and deep architectures. In: *ICML unsupervised and transfer learning*, vol 27, pp 37–50
16. Maldonado S, Weber R, Famili F (2014) Feature selection for high-dimensional class-imbalanced data sets using support vector machines. *Inform Sci* 286: 228–246
17. Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P (2010) Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *J Mach Learn Res* 11:3371–3408
18. Rashmi R, Prasad K, Udupa CBK (2021) Breast histopathological image analysis using image processing techniques for diagnostic purposes: a methodological review. *Image Sig Process* 46(1), Article 7
19. Cruz-Roa A, Basavanahally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A (2014) Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *Medical imaging 2014: digital pathology*, vol 9041, Article 904103
20. Elmore JG et al (2015) Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 313(11):1122–1132
21. Japkowicz N (2000) The class imbalance problem: significance and strategies. In *Proceedings of international conference on artificial intelligence*, pp 111–117

22. Vijayakumar T (2019) Comparative study of capsule neural network in various applications. *J Artif Intell* 1(01):19–27
23. Batra K, Sekhar S, Radha R (2019) Breast cancer detection using CNN on mammogram images. In: International conference on computational vision and bio inspired computing. Springer, Cham, pp 708–716

Investigating Data Mining Trend in Cybercrime Among Youths



Ademola Olusola Adesina , Sunday Adeola Ajagbe ,
Olakunle Sunday Afolabi , Oluwashola David Adeniji ,
and Olalekan Ibrahim Ajimobi

Abstract This study investigated the emerging trend in cyber dynamics among youths using statistical tools. The study adopted a descriptive survey design. The population of the study comprised all cybercriminals in police custody in Ibadan, Nigeria. A simple random sampling procedure was used to select four hundred and eighteen participants for the study. The results of the analysis revealed that there was a significant and positive relationship between social influence and the prevalence of cybercrime among youth in the study area, there was a significant and positive relationship between economic influence and prevalence of cybercrime among youth in Africa, there was a significant and negative relationship between political influence and prevalence of cybercrime among youth in Africa and there was a significant and negative relationship between religious influence and prevalence of cybercrime among youth in Africa. Based on the findings of the study, the conclusion was reached that data mining and machine learning have a direct pattern on the prevalence of cybercrime and associated factors. The recommendation was made that authority in charge of cybercrime should ensure that data are available in the database for easy accessibility and quality conclusion on the causes of the prevalence of cybercrime among Nigeria youth.

A. O. Adesina

Department of Mathematical Sciences, Olabisi Onabanjo University, OOU, Ago-Iwoye, Nigeria
e-mail: ademola.adesina@ouagoiwoye.edu.ng

S. A. Ajagbe (✉)

Department of Computer Engineering, Ladoke Akintola University of Technology LAUTECH,
Ogbomoso, Nigeria
e-mail: saajagbe@pgschool.lautech.edu.ng

O. S. Afolabi

Department of Computer Science, University of Abuja, Abuja, Nigeria
e-mail: afolabi.olakunle2019@uniabuja.edu.ng

O. D. Adeniji

Department of Computer Science, University of Ibadan, Ibadan, Nigeria
e-mail: od.adeniji@ui.edu.ng

O. I. Ajimobi

Department of Computer Studies, The Polytechnic Ibadan, Ibadan, Nigeria

Keywords Cybercrime dynamics · Data mining · Emerging trend · Machine learning · Cybersecurity · Youths

1 Introduction

The most common types of cybercrime among the youngsters in the area were hacking, advance fee fraud, identity theft, and cyber terrorism [1]. Across the globe, countries are striving to improve their space and advance technologically for future purposes. Human society has entered the era of information technology thanks to the rapid development of computers and information technology. It's become increasingly crucial to figure out how to extract meaningful information resources from massive amounts of data. The organization of data in a systematic manner stimulates effectiveness and efficiency in an organized environment [2]. Because of the exponential growth of Internet connections, there has been a surge in the number of cyber-attacks, many of which have sad and serious consequences. Malware is the most prevalent weapon used in cyberspace to carry out malevolent intentions, either by exploiting existing weaknesses or by taking advantage of new technology's capabilities. The development of more innovative and effective malware defense systems has been identified as a major issue by the cybersecurity community. Machine learning (ML) is a sort of learning in which the process of learning is automated. According to [3] ML is a subset of artificial intelligence (AI) that allows computers to learn directly from data, and experience for decision making. This makes computers perform tasks that are ordinarily meant for the human being to perform, instead of following pre-programmed instructions, ML systems may learn from data and carry out complex tasks.

According to [4], ML is crucial in AI research. An intelligent system that is incapable of learning cannot be considered intelligent, although intelligent systems in the past were frequently incapable of learning. They can't, for example, fix themselves in time when they make mistakes. It does not acquire and discover the essential knowledge automatically. Its logic is limited to deduction and excludes induction. As a result, they can only prove known facts and theorems; they can't discover new theorems, laws, or rules. By gaining experience, it does not improve its performance. These limits are becoming more apparent as artificial intelligence advances. ML is now one of the cores of AI in any situation where data is involved, such as the expert systems, autonomous reasoning, natural language processing, pattern recognition, computer vision, cybersecurity, intelligent robotics, and other aspects of AI have all benefited from its use [5, 6].

Innovations among computer scientists have been useful and beneficial to humanities across the globe. An invention like AI has been rapidly advancing the field of computer science and on the contrary, come with some cybersecurity issues. Solutions in cybersecurity are usually investigated with much attention on the system but limited attention on the users. Hence, this study was informed. The aim of the

study was to investigate the emerging trend in cyber-dynamics using statistical tools specific objectives are to:

- i. Examine social influence on the prevalence of cybercrime among youth;
- ii. Investigate economic influence on the prevalence of cybercrime among youth;
- iii. Investigate political influence on the prevalence of cybercrime among youth and;
- iv. Investigate religious influence on the prevalence of cybercrime among youth.

Research Hypotheses

The following hypotheses were formulated to guide the objectives of the study.

H₁: There is no significant relationship between social influence and the prevalence of cybercrime among youth.

H₂: There is no significant relationship between economic influence and prevalence of cybercrime among youth.

H₃: There is no significant relationship between political influence and the prevalence of cybercrime among youth.

H₄: There is no significant relationship between religious influence and the prevalence of cybercrime among youth.

2 Review of Related Work

The use of ML and data mining tools could help to boost the rising trend in cyber-dynamics. Cyber-activities are computer-based activities that rely on data to complete tasks. To work properly, cyber-dynamics necessitates critical information (data). Cybercrime is common among the youth in Nigeria, as it is in many other African countries. The youth engage in various cybercriminal acts such as fraud, hacking, phishing, identity theft, and so on, which have tarnished the country's image worldwide. Although there are a variety of factors that influence the prevalence of cybercrime among youth, the impact of some of these factors, such as political social and religious influences, is investigated. Since no study has looked into machine learning and data mining for emerging trends in cyber dynamics, this study aimed to fill the void.

In Onitsha South Local Government Area (L.G.A) of Anambra State, Nigeria, [7] looked into the elements that contribute to cybercrime among youths. The study used a mixed-methods approach and used a multi-stage sampling strategy to pick 522 persons aged 18 and up from the study region. The quantitative data was acquired using a researcher-developed questionnaire, In-Depth Interviews were used to gather qualitative data (IDI). The descriptive statistics were utilized to assess the quantitative data, which was processed using the Statistical Package for Social Sciences (SPSS) software. The chi-square (2) inferential statistics were used to test the hypothesis at 0.05 significant levels. The qualitative data, on the other hand, was analyzed manually using content.

To help find the solution to cybersecurity, [8] go over some of the most commonly exploited vulnerabilities in existing hardware, software, and network layers. The critiques of present state-of-the-art mitigation approaches, as well as why they work or don't work, follow. The study then went into new attack patterns in areas including social media, cloud computing, smartphone technology, and critical infrastructure. Finally, the hypotheses were discussed and future study directions which include the psychological impact of cybersecurity were recommended.

The outbreak of COVID-19 has posed new issues for the government and residents, particularly for teenagers who are constantly on the go, resulting in new social media crimes. Cybercrime is a sort of crime that takes place in a digital setting. These offences vary from intrusions of privacy to crimes that ostensibly halt the socio-economic system. A survey was used to obtain data from 265 students from high schools and universities for this study. In all surveys that used a Likert scale, activities were measured. The validity and reliability of the method are investigated using a constructional equation modelling procedure. The findings show that in the COVID-19 period, social networks have no meaningful link with teenagers' proclivity for cybercrime. In the COVID-19 era, mobile games have a minor impact on teenagers' proclivity for cybercrime, whereas parents' religious beliefs have a substantial impact on teenagers' proclivity for cybercrime. The focus of the study was not considered in the study [9].

Lamidi [9] and Ahmad et al. [10] looked into how cybercrime impacts individuals and society as a whole, with the goal of exposing fraudsters' techniques and suggesting solutions. Personal experiences (phone conversations and text messages), WhatsApp communications, and stories posted on Facebook and Naira land forums by victims were used to compile data. They were analyzed in a descriptive manner. This threat, according to the article, can be mitigated by stricter legislative punishments and government enforcement. Individuals can also recognize false proposals and avoid cyberattacks by improving their grammar skills, improving their morality as prescribed by culture, logical reasoning, and avoiding greed.

Ahmad et al. [10], Singh et al. [11] offered a novel strategy that focuses on the most common cybercrimes around the world, as well as a detailed study conducted in Nigeria among secondary school pupils. In order to battle hackers, the study revealed a novel strategy for tackling cybercrime. Cybercrime, botnets, Malvertising, ransomware, and pornography are some of the terms used to describe cybercrime. Youth involvement in cybercrime is not implausible; Nigerian youths confront unemployment, a lack of basic social amenities, poverty, hunger, and poor education. The educational system has not made it easier; according to a recent figure, 3% of young people drop out of school. The influx of young people to urban areas where it is profitable to set up shop motivates cybercrime. It's startling to learn that Nigerian teenagers travel to other nations in search of a safe haven or a location where they may carry out their activities without being discovered [12].

This result of [13] asserts that the advent and advancement of the internet, computers and mobile devices have resulted in an increase in cybercrime in Nigeria and around the world. Young people between the ages of 18 and 45 conduct cybercrime in Nigeria, taking advantage of the rapid growth of information technology

to rob private and public businesses of their hard-earned riches. The rate at which young people are becoming involved in this crime is concerning, as it is said to be having a severe impact on the country's future and giving it a terrible international image.

This research aims to fill that void. It presents findings from a study on the impact of socio-demographic characteristics such as age, gender, marital status, education, occupation, and religion on cybercrime victimization, based on a sample of 1,354 internet-active Nigerians living in the Lagos metropolis. Younger respondents, men, respondents who have never married, respondents with a higher level of education, unemployed respondents, and Christians were shown to be more likely to be victims of cybercrime. The findings of this research have important policy implications for Nigeria's fight against cybercrime and criminality [14]

According to Aghatise (2006), technological advancements have resulted in significant changes in Nigerian cultures, socialization patterns, social institutions, and social interactions. According to him, adolescents, particularly undergraduates and the unemployed, have embraced information and communication technologies to the point that they use the internet for the majority of the day. The study went on to say that cybercriminals in Nigeria are mainly between the ages of 18 and 30, and are young people who are not in secondary school but are either in higher institutions of learning or still seeking admission. Similarly, a study conducted by a group called the youth against cybercrimes and fraud in Nigeria (2018) found that one out of every five youths in most Nigerian cities is a cybercriminal [15] claimed that the anonymity and privacy provided by the internet to potential users has increased the degree of fluidity and structural complexity of the "yahoo-boys" operations in Nigeria.

Similarly, [16] argue that the majority of cybercrime in Nigeria is directed at persons rather than computer systems, and thus requires less technical competence. They also point out that human flaws like greed and gullibility are frequently used by cyber thieves, resulting in financial and psychological harm to their victims. Eze and Adeniji [17] avowed that respectable firms' machine learning technologies will likewise become attractive targets. Attackers can use poisoned data sets to force otherwise helpful systems to make wrong judgments or present misleading perspectives of the monitored environment, inflicting harm, turmoil, and disruption. It's difficult to predict whether the beneficial or negative consequences of machine learning will win out. They talked about the undeniable rise of machine learning-powered systems on both sides of the cybersecurity divide, which is irreversibly changing the internet's safety.

Arising from the review of the literature, limited research considered the use of statistical tools to investigate issues in cybersecurity and cybercrime. The common tools used for the investigation of cybercrime are machine learning techniques, and statistical tools are undermined.

3 Research Methodology

This section provides a description of the processes involved in eliciting and analyzing data for the study. Steps taken to gather and order data before their statistical analyses in order to obtain results are stated. It contains the design of the study, setting of the study, source of data, target population, sample/ sampling technique, description of the collection of data instrument, evidence of validity and reliability of the instrument, method of data collection, method of data analysis and ethical considerations. The research methodology is represented as a research approach in Fig. 1.

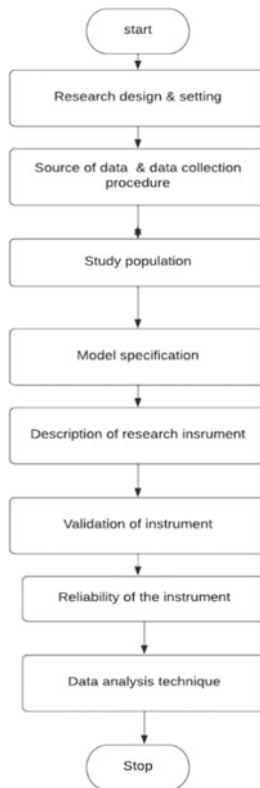


Fig. 1 Block diagram of research approach

3.1 Research Design and Setting

Generally, the research design was to specify the method and procedure for acquiring the information needed to structure the research for thorough analysis research. The survey research method was used because it will make the study flexible and it was also relevant in identifying relevant variables required to achieve the objectives of the study. For data collection, the study involves fundamental contact(s) between the researcher and the participants. The questionnaire and interview approaches were used to collect information and data. The survey research design elicited data directly from the field through the participants. Its adoption will be justified because of its tendency to elicit large data within a short time frame.

The research was carried out in a police custody setting where the targeted population can be reached. The research was conducted in Ibadan, Oyo State, at the State CID Iyaganku. Ibadan is the capital and largest city in the Nigerian state of Oyo. It is Nigeria's third-largest city, after Lagos and Kano, with a population of 3,649,000 people as of 2021 and a metropolitan area of nearly 6 million people. Geographically, it is the country's largest city.

3.2 Source of Data and Data Collection Procedure

To carry out its findings, the study approves primary and secondary data. The primary data in this study is a self-designed questionnaire that was used to elicit information on the factors that were examined. These variables include the dependent variable (prevalence of cybercrime) and the independent variable (social influence, religious influence economic influence and political influence). Secondary data was related work done on variables in context taken from the internet, magazines, and libraries.

After the approval of permission from the State CID Iyagaku, Ibadan Oyo State, copies of the questionnaire were administered to the cybercriminals in police custody. The following police stations were visited (Orita police station, Agugu police station, Airport police station, Gbagi police station and police headquarters in order to reach substantial respondents which could aid generalization in this study. The participants had been sufficiently advised on the importance of cooperating with the researcher. They were also guaranteed that their responses would be kept private. Administration of data was done in a week, 450 copies of questionnaires were administered and 426 copies of the questionnaire were retrieved. Appropriately filled copies of questionnaires (418) were sorted out and used for the analysis of the study.

3.3 Study Population

The population comprised all cybercrime convicts in the custody of the police. Youth within the age of 18–40 years are considered in this study. This research work made use of primary and secondary data. Self-designed questionnaires were used to obtain information about the study’s tested variables as primary data. These variables include social influence, religious influence economic influence and political influence on youth involved in cybercrime and prevalence of cybercrime among youth.

3.4 Model Specification

A model was a mathematical representation of reality that can take many shapes. This study’s model was created in response to the research questions. Figure 2 represent this model as it relates the influence of independent variables on the dependent variable in the study.

$$PCBC = \alpha + \beta_1P + \beta_2E + \beta_3S + \beta_4R + \beta_5E + \mu \tag{1}$$

α = Constant of the equation

β = Coefficient of the explanatory variable

Dependent Variable

PCBC = prevalence of cybercrime among youth

Independent variables

P = Political influence

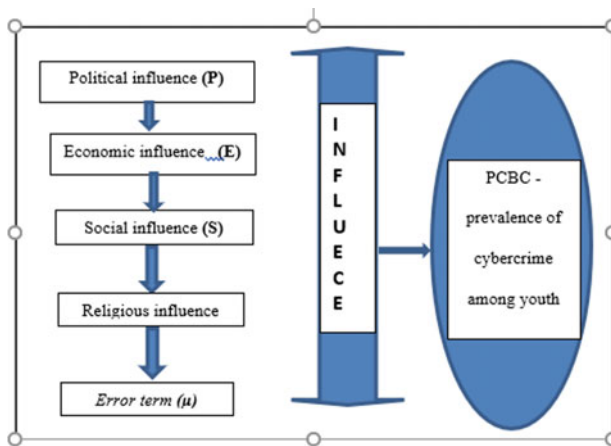


Fig. 2 Influence of independent variables on the dependent variable

E = Economic influence

S = Social influence

R = Religious influence

μ = Error term

3.5 Description of Research Instrument

There are four major independent variables in this study. The dependent variable is the user's prevalence of cybercrime among youth. This was measured by respondents' opinions on the aforementioned variables. The independent variables included a self-designed scale that addressed issues on social influence on cybercrime, economic influence on cybercrime, political influence on cybercrime and religious influence on cybercrime. This research tool is a structured questionnaire that includes demographic information as well as technical information (sections A through E). Section A of the demographic data comprised of questions based on the respondent's personal characteristics, such as gender, age group, length of engagement, educational background, and marital status.

Section B consisted of questions based on the dependent and independent variables (prevalence of cybercrime among youth). The instrument was a self-constructed questionnaire, which addressed types of cybercrime youth engaged in. the listed crime includes online fraud, identity theft, hacking, phishing, social engineering and cyber-stalking.

Section C consisted of questions that addressed the political influence on cybercrime among youth in Nigeria. The scale was a 7-item scale ranked under the following five-scale point Strongly Agree (SA); Agree (A); Undecided (U); Disagree (D) and Strongly Disagree (SD).

Section D consisted of questions that addressed the economic influence on cybercrime among Nigeria youth. It was a 10-item scale ranked under the following five-scale point Strongly Agree (SA); Agree (A); Undecided (U); Disagree (D) and Strongly Disagree (SD).

Section E consisted of questions that addressed the social influence on cybercrime among youth in Nigeria. The scale was a 10-item scale ranked under the following five-scale point Strongly Agree (SA); Agree (A); Undecided (U); Disagree (D) and Strongly Disagree (SD).

Section F consisted of questions that addressed the religious influence on cybercrime among youth in Nigeria. The scale was an 8-item scale ranked under the following five-scale point Strongly Agree (SA); Agree (A); Undecided (U); Disagree (D) and Strongly Disagree (SD).

3.6 Validation of Research Instrument

For any data collection methods via questionnaire, an instrument is valid if it measures what it purports to measure. Content validity is regarded as a crucial strategy for the attribute for which it was created. Subjecting the instrument to criticism from supervisors, experts, and authorities in the field of inquiry is an excellent technique to ensure the content validity of the questionnaire, and this is exactly what the researcher did. Expert critical judgment went a long way toward removing extraneous items and demonstrating that the instrument's content is actually addressing the proper question(s) essential to the investigation's goal.

3.7 Reliability of the Instrument

The consistency with which a test measured whatever it did define its dependability. It was usually expressed in terms of some kind of item and test reliability coefficient. The reliability of the questionnaire produced for data collection was determined using statistical procedures that included the use of testing principles such as to estimate methods and the distribution of questionnaires to a sub-sample of the population under study at one-week intervals. The investigation yielded a Cronbach alpha reliability coefficient of 0.874.

3.8 Data Analysis Technique

The data collected from the respondents were analyzed using descriptive statistics (percentage and frequencies), while Pearson Product Moment Correlation (PPMC) was used to determine the relationship between the variables in this study, and multiple regressions were used to determine the significant influence between the dependent and independent variables.

4 Results and Analysis of the Outcome

This section presents the analysis (descriptive and inferential) and interpretation of data gather in the research on the investigation of the emerging trend in cyberdynamics. The result presented and discussed in this section are those obtained from cybercriminals in police custody.

4.1 The Socio-Demographic Characteristics of Participants in the Study Area

The socio-demographic characteristic of the respondents of the study is explained using age, sex, marital status and educational status in Tables 1, 2, 3 and 4. Table 5 explained the years of engagement in cybercrime.

Table 1 shows the age distribution of the participants. The above table presented 84(20.1%) respondents within the age range 18–25 years, 250(59.8%) respondents within the age range 26–33 years and also 84(20.1%) respondents with age range 18–25 years.

Table 2 shows the gender of the participants. The table presented 362 (86.6%) respondents as male and 56(13.4%) as female. In this study, there are more male participants in the study.

Table 1 Distribution of the respondents by age

| Age (years) | Frequencies | Percentage (%) |
|-------------|-------------|----------------|
| 18–25 | 84 | 20.1 |
| 26–33 | 240 | 59.8 |
| 34–40 | 84 | 20.1 |
| Total | 418 | 100.0 |

Table 2 Distribution of the respondents by sex

| Sex | Frequencies | Percentage (%) |
|--------|-------------|----------------|
| Male | 362 | 86.6 |
| Female | 56 | 13.4 |
| Total | 418 | 100.0 |

Table 3 Distribution of the respondents by marital status

| Marital status | Frequencies | Percentage (%) |
|----------------|-------------|----------------|
| Single | 280 | 67.0 |
| Married | 124 | 29.7 |
| separate | 14 | 3.3 |
| Total | 418 | 100.0 |

Table 4 Distribution of the respondents by education status

| Education qualification | Frequencies | Percentage (%) |
|-------------------------|-------------|----------------|
| Primary education | 40 | 9.6 |
| Secondary education | 126 | 30.1 |
| Tertiary education | 252 | 60.3 |
| Total | 418 | 100.0 |

Table 5 Distribution of the length of engagement in cybercrime (in Years)

| Years of experience in cybercrime | Frequencies | Percentage (%) |
|-----------------------------------|-------------|----------------|
| 1–5 | 70 | 16.7 |
| 6–10 | 264 | 63.2 |
| 11–15 | 56 | 13.4 |
| 16–20 | 28 | 6.7 |
| Total | 418 | 100.0 |

In Table 3, the marital status of the participants was presented. The table showed that 280(67%) of the participants were single; 124(29.7%) of participants were married and 14(3.3%) were separated.

In Table 4, the educational status of the participants was indicated. The table revealed that 40(9.6%) of participants had primary education, 126(30.1%) had secondary education and 252(60.3) had tertiary education.

In Table 5, the years of engagement in cybercrime was indicated. The table showed that 70(16.7%) participants had engaged in cybercrime for 1 to 5 years, 264(63.2%) participants indicated that they had engaged for 6–10 years, 56(13.4%) participants indicated that they had engaged for 11–15 years, 28(6.7%) participants had engaged in cybercrime for 16–20 years.

4.2 Cybercrime Activities

Table 6 shows the most frequent cybercrime the youth engaged themselves in. the table presented hacking (93.3%) as the most cybercrime youth engaged in, followed by fraud (89.9%), next was identity theft (77.0%), social engineering was (56.9%), cyberstalking (53.6%) and phishing the least was (43.5%). The indication of this result was that youth engaged in cybercriminal activities like hacking, fraud and identity theft.

Table 6 shows the most engaged cybercriminal activities among youth in the study area

| S/N | Types of cybercrime | Engaged (%) | Unengaged (%) |
|-----|---------------------|-------------|---------------|
| 1 | Fraud | 376 (89.9) | 42 (10.1) |
| 2 | Identity theft | 322 (77.0) | 96 (28.0) |
| 3 | Hacking | 390(93.3) | 28(6.7) |
| 4 | Phishing | 182(43.5) | 236(56.5) |
| 5 | Social engineering | 238(56.9) | 180(43.1) |
| 6 | Cyberstalking | 224(53.6) | 194(46.4) |

Table 7 Showing the relationship between cybercrime and social influence

| Variables | N | Mean | SD | r | P | Remark |
|--------------------------|-----|-------|------|-------|-------|-------------|
| Prevalence of cybercrime | 418 | 10.78 | 3.58 | 0.549 | 0.016 | Significant |
| Social influence | | 32.00 | 5.04 | | | |

Table 8 Shows the relationship between cybercrime and economic influence

| Variables | N | Mean | SD | r | P | Remark |
|--------------------------|-----|-------|------|-------|-------|-------------|
| Prevalence of cybercrime | 418 | 10.78 | 3.58 | 0.376 | 0.000 | Significant |
| Economic influence | | 32.67 | 4.18 | | | |

4.3 Hypotheses Testing

4.3.1 The Relationship Between Cybercrime and Factors that Push Youth to It.

Table 7 shows the relationship between the prevalence of cybercrime and social influence. The table presented a significant relationship of social influence and prevalence of cybercrime among the youth, “r” = 0.549 and P = 0.016 which was less than 0.05. The null hypothesis was rejected and alternate was accepted. Hence, social influence has a significant influence on the prevalence of cybercrimes among the youth in the study area.

4.3.2 Relationship Between Cybercrime and Economic Influence

Table 8 shows the relationship between the prevalence of cybercrime and economic influence. The table presented a significant relationship of economic influence and prevalence of cybercrime among the youth, “r” = 0.376 and P = 0.000 which was less than 0.05. The null hypothesis was rejected and alternate was accepted. Hence, economic influence has a significant influence on the prevalence of cybercrimes among the youth in the study area.

4.3.3 Relationship Between Cybercrime and Political Influence

Table 9 shows the relationship between the prevalence of cybercrime and economic influence. The table presented a significant relationship of economic influence and prevalence of cybercrime among the youth, “r” = -0.219 and P = 0.000 which was less than 0.05. The null hypothesis was rejected and alternate was accepted. Hence, political influence has a negative significant influence on the prevalence of cybercrimes among the youth in the study area.

Table 9 Shows the relationship between cybercrime and political influence

| Variables | N | Mean | SD | r | P | Remark |
|--------------------------|-----|-------|------|--------|-------|-------------|
| Prevalence of cybercrime | 418 | 10.78 | 3.58 | -0.219 | 0.001 | Significant |
| Political influence | | 24.06 | 3.17 | | | |

Table 10 Showing the relationship between cybercrime and religious influence

| Variables | N | Mean | SD | r | P | Remark |
|--------------------------|-----|-------|------|--------|-------|-------------|
| Prevalence of cybercrime | 418 | 10.78 | 3.58 | -0.181 | 0.001 | Significant |
| Religious influence | | 24.06 | 3.17 | | | |

4.3.4 Relationship Between Cybercrime and Religious Influence

Table 10 shows the relationship between the prevalence of cybercrime and economic influence. The table presented a significant relationship between religious influence and prevalence of cybercrime among the youth, “r” = -0.181 and P = 0.000 which was less than 0.05. The null hypothesis was rejected and alternate was accepted. Hence, religious influence has a negative significant influence on the prevalence of cybercrimes among the youth in the study area.

4.4 Discussion of Findings

This study examined the emerging trend in cyber dynamics using statistical tools, specifically study was to examine the social influence on the prevalence of cybercrime among youth in Nigeria, to investigate the economic influence on the prevalence of cybercrime among youth in Nigeria, to investigate the political influence on the prevalence of cybercrime among youth in Nigeria, and to investigate the religious influence on the prevalence of cybercrime among youth in Nigeria. To ascertain the objective of the study six hypotheses were formulated and ran with the statistical package SPSS.

The results of the hypotheses are discussed in the tables above. This section dealt with the discussion of the outcome of the hypotheses by comparing them with other related and relevant studies.

Ho₁: tested if there is no significant relationship between social influence and prevalence of cybercrime among youth in Nigeria. The result of this hypothesis indicated that there was a significant and positive relationship between social influence and the prevalence of cybercrime among youth in the study area. The implication of the result was that youth engagement in cybercriminal activities are induced by the social environment. Some individuals in society could be the promoter of the engagement of youth in cybercrime. The result on the social influence indicated youth are

supported with facilities like phone, data and electricity to engage cybercrime, the report supported [13, 17].

Ho₂: tested if there is no significant relationship between economic influence and prevalence of cybercrime among youth in Nigeria. The result of this hypothesis indicated that there was a significant and positive relationship between economic influence and the prevalence of cybercrime among youth in the study area. The implication of the result was that youth engagement in cybercriminal activities are encouraged by hardship experienced in the economy. Parameters such as unemployment, poverty and unfavourable economic policy have pushed many youths into cybercrime in order to earn a living.

Ho₃: tested if there is no significant relationship between political influence and prevalence of cybercrime among youth in Nigeria. The result of this hypothesis indicated that there was a significant and negative relationship between political influence and the prevalence of cybercrime among youth in the study area. The negative influence indicated that lawmakers could stop the prevalence of cybercrime among youth if they enacted laws that could prevent youth's engagement in cybercrime. They could also reduce the prevalence of cybercrime by formulating policies that encouraged youth empowerment and the creation of employment opportunities for the youth. The youth have no belief in the elected politicians and they believed that politicians are self-centred and do not care about the welfare of the citizen.

Ho₄: tested if there is no significant relationship between religious influence and prevalence of cybercrime among youth in Nigeria. The result of this hypothesis indicated that there was a significant and negative relationship between religious influence and the prevalence of cybercrime among youth in the study area. Religion in this study is noted as a way of life and the supernatural doctrine has been transmitted to the youth, this work is in line with [14].

5 Conclusion and Future Direction

In this section, conclusions on the findings of the study data were gathered from 418 internet criminals, who engage in cybercrime. To fulfil the study's goal, the descriptive research method was used, and data were analyzed using descriptive statistics (like simple percentages, frequency count, mean, and PPMC). According to the findings of the study's data analysis, there was a significant and positive relationship between social influence and the prevalence of cybercrime among Nigerian youth, a significant and positive relationship between economic influence and the prevalence of cybercrime among Nigerian youth, and a significant and negative relationship between political influence and the prevalence of cybercrime among Nigerian youth.

Considering the findings of this study, a conclusion was reached on the influence of statistical tools to investigate the emerging trend in cyber-dynamics. The study through data mining revealed the pattern of relationship of data mining on the prevalence of cybercrime and associated factors. The study showed a direct/positive

pattern of relationship between the variables. More so, the use of correlation established the predictive influence of factors such as social influence, economic situation and unemployment on the prevalence of cybercrime.

The following recommendations were made based on the study's findings: Data required to gather vital information on cybercrime in Nigeria are not available, this propel the researchers to used primary data through questionnaires. This is a shortfall observed in the field. Hence, the authority in charge of cybercrime should ensure that data are available in the database for easy accessibility and quality conclusion on the causes of the prevalence of cybercrime among Nigerian youth.

References

1. Udelue, M. C., & Bentina, M. (2019). Prevalence of Cybercrimes Among Youths in onitsha South Local Government Area of Ananmbra State, Nigeria. *International Journal of Health and Social Inquiry*, 5(1), 82-106.
2. Hegde, J., & Rokseth, B. (2019). Applications of machine learning methods for engineering risk assessment – A review. *Safety Science*. doi:<https://doi.org/10.1016/j.ssci.2019.09.015>
3. Li, Y. (2014). *Application of machine learning algorithms in data mining*. China: Beijing University of Posts and Telecommunications (BUPT).
4. Teng, X., & Gong, Y. (2018). Research on Application of Machine Learning in Data Mining. *IOP Conference Series. Materials Science and Engineering*, 392(6). doi:<https://doi.org/10.1088/1757-899X/392/6/062202>
5. Ajagbe, S. A., Oladipupo, M. A., & Balogun, E. O. (2020). Crime Belt Monitoring Via Data Visualization: A Case Study of Folium. *International Journal of Information Security, Privacy and Digital Forensic*, 35–44. Retrieved from <https://library.ncs.org.ng/download/crime-belt-monitoring-via-data-visualization-a-case-study-of-folium/>
6. Udelue, M. C., Mathias, B. A., & Ezeh, S. S. (2020). socioeconomic correlates of youths involvement in cybercrime: perceptions of residents in Onitsha south L.G.A, Anambra State, Nigeria. *International Journal of Social Sciences and Humanities Reviews*, 10(3), 66–79.
7. Jang-Jaccard, J., & Nepal, S. (2014). A survey of emerging threats in cybersecurity. *Journal of Computer and System Sciences*, 80, 973–993. doi:<https://doi.org/10.1016/j.jcss.2014.02.005>
8. Li, Y., Li, J., Fan, Q., & Wang, Z. (2022). Cybercrime's tendencies of the teenagers in the COVID-19 era: assessing the influence of mobile games, social networks and religious attitudes. *Kybernetes*. doi:<https://doi.org/10.1108/K-07-2021-0582>
9. Lamidi, M. T. (2020). Investigating Cybercrime in Nigeria. *Encyclopedia of Criminal Activities and the Deep Web*. doi:<https://doi.org/10.4018/978-1-5225-9715-5.ch069>
10. Ahmad, M. A., Wisdom, D. D., & Isaac, S. (2020). An Empirical Analysis of Cybercrime Trends and Its Impact on Moral Decadence Among Secondary School Level Students in Nigeria. *The 26th iSTEAMS Bespoke Multidisciplinary Conference, Accra Ghana* . IEEE. doi:<https://doi.org/10.22624/iSTEAMS/V26P10-IEEE-NG-TS>
11. Singh, A. K., Prasad, N., Narkhede, N., & Mehta, S. (2016). Akshay Kumar Singh, Neha Prasad, Nohil Narkhede and Siddharth Mehta “Crime: Classification and Pattern Prediction”, *IARJSET*, Vol. 3, pp. 41–43. *IARJSET*, 3, 41–43.
12. Ogunjobi, O. (2020). *The Impact of Cybercrime on Nigerian Youths*. Poole, United Kingdom: Bournemouth University.
13. Ndueze, P. N., Igbo, E. U., & Okoye, U. O. (2013). Cyber Crime Victimization among Internet active Nigerians: An Analysis of SocioDemographic Correlates Demographic Correlates. *International Journal of Criminal Justice Sciences*, 8(2), 225–234.

14. Bamimore, I., & Ajagbe, S. A. (2020). Design and implementation of smart home for security using Radio Frequency modules. *International Journal of Digital Signals and Smart Systems*, 286-303. doi:<https://doi.org/10.1504/IJDSS.2020.111009>
15. Longe, O. B., Chiemeké, S. C., Fashola, S., Longe, F., & Omilabu, A. (2008). "Internet Service Providers and Cybercrime in Nigeria –Balancing Services and ICT Development. Retrieved from <https://www.intgovforum.org/cms/documents/contributions/general-contribution/2008-1/349-longe-o-b-et-al-isp-and-cybercrime-in-nigeria-igf-contributions/file>
16. Adeniji, O. D., & Olatunji, O. O. (2020). Zero Day Attack Prediction with Parameter Setting Using Bi Direction Recurrent Neural Network in Cyber Security. *International Journal of Computer Science and Information Security (IJCSIS)*, 18(3), 111-118.
17. Eze, C. V., & Adeniji, O. D. (2014). Character Proximity For RFID Smart Certificate System: A Revolutionary Security Measure ToCurb Forgery Menace. *International Journal of Scientific & Technology Research*, 3(10), 66-70.

Concatenation of SEPIC and Matrix Converter with LCL Filter for Vehicle to Grid Application



Anushka Tripathi, S. P. Soundharya, R. Vigneshwar, and M. R. Rashmi

Abstract The increased use of Electric Vehicles (EV) helps to reduce the use of fuel energy consumption. This in turn mitigates the bad environmental impact and climate changes caused by greenhouse gas and carbon dioxide emissions. At present we are moving towards Electric Vehicles and they majorly contribute to the development of the smart grid by bidirectional power flow in the connected distribution grid. Meeting peak demand is one of the major challenges that the power sector faces. Integration of electric vehicles to the grid unlocks the ability to optimize the electricity is produced and transported by turning electric cars into a virtual powerplant. Since most vehicles are in stand still condition for 90% of the time, the EVs connected to the grid via a bidirectional converter can be used during peak power demands to supply energy to the grid. When the power demand in the grid is low, the EV's battery can be charge from the grid. This paper presents a combination of SEPIC converter and Matrix converter for Vehicle to Grid(V2G) application.

Keywords Electric vehicles (EVs) · Vehicle-to-grid (V2G) · Grid-to-vehicle (G2V) · SEPIC/Zeta converter · Matrix converter · Inductor capacitor inductor (LCL) filter

1 Introduction

Electric Vehicles are an eco-friendly and efficient means of transportation as they use battery storage system instead of fuel consumption methods for running the vehicle [1–3]. They also play a crucial role through their interaction with the electric grid specially during the peak demand hours [4]. Research has proposed various converter topologies for bidirectional application which can work both in inversion as well as rectification mode [5, 6]. The DC-DC converter used in V2G and G2V applications should be capable of either bucking or boosting the voltage [7–9]. The newly proposed bidirectional SEPIC/Zeta converter in [10] for this application has

A. Tripathi · S. P. Soundharya · R. Vigneshwar · M. R. Rashmi (✉)
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: mr_rashmi@blr.amrita.edu

the capability of bilateral power flow, low conduction and switching losses as a result of Zero Voltage Switching (ZVS) and the operation of a synchronous rectifier. The bi-directional converter can be divided into two parts according to the requirement of the battery and the grid respectively, the converter can behave like a SEPIC converter when there is forward power flow, and for backward power flow it behaves like a zeta converter [11, 12]. The efficiency of the main MOSFETs and the diodes are both improved due to the ZVS operation in the improved circuit.

Lithium-Ion batteries are very widely used in electric vehicles [13]. This is due to its longer life cycle and high specific energy. Better charge transfer is obtained by charging the Li-battery with ripple currents at an optimal frequency. Hence, the use of Low Frequency Ripple Current (LFRC) to charge the Li-battery while using a bi-directional converter result in improved charge transfer frequency. Durability and efficiency of a Lithium-ion battery depends on the charging technique used. Constant Current Constant Voltage (CC-CV), Five Step Charging Pattern and Pulse Charging Method are some of the available fast charging techniques. The V2G and G2V system could interface with the EVs via a DC fast charging station. The harmonic distortion of the grid injected current must be kept at minimum and this must be ensured by the charging station. The controller must have good dynamic performance in terms of the DC bus voltage stability [3].

During V2G mode of operation, output from the battery has to be inverted to inject to the grid. Whereas, in G2V application the output from the grid has to be rectified for charging the battery. A matrix converter can be used in such applications as it is capable of working in both inversion mode and rectification mode [14–16]. To accomplish AC-AC conversion, AC-DC conversion or AC-AC conversion unidirectional switches which are connected in antiparallel in the form of H-bridge are used. Venturina method is used for obtaining the control signal for the modulation of the switches. To obtain better efficiency and output, the matrix converter uses direct IGBT switching technique [17, 18]. There are various modulation techniques for matrix converter among which the sinusoidal pulse width modulation is the simplest modulation technique. Researcher have proposed many filters after matrix converter to smoothen the voltage to a pure sinusoidal voltage of grid frequency before it is injected to the grid. In order to minimize the harmonic distortion in the final output waveform a conventional LCL filter was proposed in [19]. This helps remove the harmonics from the square pulse output of the matrix converter. The output can then be obtained as a pure sinusoidal waveform. Vehicle security is also a major concern. A block chain based communication between the vehicles was proposed in [20] to reduce the cyber attacks. In autonomous vehicles, the speed is regulated by adapting artificial intelligent techniques. Combination of genetic algorithm, memetics and adaptive direct search based optimization algorithm was proposed in [21–23] for autonomous electric vehicles.

In this paper, a SEPIC/Zeta converter is cascaded with matrix converter and LCL filter for V2G operation. Design, implementation and simulation results of the same are presented.

2 Circuit Description

The block diagram representation of system for V2G and G2V application is shown in Fig. 1. It consists of Battery, Bidirectional SEPIC/ZETA DC-DC converter, Matrix Converter which can work as an inverter and rectifier and the Grid. The direction of the power flow is also indicated in the diagram. When the grid requires energy, the power flows from the battery to the grid. For V2G operation, battery voltage of 48 V has to be boosted to 240 V which is done with the help of SEPIC Converter. SEPIC converter is preferred, since the output of SEPIC converter is not inverted. This boosted voltage is fed to the matrix converter which works as an inverter and its output is a square wave. Therefore, the harmonics in the output voltage has to be removed. An LCL filter is utilized to decrease the harmonic contents in the output of matrix converter. The grid is injected with a voltage at 50 Hz frequency.

Whenever the demand on the grid is less, the battery will be charged. In order to charge the battery, the grid supply has to be rectified. Therefore, the matrix converter works in rectification mode. The output obtained from the matrix converter is fed to the Zeta converter which bucks down the voltage to meet the battery voltage requirement. This paper presents the mode of operation of V2G application. Battery specifications is shown in Table 1.

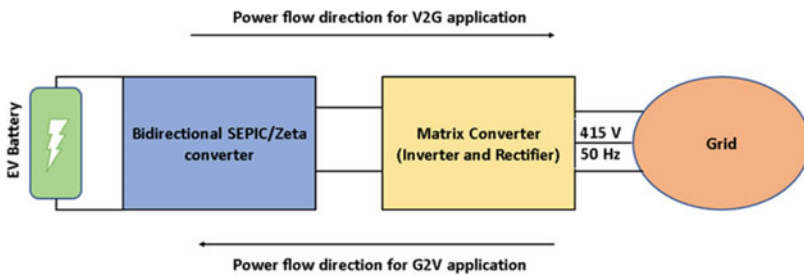


Fig. 1 Block diagram of system for V2G and G2V application

Table 1 Battery specifications

| Parameter | Value |
|---------------------|--------------------------|
| Voltage | 48 V |
| Usage/application | Electric vehicles |
| Brand | Explore synergy synocare |
| Battery capacity | 32 Ah |
| Model name/number | ESPL/4832/LI-ION |
| Max pulse discharge | 96 A |
| Storage temperature | 30 °C |
| Cycle life | 1200 |

3 Single Ended Primary Inductor Converter (SEPIC)

Figure 2 represents the circuit diagram of the SEPIC converter which comprises of two switches, two inductors and two capacitors. The SEPIC Converter’s modes of operation are given in Fig. 3.

- **Mode 1:**

This mode of operation is shown in Fig. 3a. The switches S1 and S2 are ON and OFF respectively. Initially, capacitors C1 and C2 is charged, inductor L1 gets charged from the source and the current flows through the path with lower resistance back to the source. Capacitor C1 discharges through L2 and C2 supplies the load. The circuit equations during this mode of operation are given in Eqs. (1–5)

Using Kirchoff’s voltage law:

$$V_{L1} = V_S \tag{1}$$

(For instantaneous and average voltages)

$$-V_s + V_{L1} + V_{C1} - V_{L2} = 0 \tag{2}$$

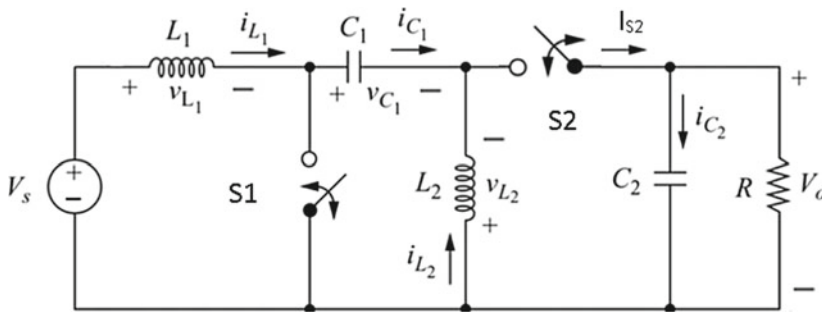


Fig. 2 SEPIC converter

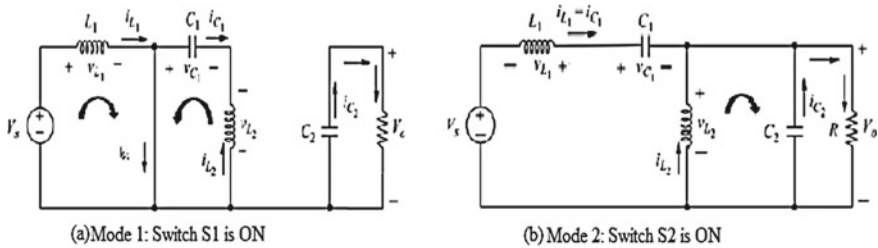


Fig. 3 SEPIC converter—modes of operation

During steady state condition, (average voltages)

$$V_{L1} + V_{L2} = 0 \quad (3)$$

$$V_{C1} = V_s \quad (4)$$

$$V_{L1} = V_{C1} = V_s \quad (5)$$

where,

V_s -Supply Voltage.

V_{L1} -Voltage across the inductor L_1

V_{L2} -Voltage across the inductor L_2

V_{C1} -Capacitor C_1 Voltage.

V_{C2} -Capacitor C_2 Voltage.

V_o -Output Voltage

- **Mode 2:**

Mode 2 operation is shown in Fig. 3b. In this mode Switch S_2 is ON and Switch S_1 is OFF. According to Lenz's law, the inductor does not allow sudden change in current. Hence, L_1 reverses voltage polarity and charges C_1 . Similarly, L_2 discharges through C_2 . In turn, C_2 discharges through the load and the polarity of the output remains same as that of the input voltage. The equations describing this mode of operation are given in Eqs. (6–7).

$$-V_s + V_{L1} + V_{C1} + V_o = 0 \quad (6)$$

$$V_{L1} = -V_o \quad (7)$$

3.1 SEPIC Converter Design

The battery voltage is 48 V. The SEPIC converter should have an output voltage (V_{out}) of 240 V to obtain AC voltage of 415 V.

- Duty cycle consideration:

The voltage input variation is assumed to be $\pm 5\%$; $V_{in} = 48$ V

Hence, $V_{in(min)} = 48 \times 0.95 = 45.6$ V, $V_{in(max)} = 48 \times 1.05 = 50.4$ V

$V_{in(min)}$ -Minimum Supply Voltage

$V_{in(max)}$ -Maximum Supply Voltage

$$\frac{V_o}{V_s} = \frac{D}{1-D} \quad (8)$$

Hence duty cycle, $D = 0.833 = 83.33\%$

$$D_{min} = \frac{V_{out}}{V_{in(max)} + V_{out}} = \frac{240}{50.4 + 240} = 0.826 \quad (9)$$

$$D_{max} = \frac{V_{out}}{V_{in(min)} + V_{out}} = \frac{240}{45.6 + 240} = 0.840 \quad (10)$$

where,

D-Duty cycle

D_{min} =Duty cycle when supply voltage is maximum.

D_{max} -Duty cycle when supply voltage is minimum

• **Inductor selection:**

The peak-to-peak ripple current (ΔI_L) is taken as 5% of the maximum input current at minimum input voltage.

Switching frequency, $F_{sw} = 30 \text{ kHz}$

$$\Delta I_L = I_{out} \times \frac{V_o}{V_{in(min)}} \times 5\% = 6.4 \times \frac{240}{45.6} \times 0.05 = 1.68 \text{ A} \quad (11)$$

Inductor values,

$$L_1 = L_2 = L = \frac{V_{in(min)}}{\Delta I_L \times F_{sw}} \times D_{max} = \frac{45.6}{13.47 \times (30 \times 10^3)} \times 0.840 = 94.78 \mu\text{H} \quad (12)$$

$$I L_{1\text{peak}} = I_{out} \times \frac{V_{out}}{V_{in(min)}} \left[1 + \frac{5\%}{2}\right] = 6.4 \times \frac{240}{45.6} \times \left[1 + \frac{0.05}{2}\right] = 34.526 \text{ A} \quad (13)$$

$$I L_{2\text{peak}} = I_{out} \times \left[1 + \frac{0.05}{2}\right] = 6.56 \text{ A} \quad (14)$$

• **SEPIC coupling capacitor selection:**

The minimum capacitance required to reduce the ripple voltage (V_{pp}) by 5% of the output voltage: $V_{pp} = 5\%$ of $V_{out} = 0.05 \times 240 = 12 \text{ V}$

$$C_1 = \frac{I_{out} \times D \times (1-D) \times 1000}{F_{sw} \times V_{pp}} = \frac{6.4 \times 0.833 \times (1-0.833) \times 1000}{30 \times 10^3 \times 12} = 2.47 \text{ mF} \quad (15)$$

$$I_{C1(rms)} = I_{out} \times \sqrt{\frac{V_o}{V_{in(min)}}} = 6.4 \times \sqrt{\frac{240}{45.6}} = 14.68A \quad (16)$$

- **Selection of Output Capacitor:**

The peak-peak ripple voltage is considered to be 5% of the input voltage. $I_{cout} = I_{c1} = 14.68A$

Effective Series Resistance (ESR)

$$ESR = \frac{V_{ripple} \times 0.5}{IL_{1peak} + IL_{2peak}} = \frac{0.05 \times 240 \times 0.5}{34.98 + 6.56} = 0.144 \Omega \quad (17)$$

$$C_2 = \frac{I_{out} \times D_{max}}{V_{ripple} \times 0.5 \times F_{sw}} = \frac{6.4 \times 0.84}{0.05 \times 240 \times 0.5 \times (30 \times 10^3)} = 29.86 \mu F \quad (18)$$

- **Switch S1:**

$$Duty\ cycle\ of\ switch\ S1 - D_{S1} = \frac{V_o}{V_{in} + V_o} = \frac{240}{48 + 240} = 0.833 = 83.33\% \quad (19)$$

- **Switch S2:**

$$Duty\ cycle\ switch\ S2, D_{S2} = \frac{V_o}{V_{in} + V_o} = \frac{48}{240 + 48} = 0.1666 = 16.66\% \quad (20)$$

$$Phasedelay = T \times D_{S1} = \frac{1}{30000} \times 0.833 = 27.7 \mu s \quad (21)$$

4 Matrix Converter

Matrix converter is a bidirectional converter which can convert AC to DC or DC to AC or AC to AC. The converter topology has 12 IGBTs as shown in Fig. 4.

Matrix converter is used in the inverting mode to convert direct current to alternating current. Switch S1, S3, S5, S4, S6, S2 is used for inverting operation and the switching waveforms are shown in Fig. 5.

Calculation of power grid inversion:

$$V_{dc} = 1.65 * V_s \quad (22)$$

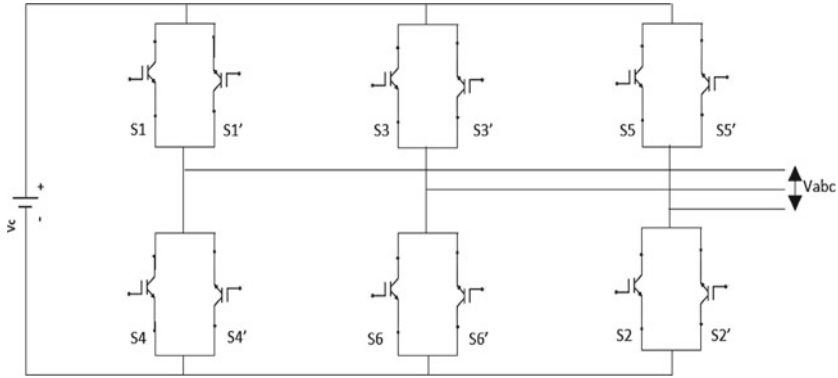


Fig. 4 Matrix converter

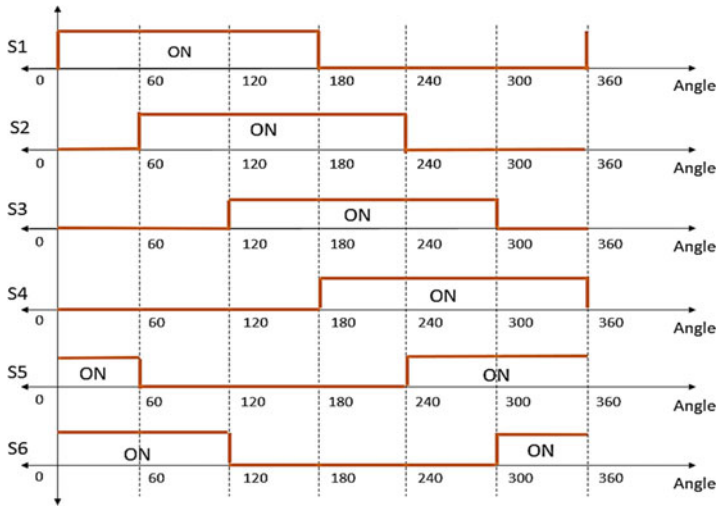


Fig. 5 Matrix converter switching waveforms

where V_{dc} -DC link voltage

$$V_s = \frac{V_{line}}{\sqrt{3}} \tag{23}$$

$$\text{Line Voltage } V_{line} = \frac{V_{dc} * \sqrt{3}}{1.65} = \frac{240}{1.049} = 228.79V \tag{24}$$

$$V = V_{dc} * \sqrt{3} = 240 * \sqrt{3} = 415.7V \tag{25}$$

5 LCL Filter

The maximum ripple current (Δi_{L1max}) can be derived as,

$$\Delta i_{L1max} = \frac{1}{8} \frac{V_{dc}}{L_1 \cdot f_{sw}} = 5\% * i_{rated} \tag{26}$$

The capacitor value C can be determined using the equation,

$$C = 15\% * \frac{P_{rated}}{2.3\pi f_{line} \cdot V_{rated}^2} \tag{27}$$

where

P_{rated} - Rated Output power

V_{rated} - Rated Output Voltage

6 Simulation Results

MATLAB/Simulink software is used to simulate the circuit. The SEPIC converter in forward mode (boosting mode) is shown in Fig. 6.

Battery voltage and current waveforms are shown in Fig. 7. SEPIC converter output current and voltages are shown in Fig. 8. Battery voltage of 48 V is boosted to 240 V and the output of the SEPIC converter is fed to the matrix converter. Matrix converter simulation circuit is shown in Fig. 9. Switching pulses for matrix converter are given in Fig. 10. Phase voltage and currents of matrix converter are shown in Fig. 11, line voltage and line currents are given in Fig. 12. FFT spectra is shown in Fig. 13. The THD is 66.9%

To reduce THD in the output voltage, LCL filter is used. The Matrix converter with an LCL filter is simulated using MATLAB as shown in Fig. 14 and its output

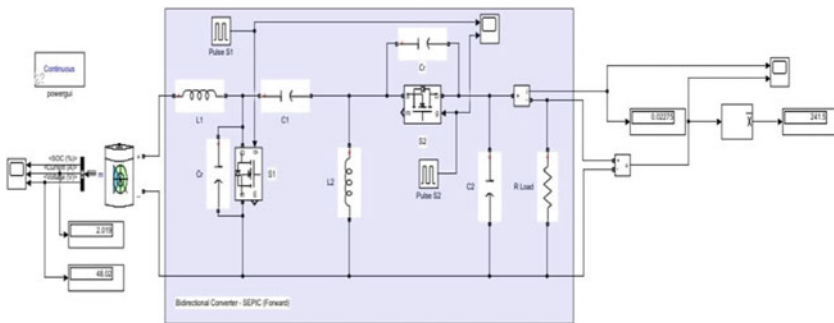


Fig. 6 Bidirectional SEPIC converter—forward mode

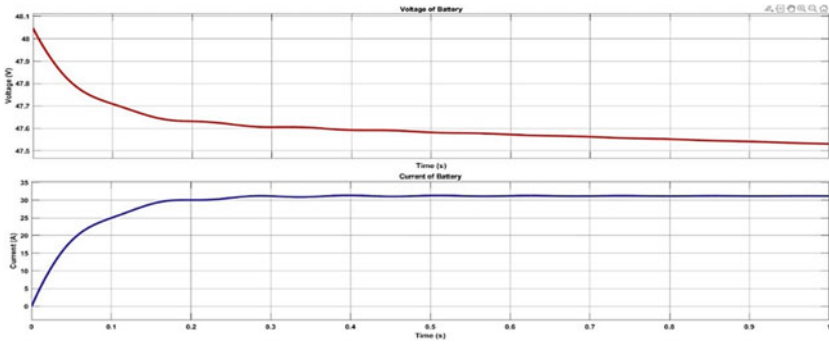


Fig. 7 Battery voltage and current waveforms

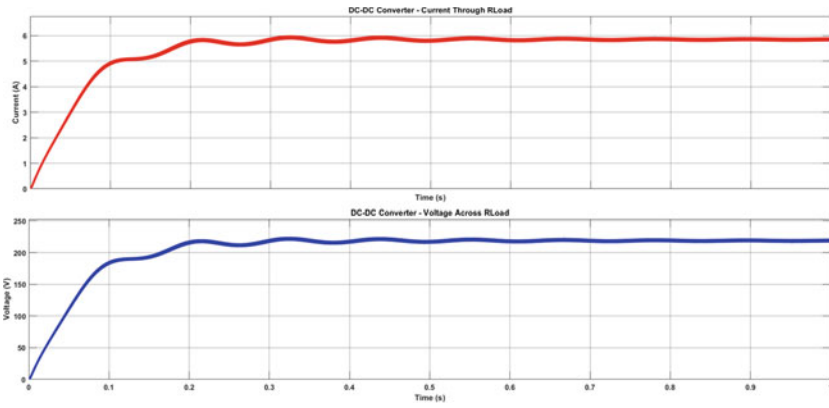


Fig. 8 SEPIC converter output current and voltage

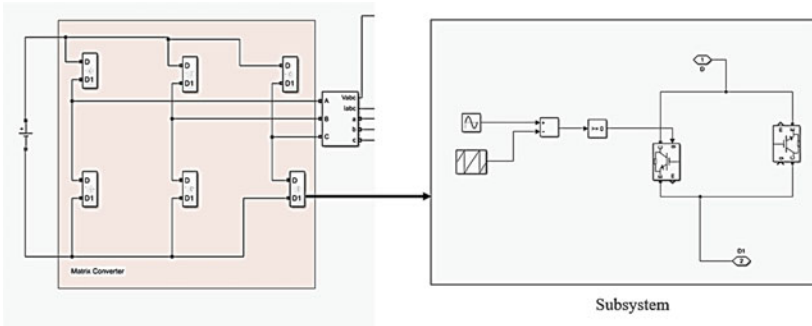


Fig. 9 Matrix converter simulation circuit

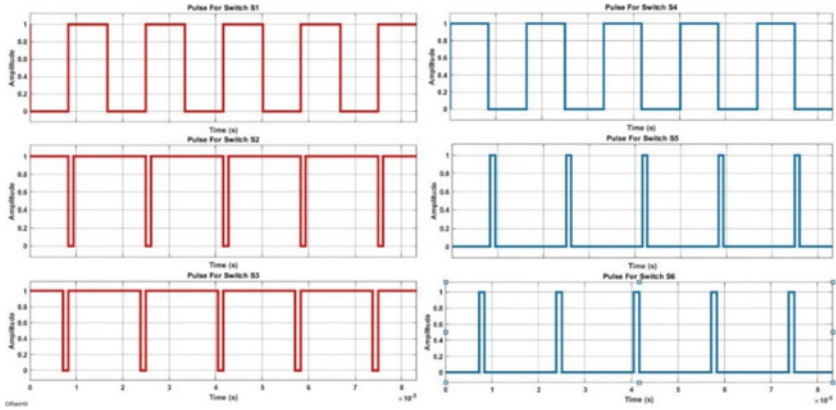


Fig. 10 Matrix converter switching pulses

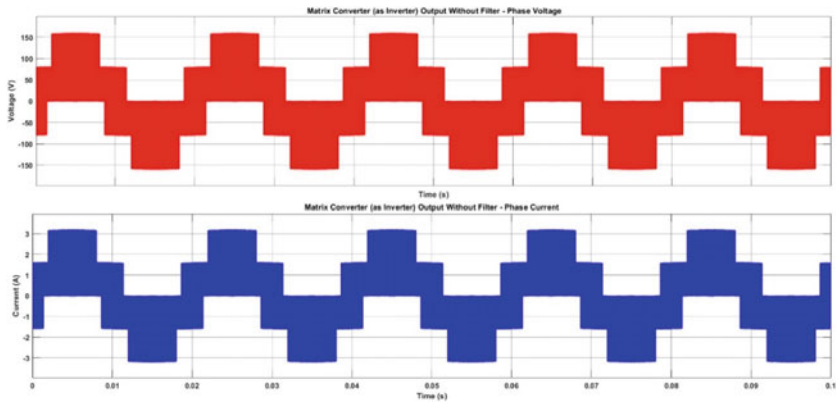


Fig. 11 Matrix converter output voltage and currents–phase

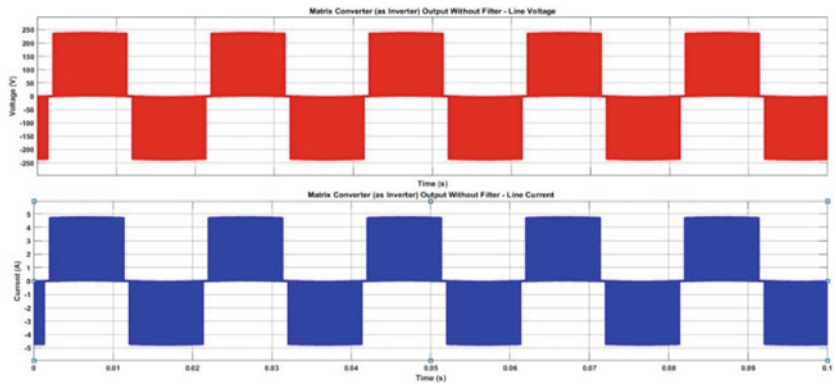


Fig. 12 Matrix converter output voltage and currents–line

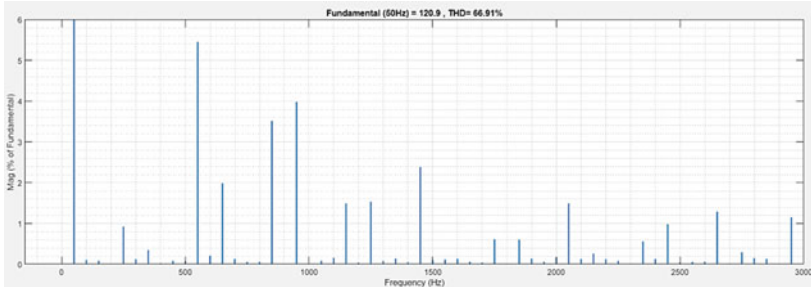


Fig. 13 FFT spectra of matrix converter output voltage without filter

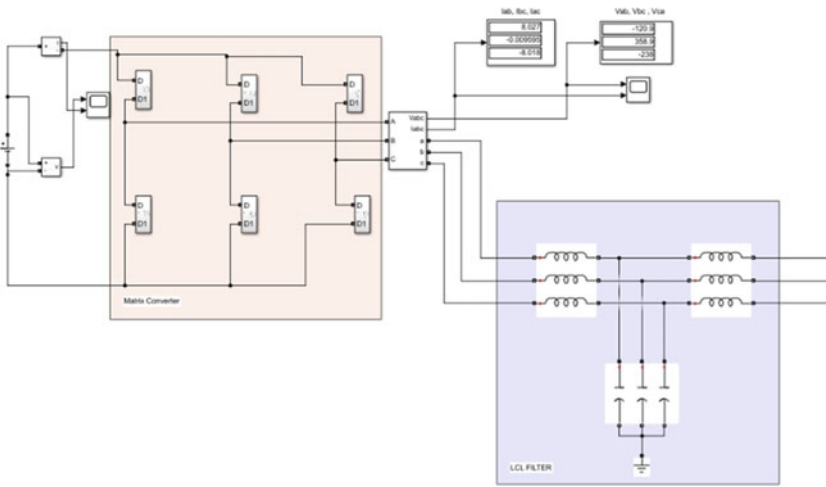


Fig. 14 Matrix converter with LCL filter

voltages and currents are shown in Figs. 15 and 16 respectively. FFT spectra with LCL filter is shown in Fig. 17. The THD is 6.77%. This voltage is injected to the grid.

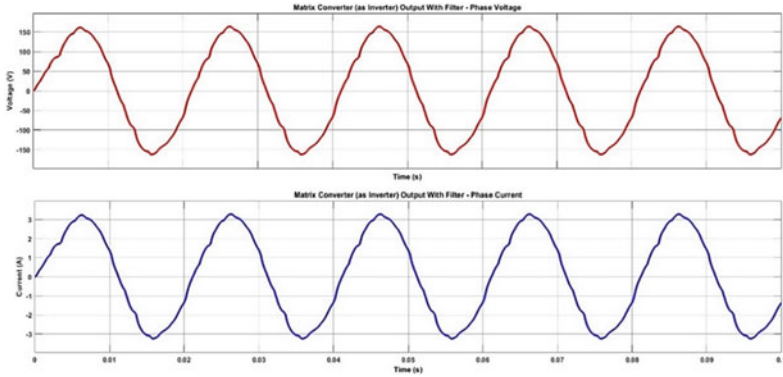


Fig. 15 Matrix converter output voltage and current with LCL filter–phase

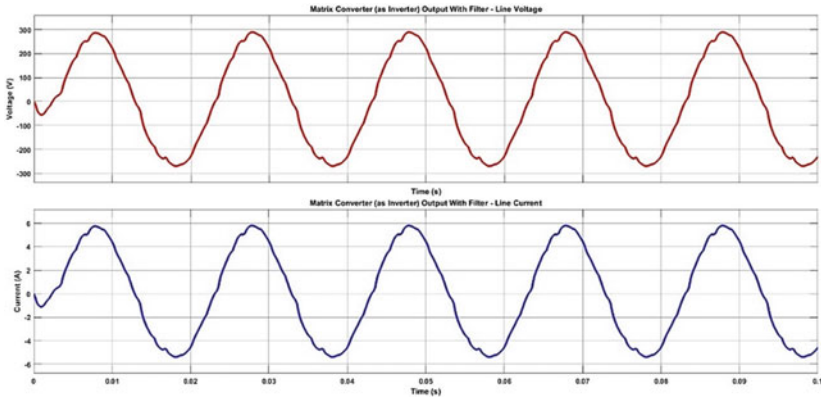


Fig. 16 Matrix converter output voltage and current with LCL filter–line

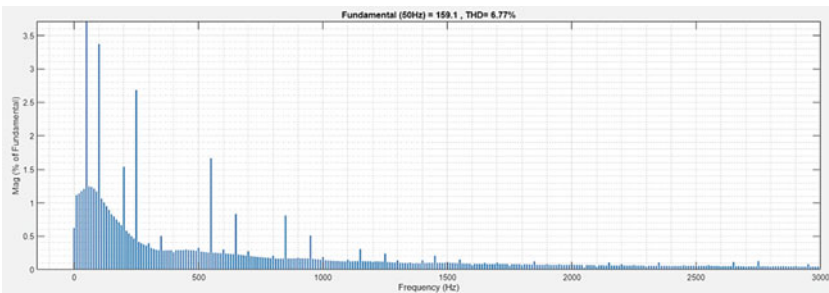


Fig. 17 FFT spectra of matrix converter output voltage with LCL filter

7 Conclusion

SEPIC converter with matrix converter was proposed in the paper for V2G application. Battery voltage of 48 was boosted to 240 V using SEPIC converter for forward mode of operation. This voltage is converted to AC voltage of 415 V using matrix converter. The THD in matrix converter was found to be 66.9% which was reduced to 6.77% using LCL filter. The work can be further extended to reduce the THD to less than 2%. The same converter configuration can be analysed for G2V mode and also fast charging technique can be implemented for battery charging.

References

1. Sai Shibu, N.B., Balamurugan, S., Arjun, D., Nidhin Mahesh, A. "Decentralized power system and future mobility: The use cases of community driven electric vehicle charging infrastructure" (2019) TESCA 2019 - Proceedings of the 2019 1st ACM International Workshop on Technology Enablers and Innovative Applications for Smart Cities and Communities, co-located with the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, ACM BuildSys, 2019, pp. 50–53.
2. Aki Vamsi Krishna & Rashmi M. R., "Wireless Power Transfer for LED Display System Using Class-DE Inverter", Lecture Notes in Electrical Engineering Volume 545, 2019, pp. 1271–1282
3. F. M. Shakeel and O. P. Malik, "Vehicle-To-Grid Technology in a Micro-grid Using DC Fast Charging Architecture," 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 2019, pp. 1–4.
4. S. Yahyazadeh, M. Khaleghi, S. Farzamkia and A. Khoshkbar-Sadigh, "A New Structure of Bidirectional DC-DC Converter for Electric Vehicle Applications," 2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), 2020, pp. 1–6.
5. Jadhav, M., & Kalkhambkar, V., "A Review on Plug-in Electric Vehicles: Services, Limitations and Impacts", *Majlesi Journal of Mechatronic Systems*, 7(2), 2018, pp.19-35
6. Pakhale, J.L., Mamatha, I., "Comparative Study of Different Current Control Techniques for Single Phase Grid Integrated Photovoltaic System", International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018, 2018, pp. 1259-1266.
7. Deepak, R. K. Pachauri and Y. K. Chauhan, "Modeling and simulation analysis of PV fed Cuk, Sepic, Zeta and Luo DC-DC converter," 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), 2016, pp. 1–6.
8. L. Yang, H. Liang and T. Liang, "Analysis and Implementation of a novel Bidirectional DC-DC Converter", *IEEE Transactions on Industry Applications*, vol. 59, no. 1, 2012, pp.422–434.
9. X. Xu, C. Zheng, C. Hu, Y. Lu and Q. Wang, "Design of Bi-directional DC-DC converter," 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), 2016, pp. 2283–2287
10. Dimna Denny C and Shahin M, "Analysis of bidirectional SEPIC/Zeta converter with coupled inductor," 2015 International Conference on Technological Advancements in Power and Energy (TAP Energy), 2015.
11. R. Kushwaha and B. Singh, "An Improved SEPIC PFC Converter for Electric Vehicle Battery Charger", *IEEE Industry Applications Society Annual Meeting*, 2019, pp.1–8.
12. Nikitha, B.N., Lekshmi, S. "A Four Switch Three-Port DC/DC/AC Converter Based on Power Decoupling Technique-Simulation Analysis", *Proceedings of the 2nd International Conference on Smart Systems and Inventive Technology, ICSSIT 2019*, pp. 412–418
13. J. S. Goud, K. R., K. R. and B. Singh, "Low Frequency Ripple Charging of Li-Ion Battery using Bidirectional Zeta DC-DC Converter to Improve Charge Transfer Efficiency," 2018 IEEE 7th International Conference on Power and Energy (PECon), 2018, pp. 62–66.

14. L. Rovere, S. Pipolo, A. Formentini and P. Zanchetta, "AC-DC Isolated Matrix Converter Charger: Topology and Modulation," 2020 IEEE Energy Conversion Congress and Exposition (ECCE), 2020
15. T. Premkumar, M. R. Rashmi, A. Suresh and D. R. Warriar, "Optimal voltage balancing method for reduced switching power losses in stacked multicell converters," 2017 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, 2017
16. Thomas, N., Jayabarathi, R., Nambiar, T.N.P., "Effect of Line Impedance and Loading on Voltage Profile in Distribution Network with Distributed Solar Photovoltaic System", 3rd International Conference on Communication and Electronics Systems, ICCES 2018, pp. 962–968.
17. N. Islam, A. Mohammed Arfi, M. A. Uz Zaman, A. Baul and A. Mohammed Abir, "Design and Implementation of a Matrix Converter Using Direct IGBT Switching Method," 2019 IEEE International Conference on Power, Electrical, and Electronics and Industrial Applications (PEEIACON), 2019, pp. 69–73.
18. Balamurugan S., Nageswari S., "Circulating current control of modular multilevel converter by wild spider foraging optimization based fractional order proportional integral derivative controller", *Journal of Intelligent and Fuzzy Systems*, 41 (2), 2021, pp. 4127-4147.
19. Nidhal Mдини, Skander Mustapha S. & Slama Belkhdja I., "Design of passive power filters for battery energy storage system in grid connected and islanded modes", *Springer Nature Switzerland Applied Sciences*, No.2:933, 2020.
20. Smys & Wang, Haoxiang, "Security Enhancement in Smart Vehicle Using Blockchain-based Architectural Framework", *Journal of Artificial Intelligence and Capsule Networks*. 3., 2021, PP. 90-100.
21. Karuppusamy, P. "Design of Adaptive Estimator for Nonlinear control system in Noisy Domain." *Journal of Electronics and Informatics* 3, no. 3 (2021): 150-166.
22. Sathesh, A. "Metaheuristics Optimizations for Speed Regulation in Self Driving Vehicles." *Journal of Information Technology and Digital World* 2, no. 1 (2020): 43-52.
23. Balasubramaniam, Vivekanadam. "Fault Detection and Diagnosis in Air Handling Units with a Novel Integrated Decision Tree Algorithm." *Journal of Trends in Computer Science and Smart Technology* , Volume 3, No. 1, 2021, 49-58.

Creation of SDIoT Testbed for DDoS Attack Using Mininet: Experimental Study



B. Keerthana, Mamatha Balachandra, Harishchandra Hebbar, and Balachandra Muniyal

Abstract Distributed Denial of Service (DDoS) attack is one of the biggest threats in Software Defined Internet of Things (SDIoT) environment. Here, we developed a proof-of-concept SDIoT testbed for DDoS attacks using Mininet Emulator. DDoS attack makes the victim's system unavailable by exhausting the system resources. For testing and proposing a solution for a DDoS attack in a SDIoT environment a testbed must be created. In our study, the testbed was created using VMware workstation, Mininet emulator, OVS switch, and Pox controller. Hping 3 tool for generating attacks, python scripts for custom topologies, Shell script for traffic capturing were used. With varying payloads, RTT (Round Trip Time), throughput and packet loss were calculated. The experimental result show that Mininet is feasible and convenient, and we conclude with a comparison between simulators and emulators providing SDIoT, along with a summary of advantages and disadvantages of using Mininet emulation in a SDIoT environment.

Keywords DDoS attack · Emulator · Mininet · Packet loss · RTT · SDIoT · Testbed · Throughput

B. Keerthana · H. Hebbar

Manipal School of Information Sciences, Manipal Academy of Higher Education, Manipal, Karnataka, India

e-mail: keerthana.s@manipal.edu

M. Balachandra (✉)

Department of Computer Science and Engineering, Manipal Academy of Higher Education, Manipal, Karnataka, India

e-mail: mamatha.bc@manipal.edu

B. Muniyal

Department of Information & Communication Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India

e-mail: bala.chandra@manipal.edu

1 Introduction

With the advancement of technology and communication networks, IoT devices play a major role in our day to day activities. Traditional networks face issues like flexibility, programmability, scalability to name a few. In order to overcome these issues, several solutions have been proposed by various researchers and industrial experts and Software Defined Network (SDN) is one of them. SDIoT is the integration of SDN and IoT devices, which provides a promising result and helps to overcome the problems faced in traditional networking. On the contrary, it has its own security issues. One of them being the DDoS attack, which aims at exhausting the resources of the victim's system and disrupt the services provided by that system. Owing to the severity of this issue, there were many solutions proposed to mitigate this problem [1, 2]. One of the major requirements for creating and testing the solutions would be the creation of a testbed. Using a testbed, the different scenarios required for the study can be created and various cyber-attacks could be simulated without the risk of affecting a real network and any of the devices in it. The testbeds help in the better understanding of the situation and the scenario used in the study. The main aim of our study, to create a testbed and develop a proof-of-concept in Software Defined networking testbed for DDoS attack using Mininet emulator. Research Questions–Q1. Is there any relation between number of packets and Round Trip Time. Q2. Is there any relation between number of attackers and packet loss. The QoS parameters chosen for the study were throughput and Round Trip Time (RTT). Understanding the characteristics of the components of an attack and the network involved is important. The output from the study were compiled and quantified, from the Mininet emulator. For this study, a two tier switch network was created with four subnets and number of nodes in the range of 1 to 20. In the scenario created, both legitimate traffic and attack traffic of varying sizes were used. Using the above mentioned setup and the work environment, the data collected was tabulated and formatted for further analysis.

2 Existing Testbeds

The requirement of a dedicated testbed for understanding the intricacies of attacks and strategies for their mitigation is essential. Even though there are various tools available for running such attacks, simulation of a large scale network with a complex system, in a small laboratory is in dire need. There is evidence in the literature suggesting that, there exist large botnets with millions of machines connected to each other that might prove to be smaller than estimated [3]. In the past decade, there have been many attempts to broaden the field of SDN [4, 5]. There are also efforts in this field to increase the robustness of the control plane in SDN [6–8].

Testing Distributed Denial of Service in SDIoT network to study the vulnerabilities, attacks characteristics and its effect to provide solution for the attack by exploring research experiments is a key factor for better understanding. Generating

Table 1 Comparison of simulators and emulators supports SDN-IoT

| Tools | License | M | OF | Sd-IoT | Wi-fi | Gui | Cli | Program | OS | RR | Free | Scale |
|---------------|----------------------------------|-----|-----|--------|-------|-----|-----|--------------|----|----|------|-------|
| Mininet | BSD | E | 1.4 | Y | Y | Y | Y | Python | Y | N | Y | M |
| Maxinet [10] | University of Paderborn, Germany | E | 1.3 | Y | N | Y | Y | Python | Y | N | Y | L |
| Netsim | TETCOS | S&E | 1.3 | Y | Y | Y | Y | Python | N | N | N | L |
| DOT [11] | NA | E | 1.3 | - | - | Y | Y | C+ + | Y | - | Y | L |
| GNS3 | GNU GPL | E | 1.3 | - | N | Y | Y | Python | Y | N | Y | L |
| Ns3 | GNU GPL | S&E | 1.3 | Y | - | Y | Y | C+ +, Python | Y | Y | Y | L |
| OpenNet [12] | GNU GPL | S&E | 1.3 | Y | Y | - | - | C+ +, Python | Y | N | Y | M |
| Estinet | NA | S&E | 1.3 | Y | Y | Y | Y | C/ C+ + | Y | N | N | L |
| Packet Tracer | Cisco | S | - | - | Y | Y | - | Python | N | N | Y | L |
| Omnet++ [13] | OpenSim Ltd | S | 1.3 | Y | Y | Y | N | C+ + | Y | Y | Y | L |
| Core [14] | BSD | E | - | - | Y | Y | - | Python | Y | N | Y | L |
| Cooja | BSD | S | - | Y | Y | Y | - | Python | Y | N | Y | L |

Abbreviations: OS—Open Source RR—Result Repeatable M –Mode NA- Not Available Y-Yes N-No L-Large M-Medium BSD—Berkeley Standard Distribution GNU–GNU General Public License

DDoS attack on live network is not feasible as it might lead to network chaos and there is no standard method to conduct DDoS attack. There are four different methods to construct a testbed viz., Mathematical Model, Hardware, Simulation and emulation methods [9], each model has its own pros and cons (Table 1).

2.1 Mathematical Model

Creating attack model is important step to understand and evaluate the effectiveness of the attack. Mathematical Models are theoretical in nature in which the systems are represented symbolically and then validated mathematically. Researchers have used mathematical model of DDoS attack with different features such as bandwidth, memory and buffer size of the server. They have used unsupervised learning algorithm to form clusters and applied mathematical correlations, and normal distributions on the clusters to understand the behavior of DDoS attack [15]. In another work, the authors had developed a mathematical model to differentiate DDoS attacks from the flash crowds using flow entropy [16]. In another work by Kriti et al. [17], they used mathematical model to represent the flow table space of a switch by using queuing theory and proposed flow-table sharing method to protect Software Defined based cloud from the Distributed Denial of Service attack. The mathematical model of detection method using fuzzy synthetic is represented by Qiao et al. [18]. Juan et al., summarized various mathematical models with respect to DDoS [19]. This model helps us to understand the behavior of malicious activity and to check the model validity. Simulated environments help us to improve the mathematical model [20].

2.2 Simulation

Through simulation the networks are modeled using a software environment. Compared to real device, simulation is safer and this also allows to analyze the behavior. In this work [21] authors have used Mininet simulated environment and generated DDoS attack using hping3 tool. Another work [22] also used Mininet to create multiple subnet and Xming applications and used to access remote client application. In another study, the authors used OMNET++ to implement SDN and IoT network. Author Ying et al. [23] used SDN-WISE-CONTIKI and Mininet to simulate SDIoT environment. Examples are OMNET++, NS3, and so on. The drawback of simulation is that the simulated router and operating systems are not modeled to undertake the stress of DDoS attack. For this reason, many researchers prefer using an emulator to do their research.

2.3 Emulation

There are various emulators available in research community as both commercial and open source options. In emulators, the flows are real-time, cost effective, slow and do not provide an accurate result. Example of emulators are NS3 and Estinet. According to our survey, with respect to emulators in SDN, most of the researchers used Mininet emulator. Each emulator has its own advantages and disadvantages. Table 2 shows the summary of possible emulators and simulators with respect to SDIoT.

In our study, of the above mentioned methods, we have chosen the emulation method for the creation of testbed and running the experiments.

Table 2 Software and hardware specifications

| Descriptors | Tools |
|-----------------------|---|
| Host CPU | Intel i5-7200 CPU @2.50 GHz 2.71 GHz |
| Host memory | 16 GB RAM |
| Host operating system | Windows |
| Virtualization | Vmware |
| Guest OS | Ubuntu 20.04.2 LTS |
| Emulator | Mininet 2.3.0 |
| Switch | ovs 2.13.3 |
| Controller | Pox 0.7.0 |
| Programming | Python 3/bash script/awk |
| Generating attack | Hping3 |

2.4 Hardware

Development of hardware testbed is fast, has higher fidelity, accurate, shared and expensive. In a study [24], the authors used Raspberry Pi to install OpenFlow and converted the device into an SDN switch. In another work [25], authors used real SDIoT testbed using RaspberryPi as IoT device and controller and Zodiac-Fx as switch, which is low cost device to detect and mitigate DDoS attack. Testing DDoS attack on these environments might lead to disruption of the real network, if attached to one. To overcome this issue, building the network system on an isolated environment testbed like PlanetLab, GENI, Emulab, DETER, which proves secure and isolated environment. These systems could also be implemented on virtual machines and the other possibility would be the use of a simulator.

3 Requirements in the Architecture and Design of a Testbed

Testbeds are created for the understanding of the function and characteristics of a known DDoS attack, and the one created here was specifically designed for this requirement. In this section, the requirements for the setup of the testbed and the details of the hardware and software used for the setup are described.

3.1 Requirements for the Setup

See (Table 2).

3.1.1 Customize the Network

A Testbed created should be able to simulate the real-world conditions in order to test the various contexts based on the available types of networks and their infrastructure. Each and every network has its own design and testing needs to be done, while considering the network as a whole. The advantage of emulators has paved for customizations in the network created which can mimic a real network scenario. Here we have used mininet emulator for creation of the network. This testbed allows a user to create custom topology based on the requirements and to customize the network parameters such as link bandwidth, delay, queue size and packet loss.

3.1.2 Provide Variety of Attack

Security plays a major role in any network and acts as one of the factors contributing to its efficiency network. While creating a testbed, the possibility of incorporating various aspects of security needs to be considered. The capability of detecting various vulnerabilities of SDIoT should be presented within the created network and these could be prone to attacks. In such testbeds, attack types such as Man-in-the Middle attack, DDoS attack, spoofing attack, Masquerade attack, etc., should be possible.

3.1.3 Rerun the Experiment

Another factor to consider would be the capability of re-running the experiment in the created network and the testbed should be capable of supporting this. It should also be able to save the experiment and replay the experiment for further improvement, along with deploying, running and stopping the experiment.

3.1.4 Traffic Generation

A testbed should be able to generate traffic flows that represents realistic user application and the required scenario. It should provide both normal and attack traffic. For the generation of traffic, the testbed should also be capable of supporting various tools that are used for the generation of the different kinds of traffic.

3.1.5 Capable to Capture the Traffic

Once the testbed is created and the traffic generation is initiated, the next step would be to study the behaviour and the characteristics of the network. This is done by capturing the traffic flow through the network. This is another important feature that is very important for further analysis and future references in troubleshooting.

3.1.6 Capable to Program and Customize

Generation of traffic in a created testbed environment includes customization and creativity that fits the requirements [26–28]. This can be achieved by programming within the testbed.

3.1.7 Monitor the Traffic

While running network traffic through a real network or through a testbed, it is essential that the traffic is monitored on a continuous basis. There are various techniques and tools that have been reported by many researchers around the world [29–31]. Choosing the appropriate monitoring system plays a critical role in the security of a network. With recent advances in networks and the adaptability of these networks, monitoring becomes challenging and accordingly, tools are being developed for efficient monitoring. The huge amounts of data generated from these networks are captured. Programs capable of handling such enormous data, such as data mining and machine learning techniques are used for the active monitoring of the networks [32]. In networks requiring real-time monitoring and handling sensitive data, programs created as said above, play a major role.

3.1.8 Analysing the Evaluation Metrics

Analysis of the network metrics could be simple in smaller networks but it could be very important in data sensitive networks such as financial institutions and healthcare networks. In such critical networks, the real-time analysis of the data generated would add very much to the security of the network and could trigger an alarm in case of any malicious traffic or just a change in the pattern of traffic flow. Intelligent systems using machine learning could read and categorize the traffic based on their learning patterns, which prove to be statistically important considering the significance [33].

3.2 *Hardware and Software Requirements*

3.3 *Experimental Environment and Setup*

Test-bed [h1-h20 = hosts, s1-s6 = switches, c0 = controller].

As shown in Fig. 1, customized network was created using python code and Mininet emulator. It consists of six ovs switches are connected to the central POX controller. These switching nodes are linked together by emulated wired Ethernet interfaces. There are 20 user nodes/Edge Devices, named as h1 to h20 with the IP addresss 10.0.0.1 to 10.0.0.20. Assuming that some hosts are compromised by the attacker. The first part is to calculate the Round Trip Time and throughput before attack and launch the attack through attacker nodes. Sequentially increase the attacker node from 1 to 4 and analyse the packet loss. Round-trip time, throughput and packet loss are obtained respectively. No SDN rules have been defined to drop or counter measure the attack.

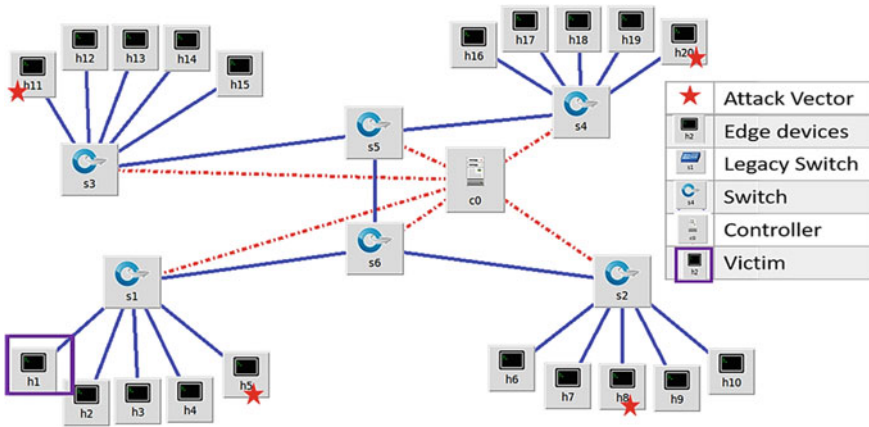


Fig. 1 Screenshot showing the network created for SDIoT

3.4 Testbed Setup

Step 1: creation of the customized network.

Created custom using python to create custom network. The pictorial representation of customized environment is shown in the Fig. 1, created using MiniEdit. It is graphical user interface editor to create customized network. Execute the custom code by typing.

```
sudo python mytopo.py.
```

The custom code is stored in mytopo.py.

Step 2: Enabling the controller.

To enable the controller, execute the command.

```
cd pox.
```

The above command is to change the current directory to pox where pox controller is installed. Enable pox controller by typing.

```
./pox.py.
```

Step 3: Generate traffic with various payload, collect the result.

In this step, hping3 tool was used to generate attack traffic with ICMP protocol and the results were collected. Created custom bash code to generate traffic and collect the result. Classification of nodes h1, h5, h6, h10 as attackers and h15 as victim. Launching of the attack by flooding the traffic from attacker’s node to victim’s node using xterm.

```
hping3 -icmp -d number of bytes of data.
```

Step 4: Analyse the result.

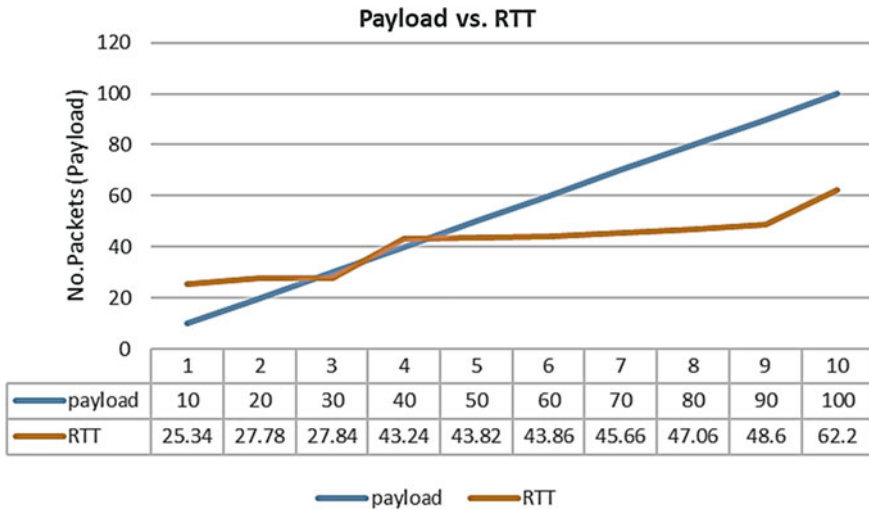


Fig. 2 Payload versus RTT

A shell script using grep, sed and awk utility was used to extract the relevant features for analysis. From the above setup and functions, the aim of the study was to address the following questions.

- Q1. Is there any relation between number of packets and Round Trip Time.
- Q2. Is there any relation between number of attackers and packet loss.

To answer the Research Question1, conducted experiment by changing various payload and comparing with their RTT. An attempt was made to look into the relation between the number of packets and the Round Trip Time (RTT), if any. As shown in Fig. 2, it is clear that there exists a positive correlation between the payload (number of packets) and RTT. The graph showed a rise in RTT as there was an increase in the number of packets.

To answer the Research question 2, we conducted experiment with various payload and calculated their throughput. Based on the results as shown in Fig. 3, it was seen that there was a positive correlation between the payload and throughput. The graph clearly shows that there is an increase in the throughput as the payload is increased. The payload, i.e., if there is an increase in the number of packets, there is an appropriate increase in the throughput.

In Fig. 4, it is seen that there is an increase in the number of packets lost as against an increase in the number of attackers.

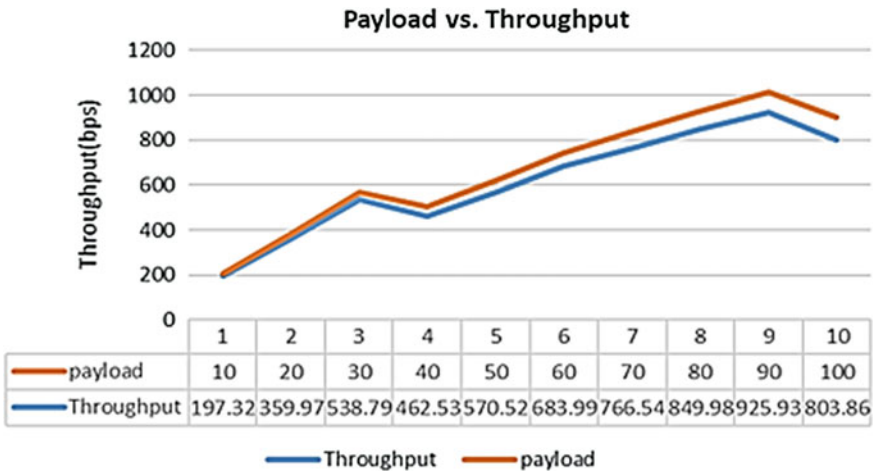


Fig. 3 Throughput versus payload

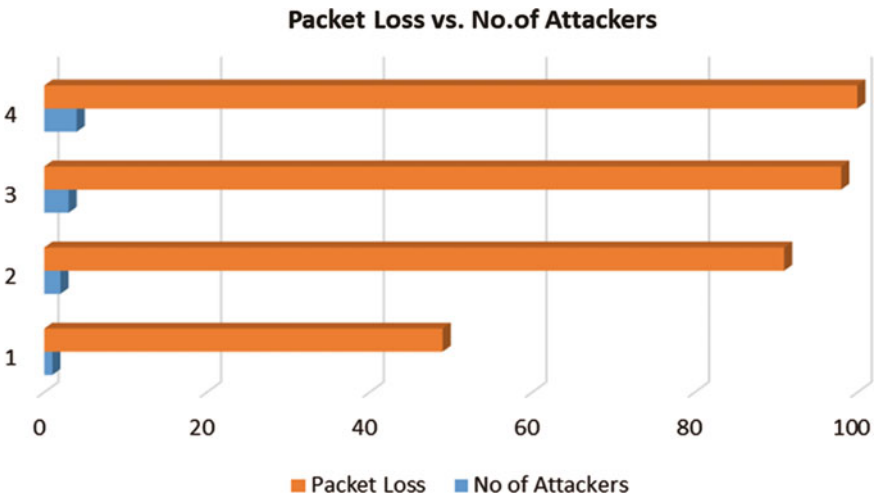


Fig. 4 Packet loss versus no. of attackers

3.5 Algorithm for Creating Testbed Setup

The summarized steps to conduct the experiment are given below:

- Step 1: creation of Python code to create custom network.
- Step 2: Enable pox controller.
- Step 3: Generate traffic with various payload, collect the result.
- Step 4: Launch attack.
- Step 5: Collection of raw data.

Step 6: Visualizing the normal flow in the network created before the attack.

Step 7: Visualizing the change in traffic flow in the network after the attack.

Step 8: Analyze the result.

4 Results

All the above scenarios show that as there is an increase in the consumption of the available resources, the parameters chosen in this study have a direct correlation to the type of input that consumes the resources. The increase in RTT, throughput and the loss of packets could be used as parameters as in studies against IDS studies.

5 Advantages and Disadvantages of SDIoT Environment in Mininet

The advantage of using Mininet is that it is open source with very good documentation. It is possible through command line to create customized topologies or use the built in topologies, by using python code or miniedit which is graphical user interface mode. Easy to use and implement. Suitable for creating medium level network. It is possible to work with wireless application using sdn-wifi. Suitable for understanding, testing and research purpose. Many researchers use Mininet.

The drawback of using Mininet is that, there is a need of heavy weight virtualization software like VMware required to install Mininet, Many additional configurations are need to use multiple vlan. It is complicated for testing complex network. The main drawback is that; result is not repeatable. To overcome the problem, run the experiments for multiple times and take average of it.

6 Proposed Solutions

Through the above experiment, provided proof of concept to implement the possibility of DDoS attack in SDIoT environment. Intrusion Detection System and studies related to DDoS attacks are continuously done by industrial researchers and academicians. There are various methods to detect DDoS attack. Attack Detect techniques are broadly classified into Signature based and Anomaly based techniques. Signature-based approaches can detect the intrusions whose attack patterns are already stored in the database. For every attack, its signature is to be created. Attacks whose patterns are not present in the database cannot be detected. It fails to handle the attacks with human interference or attacks with inherent self-modifying behavioral characteristics [34]. Signature-based systems have attack patterns defined in them. These systems

cannot detect zero-day attacks. It is also called as Misuse Detection. Anomaly detection involves the collection of data relating to the behavior of legitimate users over a period of time while Signature or Heuristics detection uses a set of known malicious data patterns (signatures) or attacks rules. Anomaly detection is broadly classifying into Statistical based, Knowledge based and Machine learning based. To handle the unknown attack use of machine learning technique could be very useful [35]. Hidden Markov model are used to detect low rate DDoS attacks [36]. In this research work [37] author proposed a novel framework to combine both traffic analysis and security monitoring in the mobile devices. This report summarizes the various machine learning algorithms used for the detection of DDoS attacks [38]. Based on literature, we hereby propose combining both statistical and machine learning provide better results. In order to mitigate such attacks, it can be programmed in such a way that the entire setup can be reprogrammed to drop the attack traffic, once it is detected. Another way of handling could be by the reduction of the bandwidth of the entire traffic passing through the setup.

7 Conclusion

This paper reports on our preliminary development of a network emulator testbed. The emulated network is software defined and controllable by the central POX-based SDN controller. The network constructs consist of virtual hosts, representing both normal users and attackers, OVS-based virtual switches, and POX as a controller. The goal of this testbed is to enable us to study the mechanism, enabled by SDN programmability, against a DDoS attack. All the emulated components in the testbed can be accommodated within a single personal computer conveniently. Based on the experiment result it is feasible and convenient to conduct DDoS attack testbed using Mininet. In future the testbed can be extended to demonstrate its usefulness in a scenario of thousands of ping processes with the largest possible payload size in SDIoT environment.

References

1. Letteri, I., G.D. Penna, and G.D. Gasperis: Security in the internet of things: botnet detection in software-defined networks by deep learning techniques. *International Journal of High Performance Computing and Networking*. 15, 3-4, 170-182 (2019)
2. Allakany, A., et al. Detection and Mitigation of LFA Attack in SDN-IoT Network. in *Workshops of the International Conference on Advanced Information Networking and Applications*. 2020. Springer.
3. Van Eeten, M., et al. The role of internet service providers in botnet mitigation an empirical analysis based on spam data. 2010. TPRC
4. Koponen, T., et al. Onix: A distributed control platform for large-scale production networks. in *OSDI*. 2010.

5. Gupta, A., et al.: Sdx: A software defined internet exchange. *ACM SIGCOMM Computer Communication Review*. 44, 4, 551–562 (2014)
6. Berde, P., et al. ONOS: towards an open, distributed SDN OS. in *Proceedings of the third workshop on Hot topics in software defined networking*. 2014.
7. Tootoonchian, A. and Y. Ganjali. Hyperflow: A distributed control plane for openflow. in *Proceedings of the 2010 internet network management conference on Research on enterprise networking*. 2010.
8. Yan, X., X. Hu, and W. Liu: SDN Controller Deployment for QoS Guarantees in Tactical Ad Hoc Networks. *Wireless Communications and Mobile Computing*. 2021, (2021)
9. Behal, S. and K. Kumar: Trends in validation of DDoS research. *Procedia Computer Science*. 85, 7-15 (2016)
10. Wette, P., et al. Maxinet: Distributed emulation of software-defined networks. in *2014 IFIP Networking Conference*. 2014. IEEE.
11. Roy, A.R., et al. Dot: distributed openflow testbed. in *Proceedings of the 2014 ACM Conference on SIGCOMM*. 2014.
12. Chan, M.-C., et al. OpenNet: A simulator for software-defined wireless local area network. in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. 2014. IEEE.
13. Varga, A. and R. Hornig. An overview of the OMNeT++ simulation environment. in *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*. 2008.
14. Ahrenholz, J., et al. CORE: A real-time network emulator. in *MILCOM 2008-2008 IEEE Military Communications Conference*. 2008. IEEE.
15. Johnson Singh, K. and T. De: Mathematical modelling of DDoS attack and detection using correlation. *Journal of cyber security technology*. 1, 3-4, 175–186 (2017).
16. Tao, Y. and S. Yu. DDoS attack detection at local area networks using information theoretical metrics. in *2013 12th IEEE international conference on trust, security and privacy in computing and communications*. 2013. IEEE.
17. Bhushan, K. and B.B. Gupta: Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. *Journal of Ambient Intelligence and Humanized Computing*. 10, 5, 1985–1997 (2019).
18. Yan, Q., Q. Gong, and F.-a. Deng: Detection of DDoS attacks against wireless SDN controllers based on the fuzzy synthetic evaluation decision-making model. *Adhoc & Sensor Wireless Networks*. 33, (2016).
19. Balarezo, J.F., et al.: A survey on DoS/DDoS attacks mathematical modelling for traditional, SDN and virtual networks. *Engineering Science and Technology, an International Journal*. (2021).
20. Saini, D.K.: Cyber defense: mathematical modeling and simulation. *International Journal of Applied Physics and Mathematics*. 2, 5, 312 (2012)
21. Long, Z. and W. Jinsong: A Hybrid Method of Entropy and SSAE-SVM Based DDoS Detection and Mitigation Mechanism in SDN. *Computers & Security*. 102604 (2022).
22. Prabakaran, S., et al.: Predicting Attack Pattern via Machine Learning by Exploiting Stateful Firewall as Virtual Network Function in an SDN Network. *Sensors*. 22, 3, 709 (2022)
23. Wang, J., Y. Liu, and H. Feng: IFACNN: efficient DDoS attack detection based on improved firefly algorithm to optimize convolutional neural networks. *Mathematical Biosciences and Engineering*. 19, 2, 1280–1303 (2022).
24. Gupta, V., K. Kaur, and S. Kaur, Developing small size low-cost software-defined networking switch using raspberry pi, in *Next-Generation Networks*. 2018, Springer. pp. 147–152.
25. Smys, S. and W. Haoxiang: A Secure Optimization Algorithm for Quality-of-Service Improvement in Hybrid Wireless Networks. *IRO Journal on Sustainable Wireless Systems*. 3, 1, 1–10 (2021).
26. Bakhshi, T.: State of the art and recent research advances in software defined networking. *Wireless Communications and Mobile Computing*. 2017, (2017).
27. Farhady, H., H. Lee, and A. Nakao: Software-defined networking: A survey. *Computer Networks*. 81, 79–95 (2015).

28. Nunes, B.A.A., et al.: A survey of software-defined networking: Past, present, and future of programmable networks. *IEEE Communications surveys & tutorials*. 16, 3, 1617–1634 (2014).
29. Keim, D.A., et al. Monitoring network traffic with radial traffic analyzer. in 2006 IEEE symposium on visual analytics science and technology. 2006. IEEE.
30. Cecil, A.: A summary of network traffic monitoring and analysis techniques. *Computer Systems Analysis*. 4–7 (2006).
31. D’Alconzo, A., et al.: A survey on big data for network traffic monitoring and analysis. *IEEE Transactions on Network and Service Management*. 16, 3, 800–813 (2019).
32. Jose, A.S., L.R. Nair, and V. Paul. Data mining in software defined networking-a survey. in 2017 International Conference on Computing Methodologies and Communication (ICCMC). 2017. IEEE.
33. Cui, L., F.R. Yu, and Q. Yan: When big data meets software-defined networking: SDN for big data and big data for SDN. *IEEE network*. 30, 1, 58–65 (2016).
34. Jyothsna, V. and K.M. Prasad, Anomaly-Based Intrusion Detection System, in *Computer and Network Security*. 2019, IntechOpen.
35. Sathesh, A.: Enhanced soft computing approaches for intrusion detection schemes in social media networks. *Journal of Soft Computing Paradigm (JSCP)*. 1, 02, 69–79 (2019).
36. Mugunthan, S.: Soft computing based autonomous low rate DDOS attack detection and security for cloud computing. *J. Soft Comput. Paradig.(JSCP)*. 1, 02, 80–90 (2019).
37. Sree Lekshmi, S., et al., Data optimization-based security enhancement in 5G edge deployments, in *Innovative Data Communication Technologies and Application*. 2021, Springer. pp. 347-362.
38. Mishra, P., et al.: A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys & Tutorials*. 21, 1, 686–728 (2018).

Automatic Eye Disease Detection Using Machine Learning and Deep Learning Models



Nouf Badah, Amal Algefes, Ashwaq AlArjani, and Raouia Mokni

Abstract Glaucoma is a serious eye disease that affects a lot of people around the world. Deep learning architectures have been widely used in recent years for image recognition tasks. In this paper, we aim to detect human eye infections of Glaucoma disease by firstly using different machine learning (ML) classifiers such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes (NB), Multi-layer perceptron (MLP), Decision Tree (DT) and Random Forest (RF), and secondly a Deep Learning (DL) model such as Convolutional Neural Network (CNN) based on Resnet152 model. The evaluation of the proposed approach is performed on the Ocular Disease Intelligent Recognition dataset. The obtained results showed that the RF and MLP classifiers achieved the highest accuracy of 77% in comparison to the other ML classifiers. On the other hand, the deep learning model (CNN model: Resnet152) provides an even better accuracy of 84% for the same task and dataset. Furthermore, we observe our best performing model produce competitive results in comparison to some state-of-the-art approaches.

Keywords Eye detect · ODIR · Machine learning · Deep learning · Glaucoma disease detection · Images recognition

1 Introduction and Related Work

Artificial Intelligence (AI) has significantly contributed to the advancement of optimal healthcare delivery. Its unique ability to recognize patterns in all forms of data has been widely recognised by both research and industry. In particular, AI has performed remarkably in pattern extraction from image data thus hugely impacting

N. Badah · A. Algefes · A. AlArjani · R. Mokni (✉)
College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University,
Al-Kharj, Saudi Arabia
e-mail: r.mokni@psau.edu.sa

R. Mokni
University of Sfax, Sfax, Tunisia

delivery of services in medicare which often requires or involves visual inspection. The task of Image recognition generally faces many challenges. These range from image resolution, image formatting, Image segmentation, and others. Besides these challenges, analysts, researchers, and developers lack domain knowledge that is often relevant in the analysis of medical data. To be more specific, domain expert knowledge is valuable in computational analysis in order to develop models whose performance is almost consistent with that of medical practitioners. For example, eye infections diagnosis analyzes several factors like patients' gender, age, other synchronous symptoms, and the patient's medical record too. Nevertheless, AI has still managed to overcome the challenges and it still has so much to offer in terms of enhancing the analysis process and the results. In this work, we focus on Glaucoma eye disease that can cause serious complications if not diagnosed in the early stages.

Glaucoma is a group of ocular disorders with multi-factorial etiology united by a clinically characteristic intraocular pressure- associated optic neuropathy. Glaucoma (in all its forms) is the leading cause of irreversible blindness worldwide [1, 2]

Several studies investigated the usage of ML and DL to detect eye diseases. For instance, Choi et al. [3] proposed a model based on MatConvNet (a MATLAB toolbox) used to implement CNN for computer vision applications. The authors applied a CNN to achieve multi-categorical classification of retinal diseases using Retina/STARE database. The data contained 10 categories that included nine retinal diseases and one normal retina. The authors stated that the best results (87.4% accuracy) were acquired when using Random Forest (RF) transfer learning based on VGG-19 architecture. The authors indicated that the DL model performance was certainly affected by the dataset's small size. Abbas et al. [4] conducted an experiment to develop the Glaucoma-deep system which is a diagnostic system that uses a pre-trained CNN model to diagnose Glaucoma eye disease. The model was trained on a selected 1200 eye fundus images collected manually from public resources. The evaluation of this model achieves an accuracy of 99%. However, the author states that the model should further be trained having achieved a low accuracy when tested on large- scale images. Soltani et al. [5] presented the glaucoma fuzzy expert system to Glaucoma disease early detection. This system was tested on a real dataset of ophthalmologic images of both normal and glaucomatous cases. The system first processed the Optic Nerve Head (ONH) images and applied the canny detector algorithm to detect the image's contours. After extracting the main parameters, the classification algorithm is applied based on fuzzy logic approaches. The system resulted in 98% accuracy. Issac et al. [6] proposed a method for the detection of glaucoma from digital fundus images. The method is based on image processing but unlike other methods, it used the discriminatory parameters of glaucoma infection as inputs to the learning algorithms. The method achieved 94.11% accuracy with 100% sensitivity. Rama Krishnan et al. [7] discussed the use of Higher-Order Spectra (HOS), Trace Transform (TT), and Discrete Wavelet Transform (DWT) features to build an automated glaucoma detection system. The system uses SVM classifiers with linear, polynomial order and Radial Basis Function (RBF). The system was trained and tested on images from the Kasturba Medical College, Manipal, India. The system resulted in 91.3% accuracy in classifying glaucoma and normal images.

Shantaiya et al. [8] proposed an automated method to extract Cup to Disc Ratio (CDR) and Ratio of neuroretinal Rim (NRR) from retinal images. The method was evaluated on 80 fundus images collected from different datasets and achieved 98.8% accuracy. Islam et al. [9] discussed the use of source and camera independent ophthalmic Disease Recognition from Fundus Image. To solve the disease detection problem, the authors implemented a unique method for detecting eight types of ocular diseases using a convolutional neural network (CNN). The datasets were collected from ODIR. The model was trained on 8000 images from 4000 cases and 1000 images from 500 cases used for testing. The evaluation of this model achieves an accuracy of 85%.

Gogineni et al. [10] proposed integrating ensemble into GoogleNet image classification technique. The authors have also used the YOLO technique for disease detection. They aimed at detecting the top five common eye diseases with higher accuracy. Users can upload a pic in a mobile or cloud application, and inbuilt AI algorithms will detect the type of eye disease with higher accuracy and thus offering prevention suggestions at an early stage without doctor intervention. The dataset contains 20 features for 5000 instances from UCI dataset. This study used the ANN algorithm and achieved an accuracy equal to 95.20%.

In Table 1, we provide a summary of prior work pertaining to the detection of eye diseases. As briefly discussed in the preceding paragraphs, we observe the superiority of DL architectures compared to other techniques, particularly the CNN-based architectures in the eye detection task. Nonetheless, the SVM appears quite competitive and in some cases outperforming the ANN like in the fundus images dataset, as shown in the presented work by Issac et al. [6]

Table 1 Summary of the research studies for eye disease detection

| Proposal | Model | Dataset | Accuracy (%) |
|--------------------------|-------------------------------------|---|-----------------------|
| Choi et al. [3] | CNN | Retina (STARE) | 87.4 |
| Abbas et al. [4] | CNN | Eye Fundus images | 99 |
| Soltani et al. [5] | Decision making algorithm | ophthalmologic images | 98 |
| Issac et al. [6] | SVM, ANN | fundus images | SVM: 94.11% ANN: 89.6 |
| Rama Krishnan et al. [7] | SVM | Digital Retinal Images | 91.67 |
| Shantaiya et al. [8] | Mean Threshold Morphological method | Retinal Images from FAU data library and MESSIDOR dataset | 98.8 |
| Islam et al. [9] | CNN | Ocular Disease Intelligent Recognition ODIR-5K | 85 |
| Gogineni et al. [10] | GoogleNet (CNN)-KNN,ANN,LR and DT. | UCI-Dataset-5000 images | 95.2 |

In this work, we investigated the detection of glaucoma disease using both Deep Learning (DL) and shallow Learning or Machine Learning techniques. DL techniques such as CNNs models have been widely used and they have paid off with outstanding results in image classification or recognition tasks. This paper uses the Ocular Disease Intelligent Recognition (ODIR) dataset containing fundus images and non-imaging data including age, gender, and di- agnostic words of 5000 patients. After preprocessing the images, we used some Machine learning models and CNN-based techniques (ResNet152 model) to build binary classifiers to recognize whether or not an image belongs to a person presenting with Glaucoma. We mitigate against overfitting concerns by applying the regularization technique.

The main contributions of this study are:

- Apply preprocessing steps using several preprocessing techniques on the raw data.
- Apply the regularization technique to overcome the problem of overfitting.
- Build binary classification models using both ML and DL architectures to automate the detection of eye disease in images.
- We perform an evaluation of our built models on the ODIR dataset.

The remaining sections of this study are organized as follows. Section 2 discusses the proposed approach. Section 3 illustrates the results of experiments and a comparison of the proposed approach with other existing studies. Finally, Sect. 4 displays the conclusion and presents some future directions.

2 Proposed Approach

ML and DL techniques are the most widely used by computer and data scientists in the data analysis, image recognition, medical imaging, and computer vision fields (e.g. [11–14]). ML is widely considered an Artificial Intelligence technique that essentially enables computer systems to learn from data and make accurate inferences and predictions. DL is a subset of ML that uses a multi-layered neural network to learn from vast amounts of data. A neural net work mimics the human brain’s functionality i.e. it contains a set of layers with neurons interconnected to perform complex mathematical operations with a final layer making predictions. Techniques in the AI domain effectively achieve tasks that range from extraction of hidden insights from diverse data, predictions in recommender systems, image and graphic detection etc. Example tasks include automated disease diagnosis from clinical data, automatic extraction of customer interests from multi-modal data in shopping application and entertainment sites, detection of fraudulent transactions from bank activity data, analysis of sentiments in product reviews from consumers (posted on social media sites, and product sites), and many more. In this paper, we adopt a neural model for binary classification in order to predict whether an eye has an infection or not. In this paper, we proposed an automatic eye disease detection using a binary classification system to distinguish between normal eyes and glaucoma, using both Machine Learning and Deep Learning models. The adapted method for the identification of eye diseases

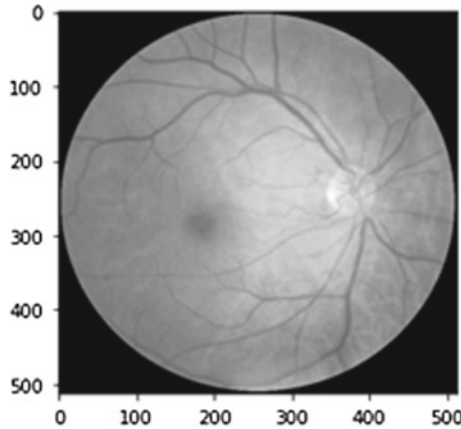


Fig. 1 A normal eye image before applying data preprocessing step

consists mainly of the following main phases: (1) Data collection, (2) data preprocessing, (3) Data Modeling using both Machine Learning and Deep Learning models and finally (4) Decision making.

2.1 Data Collection

The ODIR dataset image data is collected by Shang gong Medical Technology from Various hospitals and medical centers in China [15]. It consists of 6392 fundus images categorized under eight classes that include, Normal, Diabetes, Glaucoma, Cataract, Age-related Macular Degeneration, Hypertension, Pathological Myopia, Other diseases/abnormalities. The classes are represented by the first letter in the set. In this study, we selected example images with only two classes from this dataset such as Normal (N) and Glaucoma (G).

2.2 Data Preprocessing

The pre-processing pipeline includes resizing images to a uniform or base size . We adopt Gaussian smoothing to reduce the noise in the image i.e. enhancing the image structures at different scales as shown in Figs. 1 and 2

2.3 Data Modelling

Several techniques have been widely used for the purpose of image disease identification, in particular, Machine Learning (ML) and Deep Learning (DL) models.

Deep Learning Model. There are different types of neural networks that are used in DL such as Convolutional Neural Networks (CNNs), Recurrent Neu-ral Networks

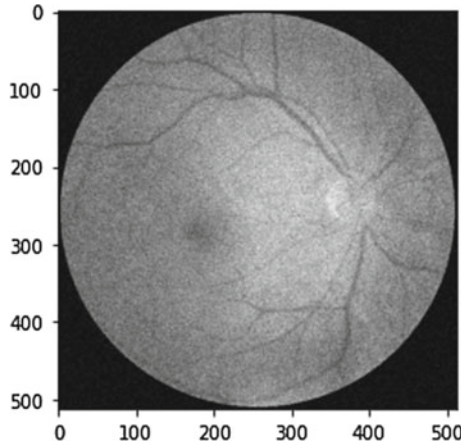


Fig. 2 A normal eye image after applying data preprocessing step

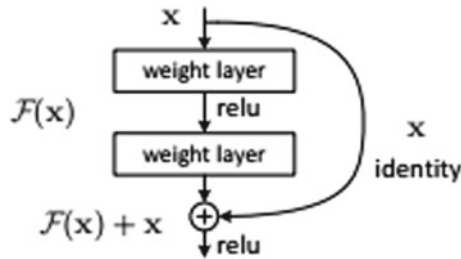


Fig. 3 Logical architecture scheme of ResNet framework

(RNNs), Long Short Term Memory (LSTM), etc. CNN’s exhibit extraordinary performance in image detection tasks because of their feature extraction ability given an image. They use a filter kernel to convolve across multi-dimensional data (representing an image) and they perform pooling which downsizes the feature map data retaining only important features [6]. ResNet is a deep learning framework used for image classification. ResNet alleviated the problem of vanishing gradients suffered when training very deep networks (i.e. ImageNet with 16–30 CNN layers). The scheme of logical blocks for ResNet architecture is shown below in Fig. 3 The ResNet model allows training deep neural networks with 152 layers.

ResNet uses the skip connections approach to ensure that the higher layers will perform at least as the lower layers and not worse. Figure 4 shows the implemented model architecture. Tensorflow and Keras platforms were used to implement a sequential model with a 0.01 learning rate. ReLU activation was used for both the input and hidden layers, whereas the output layer had a Sigmoid activation. The combination of the two activation methods enhanced the performance of our model.

We used a cross entropy loss objective to train the model for 600 Epochs using 32 as a batch size. To improve this model, we used dropout regularization after each

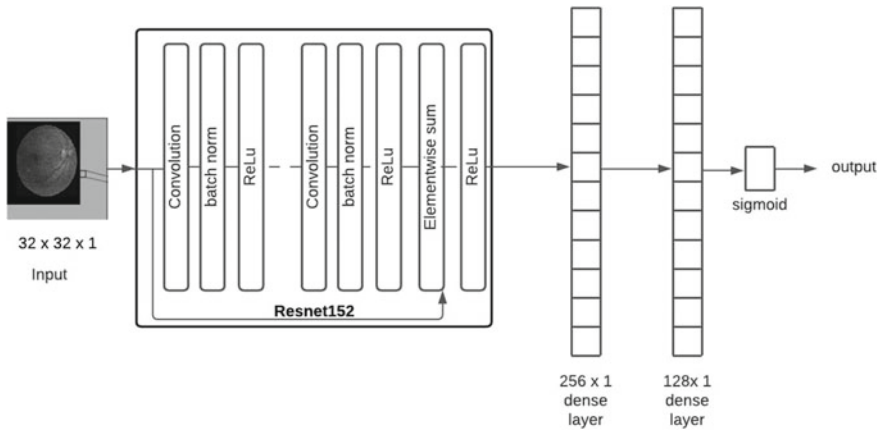


Fig. 4 ResNet architecture

Convolutional layer (dropout of 50%) to reduce over fitting. Dropout is a regularization technique for neural network models, which simply ignores randomly chosen units (i.e. neurons) during the training phase of a certain set of neurons. Ignoring precisely means that, and the neurons are ignored during either a forward or backward pass. The final output layer has one output which is the prediction of the image class. We have two classes: zero represents normal eye image, and one represents renal fundus images with Glaucoma disease.

Machine Learning Models. Machine learning models-based data modeling includes two major steps such as Feature extraction and Classification.

Feature Extraction: Feature extraction aims to extract the most discriminative features from an input image relevant for uniquely classifying the image. We use a kernel to move from left to right and from top to bottom i.e. performing a convolution operation. This operation is typically a multiplication of the matrix-shaped kernel with different parts of the input image which are also matrix-shaped resulting in a smaller number of feature representations of the input image.

Classification: Classification aims to predict the most likely label or class of a given input. For the classification step, we used some ML classifiers such as: SVM, KNN, NB, MLP, DT and RF classifiers using the Sklearn machine learning library. The models were trained to classify the images into two labels 0 for normal images and 1 for images that show Glaucoma.

3 Experiment Result and Comparison

3.1 Dataset Description

ODIR (Ocular Disease Intelligent Recognition) dataset is a collection of 6392 fundus images from approximately 5000 patients. It is collected by Shang gong Medical Technology Co., Ltd from Various hospitals and medical centers in China. The dataset includes eight labels for each image including Normal, Diabetes, Glaucoma, Cataract, Age-related Macular Degeneration, Hypertension, Pathological Myopia, Other diseases/abnormalities. The labels are represented by the first letter in the set. In this study, we selected only two classes from this dataset such as N and G labels, which respectively represent Normal and Glaucoma. The number of images of these classes is 600 for Normal and 397 for Glaucoma. We have split our dataset into 70% for training and 30% for testing to evaluate our proposed model.

3.2 Metrics Description

In this study we report classification metrics that include Precision (1), the fraction of the predicted results that is relevant, Recall (2), the fraction of the relevant results that is predicted, F1-score (3), the harmonic mean between precision and recall and, Accuracy (4), the ratio of correct predictions to the total number of all possible predictions [1]. All the above can be computed using elements of the confusion matrix that include [7]: True Positive (TP) (5) : Cases predicted to be positive (relevant) that are indeed relevant, False Positives (FP) (6): Cases falsely predicted to be positive (relevant), False Negatives (FN): Cases falsely predicted not to be relevant and True Negatives (TN) Cases correctly predicted to be negative.

$$Precision(P) = TP / (TP + FP) \quad (1)$$

$$Recall(R) = TP / (TP + FN) \quad (2)$$

$$F1 - score = 2 * (Precision * Recall) / (Precision + Recall) \quad (3)$$

$$Accuracy(Acc) = (TP + TN) / (TP + FP + TN + FN) \quad (4)$$

As additional metrics, we use in our evaluation the Area Under the Curve (AUC), which is essentially the probability that a classifier will have a random relevant case higher than a random irrelevant case. It is often visualized as the area under the receiver operating curve which is a line plot of the True Positive rate (TPR) vs the False positive rate (FPR).

$$TPR = TP / (TP + FN) \quad (5)$$

$$FPR = FP / (FP + TN) \quad (6)$$

In addition, all experiments have been evaluated on an AMD Intel core i7 processor system and are implemented using Python language.

3.3 Deep Learning Model Results

Table 2 presents the performance evaluation results for the CNN model using Resnet152 architecture in the binary classification task (Normal or Glaucoma images). As illustrated in this table, we obtained 84%, 88.2%, 81.4%, 96.3% and 79% for respectively Accuracy, F1-Score, Precision, Recall and AUC score metrics.

Figure 5 presents the confusion matrix of the CNN model from which we obtain values of TP, FP, TN, and FN to compute Precision (P), Recall (R), Macro F1, and Accuracy (Acc) reported in Table 2. $TP = 184$, implying that 184 instances were correctly classified as normal or non-Glaucoma affected eyes, $FP = 42$, implying 42 instances were incorrectly classified as Normal when they were glaucoma affected eyes, $FN = 7$ implying 7 instances were incorrectly classified as Glaucoma affected eyes when they were normal and $TN = 67$, implying that 67 instances were correctly classified to be Glaucoma affected eyes. Using the formulas presented in the metrics discussion section, we compute P, R, F and Acc presented in Table 2. The model generally performs well in detecting glaucoma in images of eyes, i.e. It effectively distinguishes normal eyes from those affected with Glaucoma.

Table 2 Results of Eye Decease detection used CNN model (Resnet152)

| Accuracy | f1-score | Precision | Recall | AUC |
|----------|----------|-----------|--------|-----|
| 84% | 88.2% | 81.4% | 96.3% | 79% |

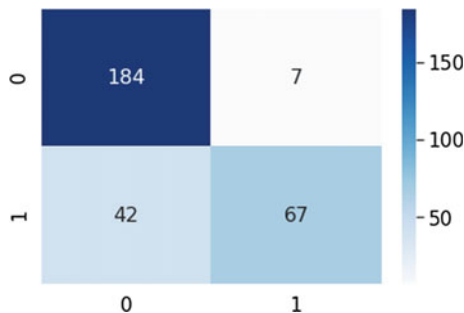


Fig. 5 CNN model confusion matrix

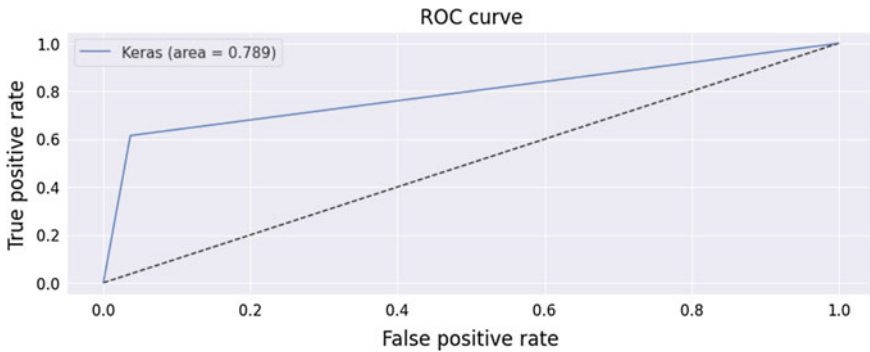


Fig. 6 The performance of CNN model using AUC metric

Table 3 Effect of varying the epoch numbers

| Epoch number | Accuracy (%) | AUC (%) |
|--------------|--------------|-----------|
| 50 | 70 | 63 |
| 100 | 68 | 66 |
| 200 | 65 | 57 |
| 400 | 77 | 71 |
| 600 | 84 | 79 |
| 700 | 78.33 | 73 |

Furthermore, we used the ROC curve to evaluate the model as shown in Fig. 6 The ROC curve is a common methodology to evaluate the model sensitivity. It examines the relationship between true and false-positive rates. In this case, the model showed great results. The recorded AUC is 78.9%.

In order to validate the choice of epoch number, Table 3 presents the results of tuning the epoch number for the CNN model. We observe the model achieve the best accuracy as well as the AUC performances with 600 epochs.

3.4 Machine Learning Model Results

We evaluated six different classification models on the task of detection of diseased eyes, particularly Glaucoma disease. For purposes of comparing the performance of all the models, we report Precision, Recall, F1-score and Accuracy, computed in a similar way to the description in the section above.

Results reported in Table 4 reveal that SVM, KNN, MLP and RF are quite competitive achieving macro F1 scores of 72% and above as well as AUC and Accuracy of 75% and above. We attribute their good performance to their superiority in classifying high dimensional data. Despite being outperformed by the prior three mentioned

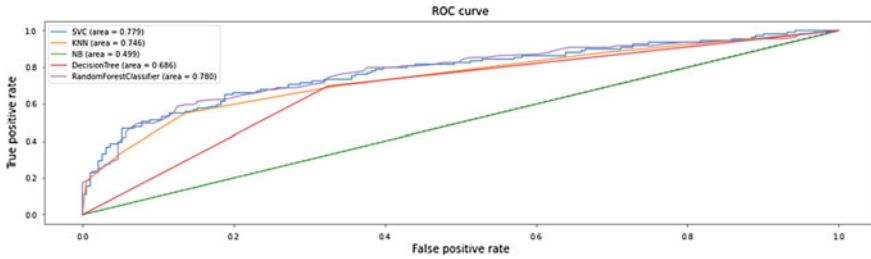


Fig. 7 ML Models ROC curve

Table 4 Results of Eye Decease detection used ML models

| Name | Macro P (%) | Macro R (%) | Macro F1 (%) | AUC | Accuracy (%) |
|------|-------------|-------------|--------------|-----|--------------|
| SVM | 75 | 72 | 73 | 78% | 76 |
| KNN | 73 | 71 | 72 | 75% | 75 |
| NB | 48 | 50 | 28 | 50% | 37 |
| MLP | 77 | 71 | 73 | * | 77 |
| DT | 67 | 69 | 67 | 69% | 68 |
| RF | 75 | 73 | 74 | 78% | 77 |

models, DTs performed moderately well with scores of 67%, 69% and 68% for Macro F1, AUC and Accuracy respectively. NB was on the other hand inefficient in this task achieving very low results in the monitored classification performance metrics. The evaluation results for these classifiers using AUC are 78%, 75%, 50%, 69% and 78% for SVM, KNN, NB, DT and RF respectively. Furthermore, for these classifiers, we obtain accuracy scores of 76%, 75%, 37%, 77%, 68% and 77% for SVM, KNN, NB, MLP, DT and RF respectively. In this Table 4, we present performance evaluation results obtained by ML models which include SVM, KNN, NB, MLP, DT and RF. As mentioned in this table, the result given by RF and SVM classifiers show the best result. In terms of AUC metric, RF and SVM classifiers outperform the rest of the other classifiers. RF classifier achieves the best results of about 78% and 77% for AUC and accuracy metrics respectively. In Table 4, in term of accuracy metric, the result given by RF and MLP algorithms are 77% show the best result. Figure 7 reveals the ROC AUC for all the models evaluated. SVC and RF were evidently the better performing models with the largest AUC of 77.9% and 78.0% respectively. KNN achieved the third best AUC. DT and NB had the least AUC of 68.6% and 49.9% respectively.

Comparing the six models based on accuracy and AUC as shown in Fig. 8, MLP and RF classifiers scored 77% accuracy while KNN scored 75%. But as mentioned in Table 4, considering AUC results RF and SVM scored the highest at 78%. That obviously indicates that SVM and RF are the best classifiers in classifying fundus images affected by Glaucoma.

Figure 8 illustrates the evaluation results of our proposed ML algorithms. Figures 9 and 10 show the confusion matrix for the two best performing algorithms, RF and MLP. A high number of instances correctly predicted to have Glaucoma (label 1) i.e. TP - 65 as well as as correctly predicted to be normal (label 0) i.e. TN - 165. The number of instances incorrectly predicted to have Glaucoma was rather low i.e. label 1 FP - 26 as well as those that the model failed to detect as Glaucoma affected eyes i.e. FN - 44. These values were used in computing the scores for Precision and

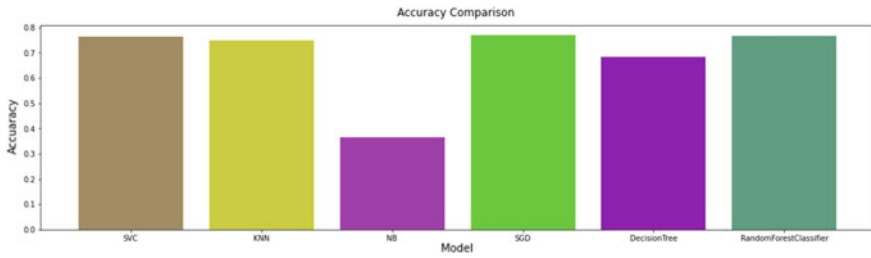


Fig. 8 ML Models performance Comparison using Accuracy metric

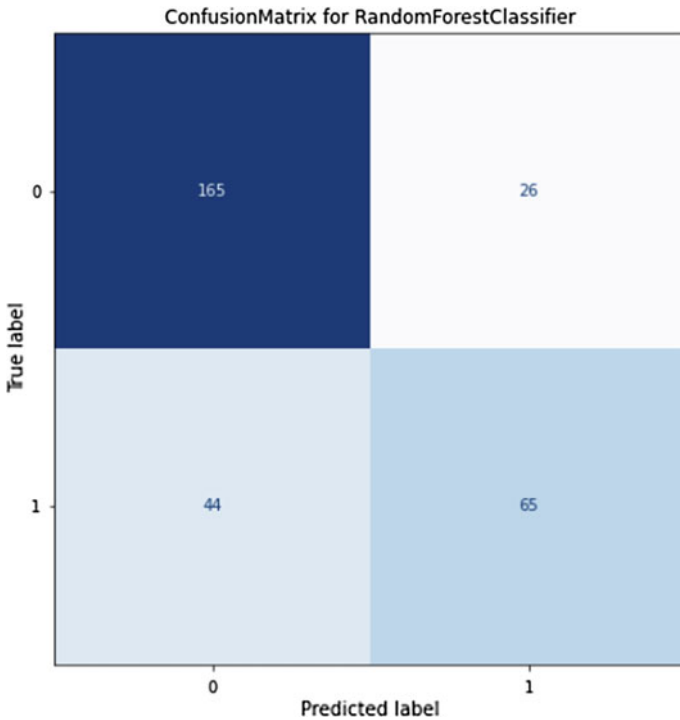


Fig. 9 RF Models Confusion Matrix

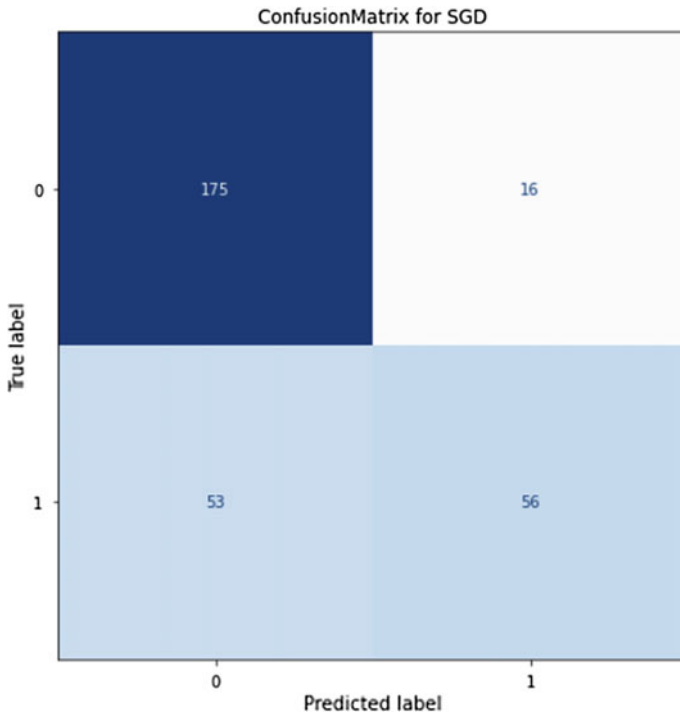


Fig. 10 MLP Models Confusion Matrix

Recall in Table 4. e.g. Precision for label 1 = $65 / (65 + 26) = 0.71$ and Recall = $65 / (65 + 44) = 0.60$.

Similarly Fig. 10 indicating MLP results reveals label 1 TP - 56, FP - 16, TN - 175 and FN - 53 i.e. Precision = $56 / (56 + 16) = 0.78$, Recall = $56 / (56 + 53) = 0.51$. MLP model uses hyperplane to classify the data as explained before. The confusion matrices in Fig. 10 shows that it predicted most of normal eye images correctly.

3.5 Deep Learning Versus Machine Learning

In Table 5, we compared the results obtained by the DL model with those obtained using the ML model algorithms. As it is revealed in this table, compared to ML algorithm, the DL outperforms ML in both considered metrics i.e. AUC and Accuracy. The results given by DL are 79%, 84% for AUC and accuracy metrics respectively, which is an improvement over 78% and 77% achieved by the RF for the same respective AUC and accuracy metrics.

Table 5 DL versus ML results comparison

| Type | AUC (%) | Accuracy (%) |
|----------------------|---------|--------------|
| DL model : Resnet152 | 79 | 84 |
| MLmodel :RF | 78 | 77 |

Table 6 Comparing with existing study summary

| Proposal | Model | Dataset | Performance (%) |
|-------------------|----------------|--|-----------------|
| Islam et al. [9] | CNN | ODIR:8000 images 90 % trainset & 10% testset | 85 |
| Our proposal work | CNN: Resnet152 | ODIR: 70%Trainset &30% Testset | 84 |

3.6 Comparison with Existing Study

In order to validate the performance of our CNN models, we compare its performance with [9] evaluated on the same dataset ODIR. The model in [9] slightly outperforms our model, however, the 1% improvement they have over our model is not statistically significant at $p = 0.05$ confidence interval, and therefore we conclude that our CNN model is a good benchmark model for automatic eye disease detection.

As shown in Table 5, the existing study divided the dataset images into 90% in the training set and 10% in the testing set and this leads to an increase and a rise of the accuracy performance (85%). But, in our study, we divided our dataset images by the logical division into 70% in the training set and 30% testing set. In addition, we used the ResNet152 model that is not used previously in the presented existing study review. We have improved this model by adding some layers such as the use of the dropout regularization after each Convolutional layer to reduce overfitting and the nonlinear transformation that is carried out by Rectified Linear Units (ReLU) as the activation function. The proposed deep learning architecture shows its competitiveness with other used models (Table 6).

4 Conclusion

This work has attempted to develop ML and DL models for an image classification task that classifies an eye as either Glaucoma infected or not. We built and evaluated a CNN model oriented with ResNet152 architecture and, indeed, it was our best performing model achieving accuracy and F1 of 84 and 88% respectively. This model was found competitive when compared to prior work. Six other models SVM, RF, KNN, and MLP evaluated on the same task using the same dataset, performed quite

well, with the best model (RF) achieving an average Macro-F1 of 78% and average accuracy of 77%. In conclusion, we recommend our DL CNN model as a good benchmark for eye disease (Glaucoma) classification tasks. For future work, we will propose to detect and identify the blood vessels of the color retina texture images and detect the infections according to the texture analysis.

References

1. Casson RJ, Chidlow G, Wood JP, Crowston JG, Goldberg I (2012) Definition of glaucoma: clinical and experimental concepts. *Clin Exp Ophthalmol* 40(4):341–349. (Roy Soc Lond. A247:529–551, April 1955)
2. Dandona L, Dandona R (2006) What is the global burden of visual impairment? *BMC Med* 4(1):6
3. Choi JY, Yoo TK, Seo JG, Kwak J, Um TT, Rim TH (2017) Multi-categorical deep learning neural network to classify retinal images: a pilot study employing a small database. *PLoS one* 12(11):e0187336
4. Abbas Q (2017) Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. *Int J Adv Comput Sci Appl* 8(6):41–5
5. Soltani A, Battikh T, Jabri I, Lakhoua N (2018) A new expert system based on fuzzy logic and image processing algorithms for early glaucoma diagnosis. *Biomed Signal Process Control* 40:366–377
6. Issac A, Sarathi MP, Dutta MK (2015) An adaptive threshold based image processing technique for improved glaucoma detection and classification. *Comput Methods Programs Biomed* 122(2):229–244
7. Krishnan MMR, Faust O (2013) Automated glaucoma detection using hybrid feature extraction in retinal fundus images. *J Mech Med Biol* 13(01):1350011
8. Shantaiya S, Gorasia S, Anwar R (2016) Early detection of glaucoma using retinal fundus images. *Imp J Interdiscip Res (IJIR)* 2(6):1525–1528
9. Islam MT, Imran SA, Arefeen A, Hasan M, Shahnaz C (2019) Source and camera independent ophthalmic disease recognition from fundus image using neural network. In: 2019 IEEE international conference on signal processing, information, communication & systems (SPIC-SCON). IEEE, pp 59–63. (Nov 2019)
10. Gogineni S, Pimpalshende A, Goddummarri S (2021) Eye disease detection using YOLO and ensemble GoogleNet. In: *Evolutionary computing and mobile sustainable networks*. Springer, Singapore, pp 465–482 2021
11. Tong Y, Lu W, Yu Y, Shen Y (2020) Application of machine learning in ophthalmic imaging modalities. *Eye Vis* 7(1):1–15
12. Gallagher B, Rever M, Loveland D, Mundhenk TN, Beauchamp B, Robertson E, Han TYJ (2020) Predicting compressive strength of consolidated molecular solids using computer vision and deep learning. *Mater Des* 190:108541
13. Sungeetha A, Sharma R (2021) Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network. *J Trends Comput Sci Smart Technol (TCSST)* 3(02):81–94
14. Vivekanadam Balasubramaniam (2021) Artificial intelligence algorithm with SVM classification using dermoscopic images for melanoma diagnosis. *J Artif Intell Capsul Netw* 3(1):34–42
15. Kaggle Ocular Disease Recognition (2020). <https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k>

An Automated System to Preprocess and Classify Medical Digital X-Rays



Sumera, K. Vaidehi, and R. Manivannan

Abstract Considering, the reality (fact) that innumerable medical (digital) images are collected at hospitals and numerous medical tests centres on a daily basis, a clear manifestation is that these images must be composed, stacked and accessed accurately for future references (CBIR). Sole motivation (reasoning) for proposition of this attempted work is to give a provision for categorizing or classifying medical X-rays automatically at macro-level (global level) to aid lab technicians (analysts) in their day job. GLCM is employed to draw out features or characteristics of (images) X-rays and resultant is utilized in building classification model by taking an advantage of SVM. Six different classes (groups) of X-ray images are taken, namely chest, foot, skull, neck, palm and spine. The proposed attempt to put forward a medical X-ray image classification process involves pre-processing of X-rays with an aim of making them fit for further processing. Digital X-rays in this current research are subjected to Pre-processing using a filtering operation (median filter), histogram filtering (or equalization) and CLAHE. The upshots of each are recorded. Subsequently, segmentation (connected component labelling), feature extraction (GLCM), classification (SVM). Lately post implementation, the outcomes vividly depict around 91% accuracy is acquired utilizing median filter accompanying GLCM and SVM whereas pre-processing images using, histogram equalization (HE) yielded 89% accuracy. Third blend (combination) of CLAHE, GLCM and SVM outshined with 96% accuracy. Consequently, CLAHE in estimation (comparison) to other Pre-processing methodologies outperformed classification results.

Sumera (✉) · K. Vaidehi

Department of Artificial Intelligence, Data Science and Computer Engineering, Stanley College of Engineering and Technology for Women, Abids, Hyderabad, India
e-mail: mohdd.ubaidd@gmail.com

K. Vaidehi

e-mail: kvaidehi@stanley.edu.in

R. Manivannan

Department of Computer Science Engineering, Stanley College of Engineering and Technology for Women, Abids, Hyderabad, India
e-mail: drmanivannan@stanley.edu.in

Keywords Contrast limiting AHE (adaptive histogram equalization) · Co-occurrence matrices · Medical image intensification · Classification · Segmentation. SVM (support vector machine) · GLCM (gray-level co-occurrence matrix)

1 Introduction

One can offer a tremendous contribution to facilitate doctors and radiologists in (their) work of clinical diagnosis by designing an approach (structure) that could automate their tasks instead of alternately letting them perform manual observations. It gets quite tedious and also incurs surplus time for medical practitioners to manually perform image interpretation. Hence, new vogues for processing of images automatically through computers and medical systems and classifying those images are welcomed now-a-days. Benefitting this needy trend of therapeutic systems, this chore concentrates on developing a methodology (or system) to automate categorization (classification) of X-ray images into six groups. These six groups of X-ray images considered in this study are chest, foot, spine, neck, skull, palm. Images considered in this study are taken from IRMA image CLEF database. This thesis comprises of-four fragments. Fragment 2 provides some information on existing trends, information on work proposed is summarized in Fragment 3, Fragment 4 shows performance assessment with respect to current developed system. Finally, experimental outcomes are handed over to Fragment 5. Fragment 6 gives conclusion.

2 Existing Trends

Enormous work has been done in past and till date by experts across globe in developing medical systems to aid and automate medical diagnosis wherein [1] introduces an effort to interpret digital images, improve diagnosis quality using techniques like median filter, histogram Equalization, for quality enhancement, PCA, K-nearest-neighbour techniques, were applied for selecting features and classification [2]. Introduces a case-study wherein analysis of malicious software is done based on machine learning approaches. Reference [3] introduces a work wherein an effort has been made to develop a CBIR system to retrieve mammographic images of breast tissue and classify them as dense, fatty, glandular using statistical features and SVM. Reference [4] introduces research on a similar area (with splitting/merging scheme) wherein contrast equalizing histogram, FD, ZM and classifiers like KNN, MLP are used. Reference [5] presents work on TRUS images which uses techniques like M3 filter, DBSCAN clustering, SVM to demonstrate a close curative system. Reference [6] presents a work of similar kind wherein classification of images (breast mammograms) based on breast mass is proposed in which segmentation is performed using fuzzy C-means, GLCM features are taken out and adaboost, back propagation, neural networks, sparse representation classifiers are used. Reference [7] presents work wherein pectoral breast muscle cropping is done where Haralicks' and Zernike

descriptors facilitate extraction of features. Classifiers like BPN with SVM operate which resulted accuracy of 95.83%. Reference [8] presents a machine learning based approach to detect/identify images with glaucoma. A series of preprocessing and morphological operations are applied followed by segmenting optic-cup part and extracting rim-disc/cup-disk ratios as features and finally classifying using SVM. Reference [9] In this abstract, collection of (Lung) images are Pre-processed and manipulated with respect to pixel values to either remove distortion, filter the image, enhance contrast for making images fit for better visualization and interpretation [10]. Automated categorizing (of microorganisms) is suggested using SVM classifier. Before performing classification, system feeds image (to module) for Pre-processing then a phase to extract features. Reference [11] introduces thesis presenting a methodology for, classification of lung images in which lung data (images) are collected via lung database with a motto to group them as cancerous/non-cancerous. Prior to the final conclusion, images undergo a set of manipulations constituting Pre-processing (median filter), segmenting (fuzzy c means), GLCM and lastly SVM. Reference [12] shows a machine learning approach to recognize symbols (mathematical symbols) which are handwritten. Reference [13] shows an implementation of a system to detect expressions from a real time dataset of faces by making use of SVM for classification of facial emotions into six categories [14]. Proposes a detection system to ascertain type of breast tissue using GLCM features and FLDA classifier that yielded an accuracy of 72.93% using features based on texture and 82.48% using cascade features [15]. This paper gives an overview on different aspects/studies in educational data mining [16]. Introduces a survey/study on Detection of outliers based on machine learning approaches by considering IOT data for analysis. Reference [17] presents a case study on spam detection based on machine-learning approaches.

Reference [18] Presents classification implementation wherein an approach to classify X-rays by applying M3-filter (better image quality), CCL for segmenting, attributes related to texture, shape are withdrawn by operating with GLCM then classification via SVM [19]. Presents a work on image fusion using guided filters [20]. Introduces a review/study on multi-focus-mage fusion using diversified mechanisms.

3 Proposed Trend

Medical (diagnostic) images accumulated from IRMA (medical image dataset) are fed to Pre-processing module wherein CLAHE (Contrast Limiting AHE (adaptive histogram equalization)) algorithm enhances image calibre (Besides this, other filtering methods like median filter, histogram Equalization; are used to examine, compare the overall system performance) accompanied by segmentation utilizing Connected component Labelling. Texture features for interested region (ROI) are pulled out using GLCM statistics. SVM model is built to which images are fed for classification. System architecture is pictorially introduced in Fig. 1.

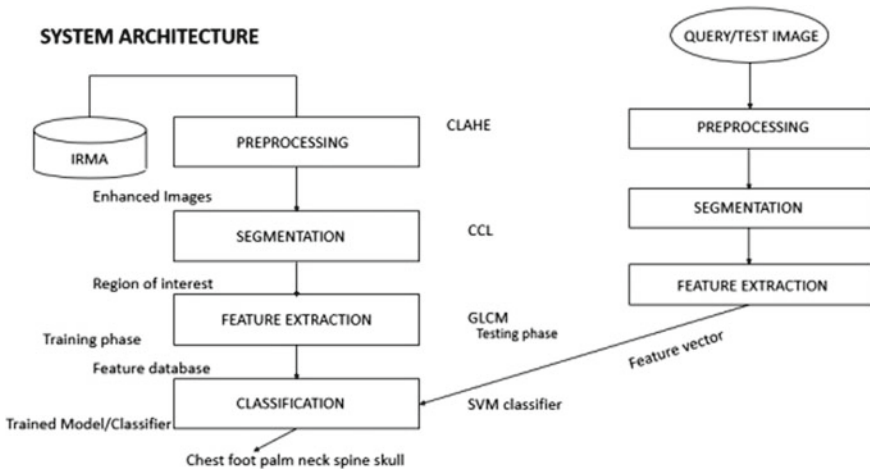


Fig. 1 Architecture of proposed trend

3.1 A Module to Improve Quality of Images-Pre-processing

For quality intensification (also termed as image enhancement), intensities (or intensity values, pixel values) of digital X-rays are modified which gives finer visualizing, interpretation and displaying experience. It lowers noise within image. One among the pitfalls of median filter is to pull out any outliers (along) with fine or minute details since its' hard for it to differentiate (distinguish) between the two. Median rate (value) will very minutely be affected by anything diminutive in size. Histogram Equalization emphasizes (stresses) on finding image's global contrast that gives too dark, too light images. Since it is not ideal for image enhancement, a variation called CLAHE is implemented here. CLAHE concentrates on finding the local contrast of image areas thereby uniformly distributing image intensities. It also clips off or limits contrast to escape noise amplification if any. Sample images demonstrating Pre-processing strategies (techniques) are shown below in Figs. 2, 3, 4, and 5.

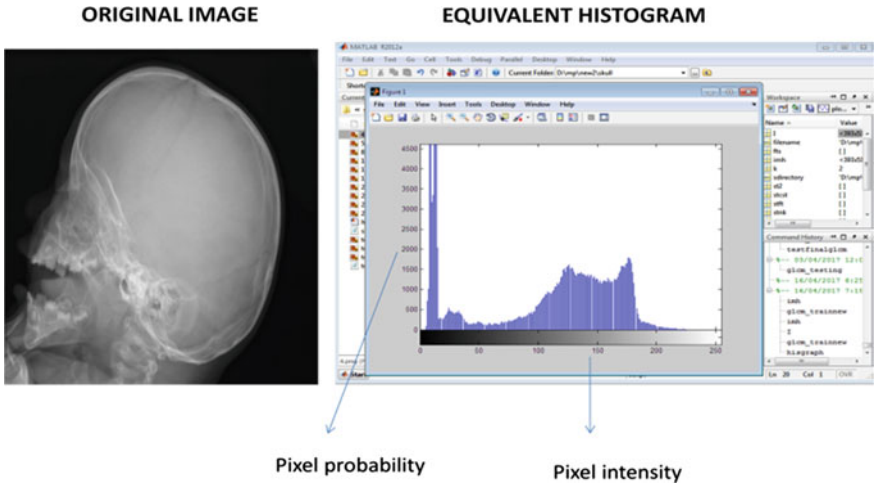


Fig. 2 Original (skull X-ray) image

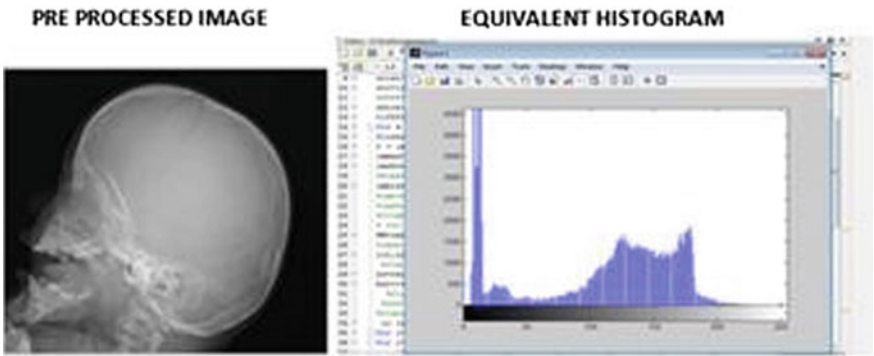


Fig. 3 Image (X-ray) after Pre-processing (median filter)

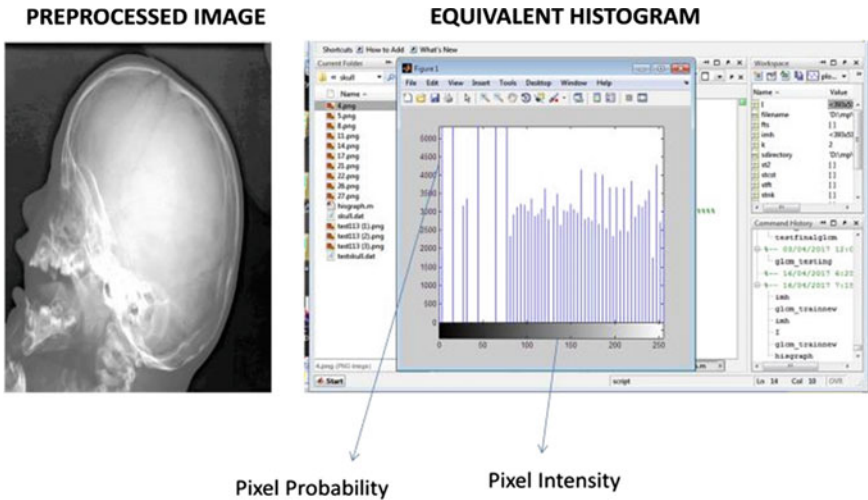


Fig. 4 Image (X-ray) after pre-processing (histogram equalization)

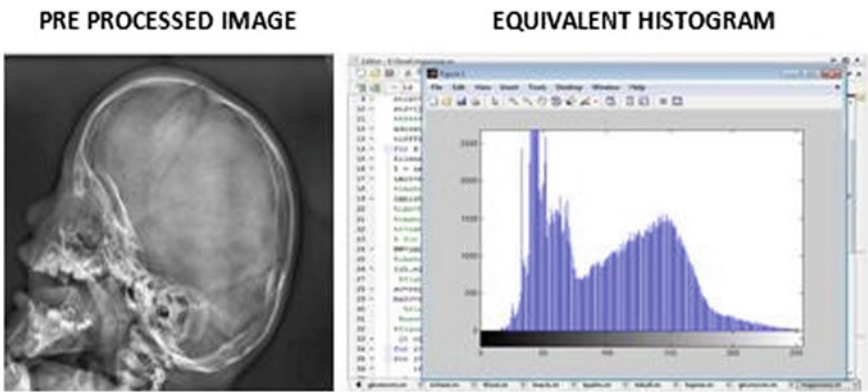


Fig. 5 Image (X-ray) after pre-processing (CLAHE)

3.2 A Module to Extract the Region of Interest-Segmentation

ROI (segmented part) whose features need to be extracted is found using CCL (Connected Component Labelling). This segmentation mode scans an X-ray, group pixels of it considering pixel connectivity. Pixels inside same group share indistinguishable characteristics. Each group of pixels are assigned a label, largest labelled component is the desired region-of-interest (ROI). Since CCL provides a very efficient, user-friendly way to segment the images thereby providing further ease to calculate and display portions of images based on 4-connectivity or 8-connectivity, this has been used in this work.

3.3 A Module to Calculate Statistics-Feature Extraction

When load (input) to algorithm is bulky to be processed (when image sizes are large enough), this input (images) can be represented as compact (feature vector). Here, a statistical methodology called GLCM (Gray Level co-occurrence matrix) is utilized for withdrawing of textural feature-based particulars of the images. GLCM matrices takes (original image's) pixel intensity co-occurrences into consideration and forms a matrix. Resultant matrix is operated to discover texture statistics (features). Proposed work presented here dealt with extraction of 22 features altogether, couple of them are: correlation, dissimilarity, contrast, variance, entropy, energy homogeneity, etc.

3.4 A Module that does the Purpose-Classification

Feature sets of X-rays (used here) computed in previous segment are forwarded to establish a classifier (SVM here). Since all six categories/classes of images are finely pre-processed and features are withdrawn using ROI, SVM here analyzes these feature sets (of images) and classifies them into six categories where images belonging to same class share similar feature values. SVM principal is depicted pictorially. See Fig. 6.

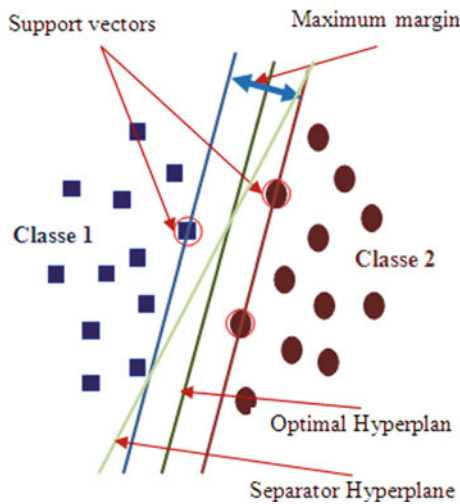


Fig. 6 SVM principal

4 Performance-Evaluation

Thankfully, Machine Learning (ML) approaches/algorithms are furnished with innumerable alternatives to calculate correctness (accuracy) of any model being developed. This study aims at calculating accuracy (how well did the classifier performed) of classifier by utilizing a confusion matrix (which is a-combination of True positives (TP), True negatives (TN), False positives (FP),False negatives (FN) of predicted versus actual task results. Here the classifiers' quality is ascertained using specificity, accuracy, sensitivity and specificity measures.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad \text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \quad \text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

5 Investigational Outcomes

Eventually, CLAHE Pre-processing worked best with GLCM, SVM offering 96% accuracy unlike median filtering, histogram filtering wherein classification result or accuracy was found as 91% ,89% respectively. Overall performance results of classifying X-ray images preprocessed using **CLAHE** are tabulated in Table 1 wherein; for the class Skull, precision, sensitivity, specificity values are 100%, 80%, 100% respectively with an accuracy of about 96.55% in contrast to the class Foot wherein the accuracy is much lesser i.e.; about 89.65%.Overall performance results of classifying X-ray images preprocessed using **Histogram Equalization** are tabulated in Table 2 wherein; for the class Neck, precision, sensitivity, specificity values are 100%, 70%, 100% respectively with an accuracy of about 94.82% in contrast to the class Spine wherein the accuracy is much lesser i.e.; about 84.48%.Overall performance results of classifying X-ray images preprocessed using **Median Filter** are

Table 1 Performance results (measures) for GLCM, SVM- with CLAHE

| X-ray- image | Precision (%) | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---------------------|---------------|-----------------|-----------------|--------------|
| Skull | 100 | 80 | 100 | 96.55 |
| Foot | 64.28 | 90 | 89.5 | 89.65 |
| Palm | 88.88 | 80 | 97.1 | 94.82 |
| Chest | 100 | 87.5 | 100 | 98.27 |
| Neck | 100 | 90 | 100 | 98.27 |
| Spine | 100 | 70 | 100 | 94.82 |
| Overall performance | 92.91 | 82.91 | 97.76 | 96 |

Table 2 Performance results (measures) for GLCM, SVM -with histogram equalization.

| X-ray image | Precision (%) | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---------------------|---------------|-----------------|-----------------|--------------|
| Skull | 77.77 | 70 | 95.83 | 91.37 |
| Foot | 80 | 80 | 95.83 | 93.10 |
| Palm | 66.66 | 60 | 93.75 | 87.93 |
| Chest | 54.54 | 75 | 90 | 87.93 |
| Neck | 100 | 70 | 100 | 94.82 |
| Spine | 66.66 | 20 | 97.91 | 84.48 |
| Overall performance | 74.27 | 62.5 | 95.9 | 89 |

Table 3 Performance results (measures) for GLCM, SVM with- median filter

| X-ray-image | Precision (%) | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---------------------|---------------|-----------------|-----------------|--------------|
| Skull | 70 | 70 | 93.75 | 89.65 |
| Foot | 100 | 70 | 100 | 94.82 |
| Palm | 66.66 | 80 | 91.66 | 89.65 |
| Chest | 100 | 62.5 | 100 | 94.82 |
| Neck | 100 | 80 | 100 | 96.55 |
| Spine | 66.66 | 20 | 97.91 | 84.48 |
| Overall performance | 83.88 | 63.75 | 97.22 | 91 |

tabulated in Table 3 wherein; for the class Neck, precision, sensitivity, specificity values are 100%, 80%, 100% respectively with an accuracy of about 96.55% in contrast to the class Spine wherein the accuracy is much lesser i.e.; about 84.48%. Figures 7, 8 and 9 presents output X-rays (images) post pre-processing, segmentation when those images were preprocessed using CLAHE, Histogram Equalization and Median Filter respectively. The performance measures discloses (shows) that CLAHE preprocessed (X-ray) along with GLCM (texture features) and SVM (classifier) gives a better/finer accuracy over other filters and serves the purpose. Hence, SVM-classification is quite acceptable for classifying X-rays (images). A bar graph (as) in Fig. 10 is opted to present a pictorial/vivid look on classification results.

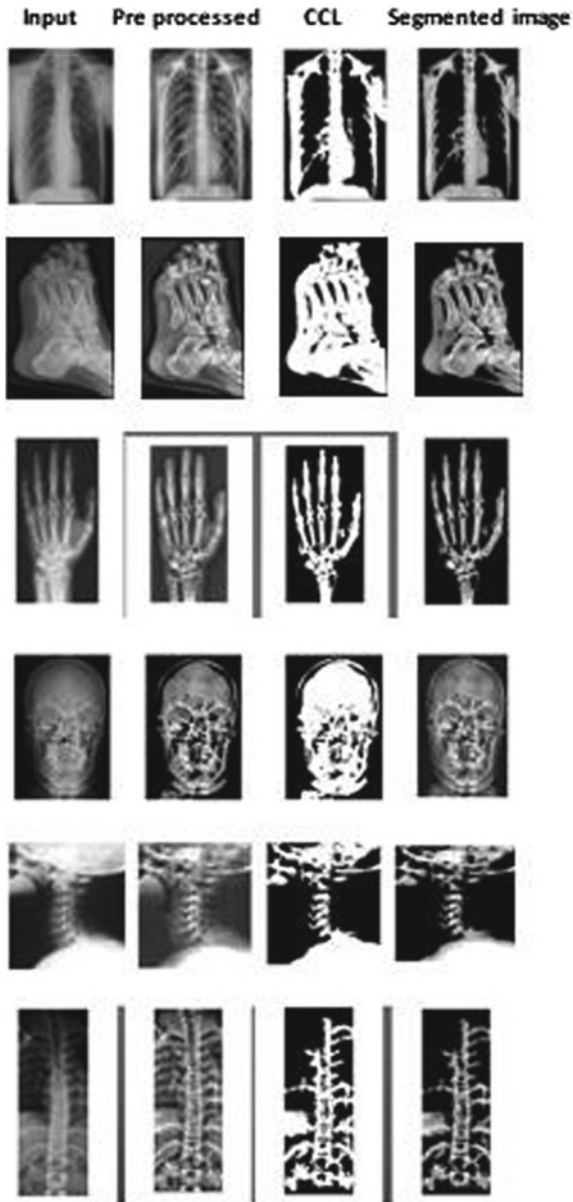


Fig. 7 Represents X-rays post CLAHE

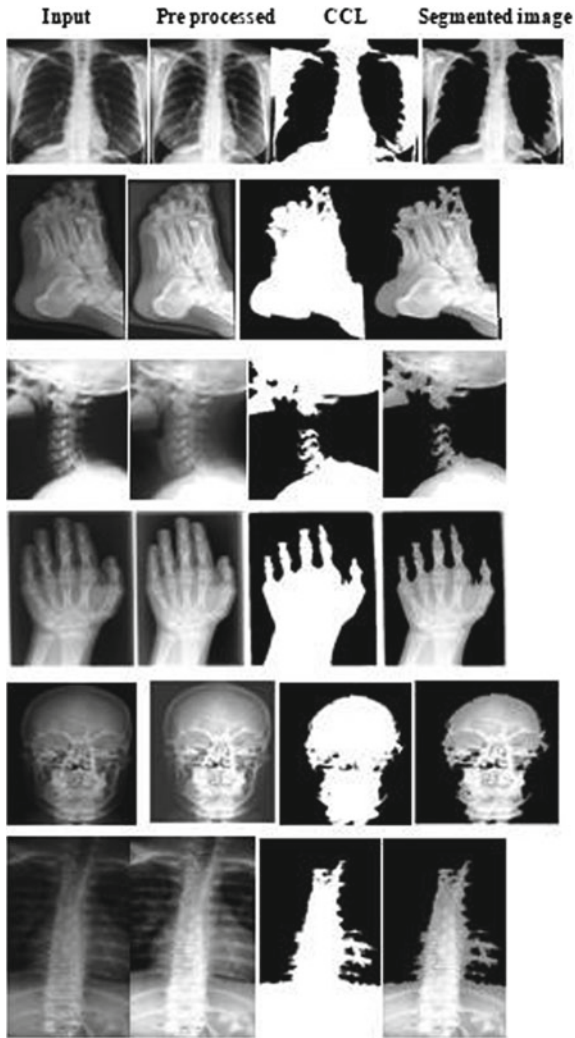


Fig. 8 Represents X-rays post histogram equalization

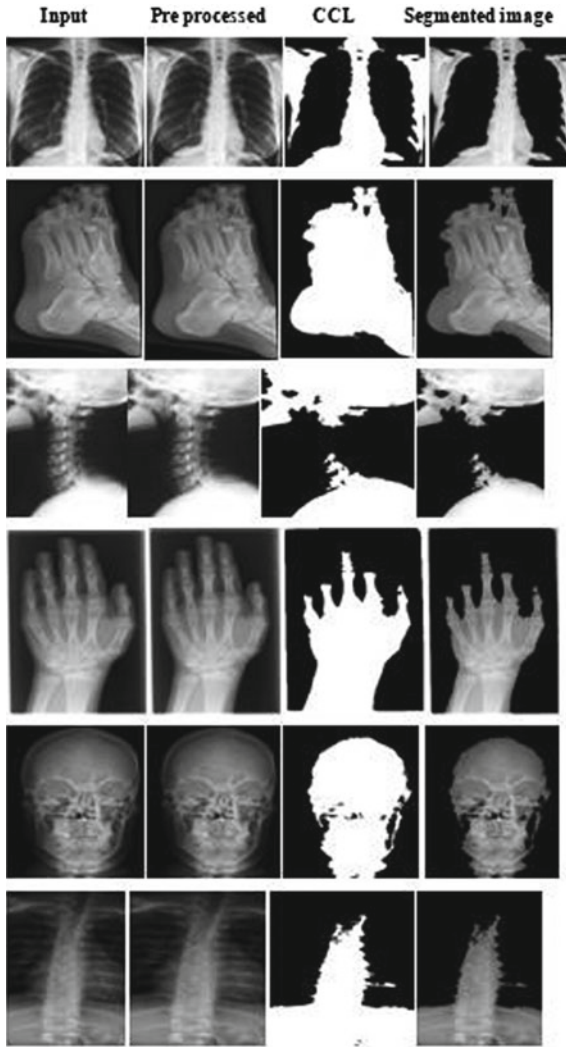


Fig. 9 Represents X-rays post median-filter

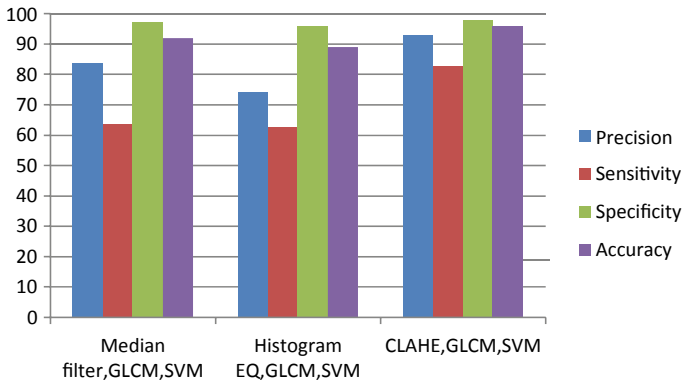


Fig. 10 Performance results (measures) for overall percentage of classified X-ray image

6 Conclusion

In Proposal described here, medical (X-rays) images of six categories namely chest, spine, palm, skull, foot, neck are acquired from IRMA medical database. These images (X-rays) are subjected to Pre-processing module wherein three different image manipulation techniques/strategies namely Median-filter, Histogram Equalization and CLAHE are implemented (or worked with) to compare their performance/results and find which Pre-processing technique works best with GLCM, SVM. Images when Pre-processed with CLAHE outshined among all three with accuracy of 96% altogether, sensitivity of 82.91%, specificity of 97.96% and precision of 92.91%. This gave a conclusion that images when Pre-processed with CLAHE gave better accuracy post classification. Future scope of this proposed work is the extension (of current work) for content based image retrieval systems (CBIR). Unlike the six groups/classes of X-rays images specified above, other parts like knee, elbow, limbs, ankle etc. can also be taken into consideration for future upgrading of medical systems. A Heterogeneous (image) types like C.T, M.R.I, U.S.G etc. can also be worked with in future.

References

1. Fang Yang, MuratHamit 'Feature-Extraction-and-classification-of-Esophageal-X-ray-images-of-Xinjiang-Kazak Nationality' Volume 2017, Article:ID 4620732.
2. Dr. M. Upender kumar, Dr. D Shravani "Novel Design of Machine Learning for Malicious Software Analysis – Malicious URL Case Study" Vol 6 Issue 4 October 2018 – December 2018 pp 292–298 International Journal of Interdisciplinary Research and Innovations (IJIRI).
3. Vaidehi, K., and T. S. Subashini. "Content- Based –Mammogram-Retrieval-based-on-Breast-Tissue-Characterization-using Statistical-Features." Research Journal of Applied Sciences, Engineering and Technology 8.7 (2014): 871–878.

4. Nooshin,jafari and Hosssein "Medical-X –ray-image-hierarchical-classification-using-a-merging-and-splitting-scheme-in feature- space" March 2013 Volume 2, Issue11.
5. R. Manavalan K. Thangaval "Evaluation-of-textural-feature-extraction-from-GLCM-for-prostate-cancer-TRUS-medical- images" 'International Journal of- Computer Applications' (No.0975 _ 8887), Volume :36– No:12, December2011
6. Vaidehi, K., and T. S. Subashini. "Automatic-characterization-of-benign-and-malignant-masses-in-mammography." *Procedia Computer Science* 46 (2015): 1762–1769.
7. V. Kaliyaperumal, S.Selvarajan "Automated-characterization- of mammographic-density-for-early-detection-of-breast-cancer risk"International-Journal-of-Simulation :Systems.Science & Technology. Feb 2014.
8. Sumera, K. Vaidehi, K., and J. Shahistha. "Automated glaucoma detection using machine learning approaches" *Turkish online journal of qualitative enquiry*"-2021.
9. S.Perumal and T.Velmurugan "Preprocessing-by-Contrast-Enhancement-Techniques-for-Medical-Images." *International- Journal -of Pure and Applied- Mathematics*" Volume 118 No. 18 2018, 3681–3688.
10. Vanitha.L. and Venmathi.A.R "Classification of- Medical-Images-Using-Support-Vector-Machine" 2011 International Conference-on-Information -and-Network-Technology.
11. Usha Kumari "Lung Cancer Image-Feature Extraction- and-Classification using 'GLCM' and 'SVM' classifier" September 2019.
12. Firdaus, Syeda Aliya, and K. Vaidehi. "Handwritten-mathematical-symbol-recognition-using-machine-learning techniques." *Advances-in-Decision-Sciences,-Image Processing, Security-and-Computer-Vision*. Springer, Cham, 2020. 658–671.
13. Sathya, R., R. Manivannan, and K. Vaidehi. "Vision-Based Personal Face Emotional Recognition Approach Using Machine Learning and Tree-Based Classifier." *Inventive Computation and Information Technologies*. Springer, Singapore, 2022. 561–573.
14. Vaidehi, K., and T. S. Subashini. "Automatic classification and retrieval of mammographic tissue density using texture features." 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO). IEEE, 2015.
15. Anjum, Nadia, and Srinivasu Badugu. "A study of different techniques in educational data mining." *Advances in Decision Sciences, Image Processing, Security and Computer Vision* (2020): 562–571.
16. Nenavath chander, Dr. M Upender, *Machine-learning-based-Outlier-Detection-Techniques-for-IOT: A comprehensive Survey* Volume 12 Issue 1 January 2021 pp 144–158 *International Journal of Advanced Research in Engineering and Technology (IJARET)* IAEME publication.
17. Amogh Deshmukh, Dr. M. Upender "Cyber-Security-Engineering-for-Malware-Analysis – Machine Learning for Spam detection case study" Vol 10 Issue 10 October 2019 pp 301–306 *Journal of Engineering Sciences JES* ISSN 0377-9254.
18. Sumathi Ganesan ,T.S. Subashini "Classification-of-medical- X-ray images-for-Automated-annotation" *Journal-of-Theoretical and-Applied-Information-Technology* 31, May 2014. Vol. 63, No.3.
19. Dulhare, Uma N., and Areej Mohammed Khaleed. "Taj-Shanvi Framework for Image Fusion Using Guided Filters." *Data Management, Analytics and Innovation*. Springer, Singapore, 2020. 419–427.
20. Dulhare, Uma, Areej Mohammed Khaled, and Mohd Hussam Ali. "A review on diversified mechanisms for multi focus image fusion." *Proceedings of International Conference on Communication and Information Processing (ICCIP)*. 2019.

PSTN Enterprise Collaboration Systems



P. S. JosephNg, R. C. Yeo, J. H. Wong, P. H. Tan, Y. H. Yong, and S. L. See

Abstract Both intranet and extranet can be interchanged in Enterprise Collaboration System (ECS). However, it still has some limitations and challenges are faced by the organization due to connection issues. System consistency, safety, and security issues are the main concerns in this discussion on how the system can ensure information protection in secured status. To solve these issues, it is suggested that ECS should use the other connection which is called PSTN connectivity that is used to connect collaboration systems with the network. Especially in this pandemic period, almost 90% of the employees are required to work from home; this is indirectly challenging the connection to the company ECS, considering the stability of the system. Email is one of the most popular communication tools used in the enterprise. The company will create a unique email address for their employee used in ECS. However, problems such as message overloading, spam, and so on may occur. Therefore, LightWeight Semantics (LWS) approach is suggested to plug into an email to further process email messages. This is indirectly helping the searching process to become more direct to obtain the expected search results via semantic search. Legacy in LWS is connected to avoid any data loss and to study the previous data storing style.

P. S. JosephNg (✉) · R. C. Yeo · J. H. Wong · P. H. Tan · Y. H. Yong · S. L. See
Faculty of Data Science and Information Technology, INTI International University, Nilai,
Malaysia

e-mail: joseph.ng@newinti.edu.my

R. C. Yeo

e-mail: I19017134@student.newinti.edu.my

J. H. Wong

e-mail: I20019131@student.newinti.edu.my

P. H. Tan

e-mail: I20019398@student.newinti.edu.my

Y. H. Yong

e-mail: I20019139@student.newinti.edu.my

S. L. See

e-mail: I20019214@student.newinti.edu.my

Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

Keywords Lightweight semantics · LWS · ECS · Intranet · Extranet · Collaboration

1 Introduction

The Enterprise 2.0, or Enterprise Collaboration system, involves the use of intranet and extranet to the enterprise networks. Intranet in the enterprise collaboration system allows the enterprise for communication between teams and individuals in the Local Area Network (LAN) or Wide Area Network (WAN). The intranet in ECS acts as a private network on the enterprise where most of the information are relevant to the enterprise, mostly exclusive and sensitive information. A006E extranet is an extended version of the intranet. By using the extranet, the enterprise provides an important improvement in communication with the customer or external stakeholders. One of the good examples of extranet in ECS is the online service such as social-based Customer Relationship Management (CRM) system [2]. It provides a communication platform for employees and clients, allowing them to provide and raise the level of innovativeness in customer services, sharing knowledge regardless of time or place.

In the business competition, an extranet is used as a market strategy for many enterprises. With the extranet, enterprises can receive the real-time updation of data and information for data analytics. Extranet in ECS also allows the enterprise to quickly communicate with clients, and to keep online contact with customers or external stakeholders. It brings a lot of benefits to the enterprise such as creating new opportunities for acquiring and servicing the existing and prospective customers [2]. Besides, the extranet is efficient in the sharing of knowledge with external stakeholders. The enterprise can acquire the benefits from external stakeholders' information such as improving the enterprise's competitive strategy. However, it is a double-edged sword as the external stakeholder can also get some of the enterprise's information for their business strategy (see Fig. 1).

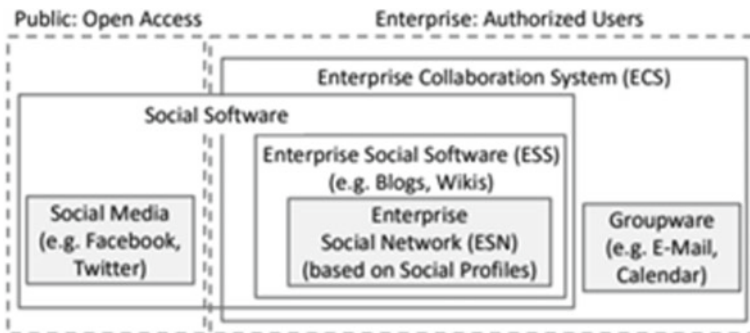


Fig. 1 Interrelationships between the social intranet and the enterprise collaboration system

1.1 Problem Statement, Question and Objective

Many organizations are using Intranet as their network solution to transfer data via the internal network without using the internet. An extranet is also used by organizations for communicating with the supplier and others who are not the employees of the organization. Enterprise Collaboration System (ECS) or Enterprise 2.0 is also known as intranet with social software [3]. Therefore, a conceptual solution is provided to enhance the intranet to modern intranet by considering the weaknesses of the current intranet.

Research Questions

- RQ1: What are the differences between Groupware, Social intranet, and ECS extranet use in business intelligence?
- RQ2: How does the Enterprise Collaboration System help an organization and what are the characteristics and functionalities of the system?
- RQ3: What are the solutions that can improve ECS from its disadvantages into a modern intranet for future use?

The questions stated above, are solved through the objectives of this work.

Research Objective

- R01: To identify the differences between the type of networks used in a business intelligence environment.
- R02: To identify the benefits of ECS that is brought to an organization, from the perspective of company branding and financial savings.
- R03: To enhance the current ECS limitation by suggesting a conceptual idea, to improve its disadvantages into an increased efficiency.

The objectives above, clearly states about what is expected to be performed.

Research Hypothesis

- RH1: Security and protection of the ECS system can be further enhanced up to 90% to avoid data loss.
- RH2: To reduce the time on searching information from email and to obtain the expected result.
- RH3: New conceptual solutions are suggested that can enhance costing, branding, and efficiency.

2 Literature Review

Groupware refers to the traditional computer software that runs in the enterprise's intranet to help employees make correspondence and management on the workgroup exercise [1]. It provides communication tools, conferencing tools, and collaborative management tools which allow the employees in the enterprise to exchange

documents, send emails, share database access, online meeting, view information of other employees, calendaring and more. It has become a far more efficient and effective technique of communication between groups and individuals than the previous written documentations or electronic methods [1]. It brings the benefits of fast and easy access to information to improve enterprise productivity. Besides, groupware also allows the employees of the enterprise to communicate easily such as, the employee can join the voice conferencing or video conferencing, over the intranet on ECS [1]. Thus, it can improve the enterprise collaboration, since it reduces the travel time and cost for gathering meetings, which has become a requisite especially during the Covid pandemic.

Social Intranet is an intranet that uses multiple social media tools for employees to use as a collaboration for sharing knowledge with others [3]. The features of social media can be blogs, wikis, discussion forums, etc. Social intranet helps the enterprise to keep in contact with their customers on different platforms of social media and also helps them in collecting information for data analytics to improve their business. The main difference between the Enterprise Collaboration system (ECS) and social intranet is that ECS combines the functionality of the Enterprise Social Software (ESS) with the traditional groupware components to help the communication, collaboration, and sharing knowledge within the enterprise. Since ECS is used as an extended version of the intranet for the enterprise, it can be viewed as an extranet of the enterprise.

ECS is extremely helpful as a communication tool among employees that removes the physical barriers where an employee can work isolated not depending on the workplace [6]. The journal about the history of collaboration systems where collaboration system was implemented in 1995, where the function of the system was described as “A Method for Developing and Applying Metrics Profiles for the Benefits Management of Enterprise Collaboration Platforms” [9]. BRM gives the benefits of answering the quantitative metrics. With BRM, the data was processed into rich data for future studies, this indirectly benefitted scorecards and ECS efficiency [11]. The TDMS architecture was illustrated to form a collaboration system for the use of industry 4.0 focusing on the supply chain [13]. Separating the Small-Medium Enterprise (SME) into demand-driven, supplier-driven, and partner-driven cooperation, emphasized different methods used in different situation [14]. The purpose of the Identification of Requirements for Enterprise Social Software (IRESS) framework was to assist the company to find out the functions needed in software requirements [15] (see Table 1).

Table 1 Literature review

| No. | Article title | Advantages | Limitations |
|-----|---|--|--|
| 1 | “Service-oriented collaboration framework based on the cloud platform and critical factors identification” [4] (2021) | <ol style="list-style-type: none"> 1. Benefits in sharing information with the cost-effective solution provided by chain synergy 2. Connecting cloud platform to innumerably isolated information among company group members | <ol style="list-style-type: none"> 1. Coordination issues in the service value chain are not implemented in detail 2. Cloud platform required maintenance service quite frequently |
| 2 | “Professional Network Analytics Platform for Enterprise Collaboration” [5] (2017) | <ol style="list-style-type: none"> 1. Utilize the voluminous data stored in databases to connect the system in a large company | <ol style="list-style-type: none"> 1. Hierarchical levels are not involved in network visualization |
| 3 | “EPIC Framework for Enterprise processes integrative collaboration” [7] (2017) | <ol style="list-style-type: none"> 1. Addressing the concept of the gap to encapsulate the integrative holistic view in a complex system 2. Ensure the information safety issues with advanced integration testing based on V&V tests | <ol style="list-style-type: none"> 1. Insufficient explanation about the system architecture and operational processes are not discovered 2. Run-time issues required the user to install another system to connect to ESE |
| 4 | “Improving Inter-Enterprise Collaboration with Recommendation Tool based on Lightweight Semantics in Emails” [8] (2018) | <ol style="list-style-type: none"> 1. Conceptual solution on LWS where it allows the company system to extract data from email 2. Variety of searching methods to search for the accurate document | <ol style="list-style-type: none"> 1. The entities are created only based on the big data, whereas small data could not generate the entities 2. The search engine reaches only up to 71% of the accuracy |
| 5 | “Traces of design activity: the design of coordination mechanisms in the shaping of enterprise collaboration systems” [10] (2018) | <ol style="list-style-type: none"> 1. Coordinate mechanism technique is used to analyse the cooperative activities 2. Provide a new form of CM implemented in ECS to enhance the user-friendly interface | <ol style="list-style-type: none"> 1. Spatial facets of ECS are not discussed in this paper and the overtime issue may lead to system failure |
| 6 | “Social Collaboration Analytics Framework: A framework for providing business intelligence on collaboration in the digital workplace” [12] (2021) | <ol style="list-style-type: none"> 1. Describe in detail the eight phases of working step in a collaboration that uses SCA 2. A structural base of documentation is used in this SCA framework where visualization and interpretation are extremely useful | <ol style="list-style-type: none"> 1. Additional SCAF is not provided with those that does not work independently 2. SCA is limited for the use of matrices to receive information in deep interpret data mining and data analysis |

3 Results and Findings

3.1 *Enterprise Collaboration System Helps in the Enterprise*

An enterprise collaboration system is an information system whose function is to collect, process, store, and share information among individuals in an enterprise. Employees can easily access and update in real-time at any time. This can help employees communicate well and complete their tasks more effectively. Enterprise collaboration system can bring the following benefits to enterprises:

1. **Improve communication between employees.** The enterprise collaboration system has private messaging and communication technology to communicate with all employees, which is very important for employees. In terms of communication, employees do not need to use other social platforms to communicate, which can reduce the confusion that may arise when using different social platforms. For example, using multiple social platforms to communicate with colleagues at the same time can cause difficulties when looking for previous information.

In addition, since all employees are on the same platform, this can help employees communicate more easily with their private messaging services without having to obtain other employees' mobile phone numbers or social platforms. This can help employees in different departments to communicate at work without the need to communicate through a third party. Since the information may be incomplete or too much time spent when communicating through a third party, the enterprise collaboration system can help employees in different departments to communicate more efficiently.

When faced with a problem or in a discussion, the enterprise collaboration system can also help the enterprise to get more different ideas as a reference through the comments of the employees, which can effectively help the enterprise to be more comprehensive in deciding.

2. **Cooperation among employees.** The enterprise collaboration system allows multiple employees to perform one job at the same time. When the work is assigned, the information entered, deleted, or modified by the employees in the project will also be synchronized. This can help employees collaborate and make changes immediately when any problems are found. In addition, since all employees can simultaneously understand the work progress, this will also encourage mutual help and supervision between employees to ensure that the work can be completed smoothly.
3. **Simplify work.** Enterprise collaboration systems are all carried out using the network, so data damage or loss due to paper data can be reduced. This can also make it easier for employees to find past information, and the boss can also know the progress of the project and the work status of the employees without asking the employees.

System Limitation: Enterprise collaboration systems have brought a lot of benefits to many companies today, but no matter how good the system is, there will be some problems, for instance:

1. **Safety.** Since the enterprise collaboration system is a system with all employee accounts, this will cause a large amount of company information to be leaked when the employee's computer system is implanted with a virus. The enterprise collaboration system cannot prevent hackers from being attacked.
2. **System Consistency.** The enterprise collaboration system is a system used in the enterprise. Therefore, when the system is attacked or paralyzed or needs to be repaired, all employees will be completely unable to use all the stored data. Work with ECS and return to the traditional model of work, causes a lot of losses to the enterprise.
3. **Cost.** When the enterprise collaboration system has to be installed, the enterprise needs to spend a lot of money. In addition, the security of the system cannot be ignored; so the monthly maintenance fee will also bring a lot of expense to the enterprise.

3.2 Current Enterprise Collaboration System Challenge

The first condition of the system is to build a network between employees based on different types of data. These data represent different types of relationships and are stored in multiple databases in the enterprise. Different types of data are past work-related information such as the alumni of the educational institution, the work experience of the organization, and the data of the working magistrate. Such data cannot be used in the same way as recorded in these systems. Therefore, these ordinary meeting data and project allocation data can establish the relationship between employees and project social networks, but they cannot be used by the system at the same time because they are essentially different. They should need data to link these different contents. For the system to use these different contents, all the platforms can be converted into a unified format to ensure that the format of all data is unified for better identification, and the data can also be used by other components of the system. In the initial stage of the system, the three types of relationships between configuration files, matrices, and events are determined. And briefly describing the entity's data, the data of employees, projects, departments, and customers are all representatives of physical data. The data of the relationship between two entities is called a matrix. For example, employees work for projects, employees need to serve customers, and other events involve tracking data systems.

The second challenge of this system is the need to manage a large number of data feeds from different operational databases. The system needs to try to understand the data because the lack of data understanding will cause the system to not know the importance of the part of the data and will not back it up. The data may not be stored correctly. More importantly, the large data sources used by the system are dynamic and heterogeneous, and they cannot be stored or managed in any single

database management system. The system needs to meet this challenge, and a data service centre has to be built. This data centre is a heterogeneous database functional component with all the functions required by the system, so that other components can access data conveniently and quickly. In addition, the most advanced methods are applied to make the system distributed, where some functions work as microservices, and they can work together.

The third challenge of the system is that the results of various analyses should serve the interests of multiple users, including managers and employees of different roles and levels. Different users have different requirements and expectations for analysis results, ranging from helping users find and establish new contacts to form new project teams. To meet this challenge, the system needs to perform various calculations on the platform, including building professional networks of employees based on different data sources, analysing these networks, including the social distance between employees, and presenting them to various applications.

3.3 ECS Understands the Connections of Others and Visualize Abilities

Every employee knows something, and every employee has his social circle. If these data can be visualized, it will be helpful to the employees of the company. In the new system, the system's archives need to store the connections and capabilities of the entire enterprise, and everyone who uses it can learn from it. The file records in detail the employees, the colleagues he added to his network, the team he is currently in, the members of the team, the positions he has held, and the abilities of his colleagues.

3.4 ECS Find Contacts and the Shortest Path

Finding contacts is common but also a very important function. Employees can search for colleagues they might be interested in contacting. The search function in the new system allows employees to use various parameters to find information, such as specific customers, specific projects, capabilities, specific locations, and specific levels of work. This enables the employees of the company to quickly find the desired target and complete the current task as soon as possible. The shortest path function means that employees can use the shortest path to reach the target contact. The system will calculate the social distance between two employees to get an intuitive "shortest path view".

4 Conceptual Solution

4.1 Recommendations for New Connections

The main purpose of the recommendation function of the new system is to encourage employees to establish connections with others who may not necessarily be similar to them but can add important professional value to them. The recommendation engine recommends new contacts based on different parameters such as the employee's job, ability, location, role, current contact, or other preferences. In simple terms, the recommendation engine will make recommendations based on the three factors:

- (i) Personal profile data of employees
- (ii) Project social network data
- (iii) Log data.

In addition to recommending a set of recommended contacts, the recommendation system also shows the basic principles of these recommendations. When employees move to new projects, move to new locations, acquire new capabilities, and establish new relationships, their preferences for new contacts change. The system detects such changes and adjusts the recommendations accordingly [5].

Enterprise collaboration system based on Cloud. With the increasing competition in the market and the reduction of product interest rate space, companies need to shift their focus. It is a good choice to gradually move the business to higher value-added service areas. Nowadays, after-sales service is the source of differentiation and business opportunities. It can create 30% of the profit for the company that exceeds the product interest rate. Since the products and services provided by various companies are different, they can integrate these into a strategy to realize an after-sales service value chain, which includes suppliers, manufacturers, recyclers, repairers, etc., thereby gaining competitive advantage across the entire value chain.

The system based on platform services has proven to be efficient and cost-effective. In addition to the collaboration of a single enterprise's internal work, the ECS is also a system used by after-sales services such as product operation, maintenance, recycling, and remanufacturing. The system based on the cloud platform can realize the collaboration and interaction between the enterprises in the entire value chain and can use the cloud platform to create value.

Enterprise collaboration systems based on cloud platforms can bring three main functions in the value chain described above. The first is the virtualization and servicing of resources. Any enterprise that meets the access conditions is a potential service provider. They can provide their service resources and capabilities, virtualize and encapsulate them and publish them on the cloud platform to any users who need or are interested can get services from the platform. The second function is to match supply and demand. The cloud platform provides intelligent search and matching based on the needs of the identified or received users, and select the best supplier based on specific regulations and optimization. The third main function provided by

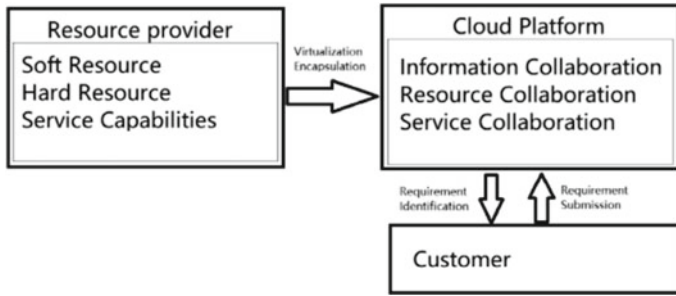


Fig. 2 Cloud platform service value chain

the system is also the basic function of the system, which is the implementation of collaborative services. Each supplier and its partners are customer-oriented together to provide a collaborative service [4] (see Fig. 2).

LightWeight Semantic (LWS) approach. No matter how, email is the first tool preferred by many organizations; but insufficient function such as recall/delete messages is not available and searching an old information will take a long time as the messages are not organized structure based. Semantic interoperability of LWS is used to take this obstacle based on the Virtual Enterprise Paradigm (VEP).

To be specific, LWS implement the tags/annotations (partition of pages/blogs) into data documentation. The tags can be created by the user or automatically generated. Most importantly, LWS can be linked to ECS from reacquiring legacy data. Considering the data protection issue, tags can be set to be invisible or removed; only the organization has the permissions to access. The name of the entities is processed into more human understanding; for example, searching “New York” by entering the keyword “NYC” or “NY”.

The objective is to utilize LWS in cooperation with email communication which can connect to the ECS used in the enterprise. This is an additional feature that supports businesses in searching documents by semantic search and semantic recommendations. Note that there is not any restriction to use LWS; users can use the system with their email on desktop, or even mobile devices.

Referring to the figure above, the attachment stripper interprets the messages from a mail server. On the other side, legacy stored in the ECS database is encapsulated by Adapter. The centre of the Virtual Common Repository (VCR) is to process the data received from the attachment stripper and adapter. Moving on to the indexer that contains two components, the document parser which responsible to check the format of the document (doc/pdf); whereas NER is to extract the information from the document parser. Throughout indexer, solr document and metadata are generated. Solr document is created for later use in searching purposes using Apache Solr; metadata is stored in a graph which is also known as a semantic network that stores data translated from indexer to develop LWS graph. Each document is then categorized as a document node which includes sentences or paragraph nodes, and each of the nodes is assigned by an entity’s presentation. Each entity is created only once

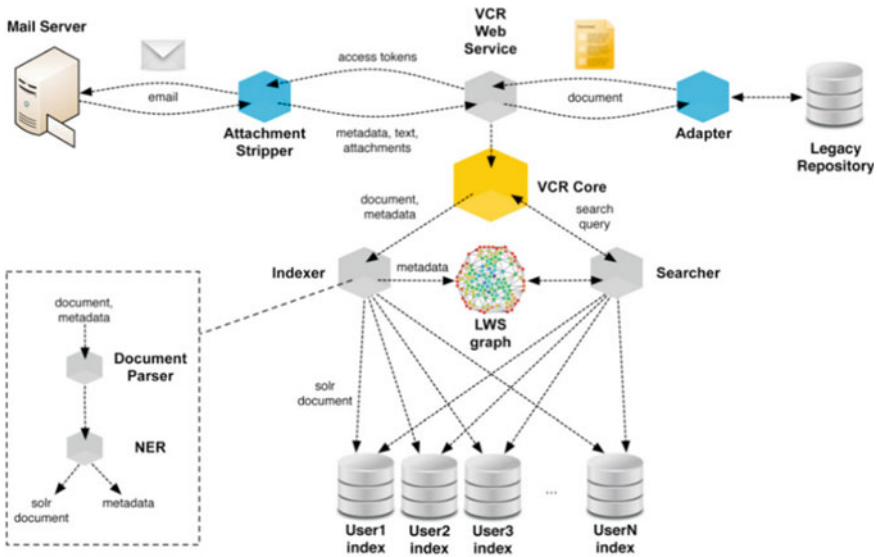


Fig. 3 LWS architecture

to categorize multiple documents to avoid duplication. Activity graph can track the document status whether is updated, deleted, or downloaded. Next, the searcher on its name is to search all the relevant documents. Searcher utilizes the data interpreted by Indexer either perform text search from user indexes or semantic search from LWS graph [8] (see Fig. 3).

Now it is known that not only the information is extracted from emails, but also attachments in the email such as video, document in excel, word, or PDF. The search results obtained can be described as the below descriptions. Entity search on its name is to search the information via the entities searching. One of the examples is that the user wishes to search on “Feedback on Transportation services”, then several entities may appear such as Transportation, Feedback, Complaint which are relevant to the searching requirement. Targeted entity search is to search the information which only focuses on one entity type. Multi-entity-based search is the combination of both entity search and targeted entity search. The search result is obtained from the semantic graph where the user can search different topics and different entities at the same time to check for the expected results [8].

5 Conclusion and Limitation

In the LWS approach, the user does not require to search information one by one from the email. With LWS, it does not require the organization to use the new system but to link the email to ECS. All the information in the email is transferred and

organized in LWS. Through the different search engine methods, user can easily find the most relevant information, which indirectly reduce the time spent on searching information as well as enhance the way of documenting information from the email.

As mentioned in [8], by using this technique, the percentage of receiving expected results which meet the user needs is 71% on average. However, this kind of searching required the user to understand the keyword of the entities. The entities will create only if there is huge information about the same topic, and small information available will be an issue in searching. There is a problem where some of the entity is not recognized by Entity Extraction tools. The percentage of assigning wrong entities to each tag may be a problem where users would not receive correct search results is 29%. Since this is a conceptual solution that is not implemented in real life, how it is connected in terms of architecture, hardware, and software used is not known.

6 Future Work

In the future, more about how it works in a business intelligence system might be discovered. Moreover, the connection issue on how each component is stated in LWS whether it is using a wireless, physical cable, or streamlining to the internet might be determined.

References

1. S. Prakash, S. Joshi, T. Bhatia, S. Sharma, D. Samadhiya, R. R. Shah, O. Kaiwartya and M. Prasad, "Characteristic of enterprise collaboration system and its implementation issues in business management," *International Journal of Business Intelligence and Data Mining*, vol. 16, no. 1, pp. 49–65, 2020.
2. K. Kolasinska-Morawska, L. Sulkowski and P. Morawski, "Agility in Customer Service Using Cloud-Based CRM Systems and Enterprise Collaboration Tools," *Economic and Social Development: Book of Proceedings (2017)*, pp. 72–81, 2017.
3. F. Schwade and P. Schubert, "Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities," *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp. 401–410, 2017.
4. X. Liu, Q. Deng, G. Gong, M. Lv, and C. Jiang, "Service-oriented collaboration framework based on the cloud platform and critical factors identification," *Journal of Manufacturing Systems*, vol. 61, no. 13, pp. 183–195, Oct. 2021, doi: <https://doi.org/10.1016/j.jmsy.2021.09.007>.
5. G. Gopalakrishnan, K. Benhur, A. Kaushik, and A. Passala, "Professional Network Analytics Platform for Enterprise Collaboration," *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, vol. 4, no. 4, Feb. 2017, doi: <https://doi.org/10.1145/3022198.3023264>.
6. N. Koceska and S. Koceski, "THE IMPORTANCE OF ENTERPRISE COLLABORATION SYSTEMS DURING A PANDEMIC," 2020. Accessed: Nov. 06, 2021. [Online]. Available: <http://www.aebjournal.org/articles/0804/080403.pdf>.

7. M. Sitton and Y. Reich, "EPIC framework for enterprise processes integrative collaboration," *Systems Engineering*, vol. 21, no. 1, pp. 30–46, Jan. 2018, doi: <https://doi.org/10.1002/sys.21417>.
8. M. Šeleng, Š. Dlugolinský, L. Hluchý, and W. Gräther, "Improving Inter-Enterprise Collaboration with Recommendation Tool based on Lightweight Semantics in Emails," *Procedia Computer Science*, vol. 138, no. 6, pp. 486–491, 2018, doi: <https://doi.org/10.1016/j.procs.2018.10.067>.
9. K. Smolander, M. Rossi, and S. Pekkola, "Heroes, contracts, cooperation, and processes: Changes in collaboration in a large enterprise systems project," *Information & Management*, vol. 58, no. 2, p. 103407, Mar. 2021, doi: <https://doi.org/10.1016/j.im.2020.103407>.
10. C. S. Nitschke and S. P. Williams, "Traces of design activity: the design of coordination mechanisms in the shaping of enterprise collaboration systems," *Procedia Computer Science*, vol. 138, no. 7, pp. 580–586, 2018, doi: <https://doi.org/10.1016/j.procs.2018.10.078>.
11. S. Grams, F. Schwade, J. Mosen, and P. Schubert, "A Method for Developing and Applying Metrics Profiles for the Benefits Management of Enterprise Collaboration Platforms," *Procedia Computer Science*, vol. 181, no. 181, pp. 553–561, Jan. 2021, doi: <https://doi.org/10.1016/j.procs.2021.01.202>.
12. F. Schwade, "Social Collaboration Analytics Framework: A framework for providing business intelligence on collaboration in the digital workplace," *Decision Support Systems*, vol. 148, p. 113587, May 2021, doi: <https://doi.org/10.1016/j.dss.2021.113587>.
13. JosephNg, P. S., & Eaw, H. (2021). Making Financial Sense from EaaS for MSE during Economic Uncertainty. *Future of Information and Communication Conference* (pp. 976–989). Vancouver, Canada: Springer Advances in Intelligent Systems and Computing.
14. JosephNg, P. S. (2021). Economic Turbulence and EaaS Grid Computing. *Int. J. of Business Forecasting and Marketing Intelligence*, 7(1), 33–52
15. JosephNg, P. S., & Eaw, H. C. (2022). Still Technology Acceptance Model? Reborn with Exostructure as a Service. *International Journal of Business Information Systems*, forthcoming.
16. JosephNg, P. S. (2019). EaaS Infrastructure Disruptor for MSE. *International Journal of Business Information System*, 30(3), 373–385.
17. JosephNg, P. S. (2018). EaaS Optimization: Available yet hidden information technology infrastructure inside the medium-size enterprise. *Technological Forecasting and Social Change*, 132(July), 165 - 173.
18. JosephNg, P. S., & Kang, C. M. (2016). Beyond Barebones Cloud Infrastructure Services: Stumbling Competitiveness During Economic Turbulence. *Journal of Science & Technology*, 24(1), 101–121.
19. JosephNg Poh Soon, Kang Chon Moy, Ahmad Kamil Mahmood, Wong See Wan, Phan Koo Yuen, Saw Seow Hui, Lim Jit Theam (2016), EaaS: Available yet Hidden Infrastructure inside MSE, 5th International Conference on Network, Communication, and Computing, Kyoto, Japan, ACM International Conference Proceeding Series, 17–20.
20. PS, J. N., Kang, C. M., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. (2016). Exostructure Services for Infrastructure Resources Optimization. *Journal of Telecommunication, Electronic and Computer Engineering*, 8(4), 65–69.
21. JNP Soon, WS Wan, PK Yuean, LE Heng & LJ Theam (2015). Barebone Cloud IaaS: Revitalisation Disruptive Technology, *International Journal of Business Information Systems*, 18(1), 107–126.
22. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. H. (2014). IaaS cloud optimisation during economic turbulence for Malaysia small and medium enterprises. *International Journal of Business Information Systems*, 16(2), 196–208.
23. Joseph, N. P.S., Mahmood, A. K., Choo, P. Y., Wong, S. W., Phan, K. Y., & Lim, E. H. (2013). Battles in volatile information and communication technology landscape: The Malaysia small and medium-size enterprise case. *International Journal of Business Information Systems*, 16(2), 196–208.

24. S. Cisneros-Cabrera, G. Pishchulov, P. Sampaio, N. Mehandjiev, Z. Liu, and S. Kununka, "An approach and decision support tool for forming Industry 4.0 supply chain collaborations," *Computers in Industry*, vol. 125, no. 125, p. 103391, Feb. 2021, doi: <https://doi.org/10.1016/j.compind.2020.103391>.
25. T. Brink, "SME routes for innovation collaboration with larger enterprises," *Industrial Marketing Management*, vol. 64, no. 64, pp. 122–134, Jul. 2017, doi: <https://doi.org/10.1016/j.indmarman.2017.01.010>.
26. J. Glitsch and P. Schubert, "IRESS: Identification of Requirements for Enterprise Social Software," *Procedia Computer Science*, vol. 121, no. 121, pp. 866–873, Jan. 2017, doi: <https://doi.org/10.1016/j.procs.2017.11.112>.

Author Index

A

Abhyankar, Anant, 493
Ângelo Lellis Moreira, Miguel, 521
Adekeye, Deji Babatunde, 319
Adeniji, Oluwashola David, 319, 725
Adeoye, Adeyinka, 697
Adesina, Ademola Olusola, 319, 725
Adetunla, Adedotun, 697
Adewolu, Joshua, 697
Afolabi, Olakunle Sunday, 725
Aggarwal, Vishesh, 673
Ahire, Mahesh, 493
Ahuja, Harjis, 65
Ajagbe, Sunday Adeola, 319, 725
Ajimobi, Olalekan Ibrahim, 725
Akilandeswari, J., 389
Al-Maari, A. A. A., 377
Al-Sofi, S. M., 451
Almujally, Nouf, 463
Alrashidi, Huda, 463
Amisha, 309
Ananth, P., 53
Annapurna, K., 633
Antón-Sancho, Álvaro, 139
Aphale, Ishan, 571
Araújo Costa de, Igor Pinheiro, 219
Aremu, Bashiru, 23
Arumugam, G., 353
Arumugaperumal, S., 421
Arun Kumar, U., 659

B

Balachandra, Mamatha, 759
Banothu, Nagateja, 709

Bapat, Pranav, 571
Bapna, Ruchita, 65
Basumatary, Thulunga, 547
Bhase, Gargee, 65
Bhatlawande, Shripad, 493, 509
Bhuvanewari, A., 407
Bijane, Bhumika, 509
Bisne, Chaitanya, 77
Bremm De Carvalho, Eric, 521
Busygin, Volodymyr, 205

C

Calatayud, David G., 139
Chadgal, Ankush, 493
Charde, Geetai, 509
Chauhan, Prithviraj, 151
Chine, Abhijit, 509
Costa, Igor Pinheiro de Araújo, 245
Cristin, R., 295

D

Daniel Ullmann, Thomas, 463
Daniya, T., 295
Darekar, Vidya, 77
Dayana, R., 585
Deepika, J., 389
Desai, Amogh, 437
Deshmukh, Amberish, 77
Deshmukh, Chinmay, 309
Deshpande, Atharva, 571
Devaneyan, S. Pradeep, 169
Doriya, Rajesh, 39

E

Ekbote, Ninad, 571
 Ezhil Roja, P., 479

F

Fernández-Arias, Pablo, 139
 Fernando, Xavier, 23
 Francisco Simões Gomes, Carlos, 521

G

Gaikwad, Kunal, 309
 Gajbhiye, Akash, 309
 Glory Thephoral, J., 345
 Gobhinath, S., 659
 Gomes, Carlos Francisco Simões, 219, 245
 Gupta, Neetesh, 363

H

Hebbar, Harishchandra, 759
 Hinge, Rutuja, 77

I

Ibrokhimov, Aziz, 563
 Ige, Olubunmi, 697

J

Jain, Saurabh, 39
 Jain, Vishesh, 673
 Jana, Radha Krishna, 601
 Jayashri, N., 1
 Jeya Pandian, M., 345
 Jindal, Rajni, 673
 JosephNg, P. S., 377, 451, 803
 Joy, Mike, 463
 Jyothi, C., 95

K

Kadhun, Methaq, 463
 Karthick Kumar, A., 585
 Karthik, Durga, 345
 Karthika, D., 1
 Karumanchi, Mani Deep, 169
 Keerthana, B., 759
 Khandare, Hrishikesh, 39
 Khaydarov, Elshod, 197
 Khylyko, Maksym, 205
 Kozenkova, Vladyslava, 205
 Krishnendhu, S. P., 547

Kshirsagar, Ankita, 363
 Kurian, Nitheesh, 333

L

Lai, S. C., 451
 Lim, E. H., 377
 Lim, J. T., 377, 451

M

Mamier, Fernando Benedicto, 245
 Madake, Jyoti, 493, 509
 Mahendran, G., 659
 Mahesha, Y., 9
 Maity, Saikat, 601
 Malarvezhi, P., 585
 Mane, Vijay, 309
 Manivannan, R., 789
 Matviichuk, Andrii, 205
 Misbha, D. S., 479
 Mohandas, Prabu, 547
 Moreira, Érick Pinto, 245
 Moroz, Boris, 205
 Mubarak, D. Muhammad Noorul, 621
 Muniyal, Balachandra, 759

N

Nagaraju, C., 9
 Nandhini, M., 407
 Nisha Jenipher, V., 53
 Nithya, K., 53

O

Oguns, Yetunde Josephine, 319
 Oladipupo, Matthew Abiola, 319

P

Palwankar, Shrinath, 151
 Parekh, Virang, 437
 Pathan, Mahabub Subhani, 633
 Patil, Sarang, 151
 Pham, Van-Truong, 533
 Phan, K. Y., 377, 451
 Phillips, Celsy, 151
 Pimperkhede, Sameer, 571
 Prabu, M., 709
 Pradeep, D., 407
 Preethi, M., 421
 Pushpa Rani, M., 23

R

Rashmi, M. R., 743
Revathy, G., 345
Roshini Begum, A., 407
Roy, Archisman, 685

S

Santha Subbulaxmi, S., 353
Santos dos, Marcos, 219, 245, 521
Santos Hermogenes dos, Lucas Ramon, 219
Sasikumar, P., 123
Sathiamoorthy, S., 123
See, S. L., 803
Sheeba, Adlin, 53
Sheeba, J. I., 169
Shekokar, Narendra, 65, 437
Shilaskar, Swati, 493, 509
Shvachych, Gennady, 205
Soundharya, S. P., 743
Sreedevi, B., 345
Sreeja, S., 621
Sriram, V., 673
Sumera, 789
Supriya, M., 95
Surya, U., 53
Swetha, N., 407

T

Tan, P. H., 803
Taralekar, B. G., 151
Taralekar, Bharat, 77

Telles, Charles Roberto, 685
Tran, Thi-Thao, 533
Trinh, Minh-Nhat, 533
Tripathi, Anushka, 743

U

Unadkat, Utsav, 437

V

Vadamodula, Prasad, 295
Vadivukkarasi, K., 333, 585
Vaidehi, K., 789
Velayutham, C., 421
Vergara-Rodríguez, Diego, 139
Verma, Bhupendra, 363
Vigneshwar, R., 743
Vigneshwari Reddy, Mainampati, 547
Vilarinho Terra, Adilson, 521
Vinora, A., 53

W

Williams, Andy E., 233
Wong, J. H., 803
Wyawahare, Medha, 571

Y

Yeo, R. C., 803
Yong, Y. H., 803